



US012390514B2

(12) **United States Patent**
Maianti et al.

(10) **Patent No.:** US 12,390,514 B2
(b4) **Date of Patent:** Aug. 19, 2025

(54) **CANCER VACCINE**(71) Applicant: **President and Fellows of Harvard College**, Cambridge, MA (US)(72) Inventors: **Juan Pablo Maianti**, Revere, MA (US); **David R. Liu**, Lexington, MA (US)(73) Assignee: **President and Fellows of Harvard College**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 876 days.

(21) Appl. No.: **16/492,534**(22) PCT Filed: **Mar. 9, 2018**(86) PCT No.: **PCT/US2018/021880**

§ 371 (c)(1),

(2) Date: **Sep. 9, 2019**(87) PCT Pub. No.: **WO2018/165631**PCT Pub. Date: **Sep. 13, 2018**(65) **Prior Publication Data**

US 2021/0196809 A1 Jul. 1, 2021

Related U.S. Application Data

(60) Provisional application No. 62/469,219, filed on Mar. 9, 2017.

(51) **Int. Cl.**

<i>A61K 38/50</i>	(2006.01)
<i>A61K 31/7088</i>	(2006.01)
<i>A61K 38/46</i>	(2006.01)
<i>A61K 39/00</i>	(2006.01)
<i>A61K 45/06</i>	(2006.01)

(52) **U.S. Cl.**

CPC	<i>A61K 38/50</i> (2013.01); <i>A61K 31/7088</i> (2013.01); <i>A61K 38/465</i> (2013.01); <i>A61K 39/0011</i> (2013.01); <i>A61K 39/001106</i> (2018.08); <i>A61K 39/001113</i> (2018.08); <i>A61K 39/001115</i> (2018.08); <i>A61K 39/001156</i> (2018.08); <i>A61K 39/001157</i> (2018.08); <i>A61K 39/001182</i> (2018.08); <i>A61K 39/001188</i> (2018.08); <i>A61K 39/001191</i> (2018.08); <i>A61K 39/001192</i> (2018.08); <i>A61K 45/06</i> (2013.01); <i>A61K 39/001149</i> (2018.08); <i>A61K 2039/5154</i> (2013.01); <i>A61K 2039/53</i> (2013.01); <i>A61K 2039/58</i> (2013.01); <i>C07K 2319/00</i> (2013.01); <i>C12Y 304/22062</i> (2013.01)
-----------	---

(58) **Field of Classification Search**

None

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,182,449 A	1/1980	Kozlow
4,186,183 A	1/1980	Steck et al.
4,217,344 A	8/1980	Vanlerberghe et al.
4,235,871 A	11/1980	Papahadjopoulos et al.
4,261,975 A	4/1981	Fullerton et al.
4,485,054 A	11/1984	Mezei et al.
4,501,728 A	2/1985	Geho et al.
4,663,290 A	5/1987	Weis et al.
4,737,323 A	4/1988	Martin et al.
4,774,085 A	9/1988	Fidler
4,797,368 A	1/1989	Carter et al.
4,837,028 A	6/1989	Allen
4,873,316 A	10/1989	Meade et al.
4,880,635 A	11/1989	Janoff et al.
4,889,818 A	12/1989	Gelfand et al.
4,897,355 A	1/1990	Eppstein et al.
4,906,477 A	3/1990	Kurono et al.
4,911,928 A	3/1990	Wallach
4,917,951 A	4/1990	Wallach
4,920,016 A	4/1990	Allen et al.
4,921,757 A	5/1990	Wheatley et al.
4,946,787 A	8/1990	Eppstein et al.
4,965,185 A	10/1990	Grischenko et al.
5,017,492 A	5/1991	Kotewicz et al.
5,047,342 A	9/1991	Chatterjee
5,049,386 A	9/1991	Eppstein et al.
5,079,352 A	1/1992	Gelfand et al.
5,139,941 A	8/1992	Muzyczka et al.
5,173,414 A	12/1992	Lebkowski et al.

(Continued)

FOREIGN PATENT DOCUMENTS

AU	2012244264 A1	11/2012
AU	2012354062 A1	7/2014

(Continued)

OTHER PUBLICATIONS

Mahoney et al., "The Next Immune-Checkpoint Inhibitors: PD-1/PD-L1 Blockade in Melanoma" 37(4) Clinical Therapeutics 764-782 (Year: 2015).*

(Continued)

Primary Examiner — Nancy J Leith

(74) Attorney, Agent, or Firm — Wolf, Greenfield & Sacks, P.C.

(57) **ABSTRACT**

Provided herein are systems, compositions, and methods for generating immunogenic peptides or epitopes from tumor associated antigens (e.g., *in vivo* or *ex vivo*). Polynucleotides (e.g., genes) encoding the tumor associated antigens may be edited at selected target sites by nucleobase editors comprising a catalytically-inactive Cas9 and a cytosine deaminase, leading to the expression of heteroclitic or cryptic peptides that are more immunogenic than the native peptide derived from the tumor associated antigens. The heteroclitic or cryptic peptide elicit strong tumor-specific immune response (e.g., T-cell response or B-cell response), which inhibits tumor growth and metastasis.

23 Claims, 5 Drawing Sheets

Specification includes a Sequence Listing.

US 12,390,514 B2

Page 2

(56)

References Cited

U.S. PATENT DOCUMENTS

5,223,409 A	6/1993	Ladner et al.	7,595,179 B2	9/2009	Chen et al.
5,244,797 A	9/1993	Kotewicz et al.	7,638,300 B2	12/2009	Schultz et al.
5,270,179 A	12/1993	Chatterjee	7,670,807 B2	3/2010	Lampson et al.
5,374,553 A	12/1994	Gelfand et al.	7,678,554 B2	3/2010	Liu et al.
5,405,776 A	4/1995	Kotewicz et al.	7,713,721 B2	5/2010	Schultz et al.
5,436,149 A	7/1995	Barnes	7,771,935 B2	8/2010	Liu et al.
5,449,639 A	9/1995	Wei et al.	7,794,931 B2	9/2010	Breaker et al.
5,496,714 A	3/1996	Comb et al.	7,807,408 B2	10/2010	Liu et al.
5,512,462 A	4/1996	Cheng	7,851,658 B2	12/2010	Liu et al.
5,580,737 A	12/1996	Polisky et al.	7,915,025 B2	3/2011	Schultz et al.
5,614,365 A	3/1997	Tabor et al.	7,919,277 B2	4/2011	Russell et al.
5,652,094 A	7/1997	Usman et al.	7,993,672 B2	8/2011	Huang et al.
5,658,727 A	8/1997	Barbas et al.	7,998,904 B2	8/2011	Liu et al.
5,668,005 A	9/1997	Kotewicz et al.	7,999,071 B2	8/2011	Schlom et al.
5,677,152 A	10/1997	Birch et al.	8,012,739 B2	9/2011	Schultz et al.
5,767,099 A	6/1998	Harris et al.	8,017,323 B2	9/2011	Liu et al.
5,780,053 A	7/1998	Ashley et al.	8,017,755 B2	9/2011	Liu et al.
5,830,430 A	11/1998	Unger et al.	8,030,074 B2	10/2011	Schultz et al.
5,834,247 A	11/1998	Comb et al.	8,067,556 B2	11/2011	Hogrefe et al.
5,835,699 A	11/1998	Kimura	8,114,648 B2	2/2012	Schultz et al.
5,844,075 A *	12/1998	Kawakami	8,173,364 B2	5/2012	Schultz et al.
		A61P 35/00	8,173,392 B2	5/2012	Schultz et al.
		530/328	8,183,012 B2	5/2012	Schultz et al.
			8,183,178 B2	5/2012	Liu et al.
			8,206,914 B2	6/2012	Liu et al.
			8,354,380 B2	1/2013	Liu et al.
5,849,548 A	12/1998	Haseloff et al.	8,361,725 B2	1/2013	Russell et al.
5,851,548 A	12/1998	Dattagupta et al.	8,394,604 B2	3/2013	Liu et al.
5,855,910 A	1/1999	Ashley et al.	8,420,104 B2	4/2013	Charneau et al.
5,856,463 A	1/1999	Blankenborg et al.	8,440,431 B2	5/2013	Voytas et al.
5,962,313 A	10/1999	Podsakoff et al.	8,440,432 B2	5/2013	Voytas et al.
5,981,182 A	11/1999	Jacobs, Jr. et al.	8,450,471 B2	5/2013	Voytas et al.
6,015,794 A	1/2000	Haseloff et al.	8,492,082 B2	7/2013	De Franciscis et al.
6,057,153 A	5/2000	George et al.	8,546,553 B2	10/2013	Terns et al.
6,063,608 A	5/2000	Kotewicz et al.	8,569,256 B2	10/2013	Heyes et al.
6,077,705 A	6/2000	Duan et al.	8,586,363 B2	11/2013	Voytas et al.
6,099,857 A	8/2000	Gross	8,673,612 B2	3/2014	Klatzmann et al.
6,156,509 A	12/2000	Schellenberger	8,680,069 B2	3/2014	de Fougerolles et al.
6,183,998 B1	2/2001	Ivanov et al.	8,691,729 B2	4/2014	Liu et al.
6,355,415 B1	3/2002	Wagner et al.	8,691,750 B2	4/2014	Constien et al.
6,416,997 B1	7/2002	Mir-Shekari et al.	8,697,359 B1	4/2014	Zhang
6,429,298 B1	8/2002	Ellington et al.	8,697,439 B2	4/2014	Mangeot et al.
6,453,242 B1	9/2002	Eisenberg et al.	8,697,853 B2	4/2014	Voytas et al.
6,479,264 B1	11/2002	Louwrier	8,709,466 B2	4/2014	Coady et al.
6,503,717 B2	1/2003	Case et al.	8,728,526 B2	5/2014	Heller
6,534,261 B1	3/2003	Cox, III et al.	8,729,038 B2	5/2014	Gruber et al.
6,558,671 B1	5/2003	Slingluff et al.	8,741,279 B2	6/2014	Kasahara et al.
6,589,768 B1	7/2003	Kotewicz et al.	8,748,667 B2	6/2014	Budzik et al.
6,599,692 B1	7/2003	Case et al.	8,758,810 B2	6/2014	Okada et al.
6,607,882 B1	8/2003	Cox, III et al.	8,759,103 B2	6/2014	Kim et al.
6,610,522 B1	8/2003	Kotewicz et al.	8,759,104 B2	6/2014	Unciti-Broceta et al.
6,689,558 B2	2/2004	Case	8,771,728 B2	7/2014	Huang et al.
6,716,973 B2	4/2004	Baskerville et al.	8,790,664 B2	7/2014	Pitard et al.
6,824,978 B1	11/2004	Cox, III et al.	8,795,965 B2	8/2014	Zhang
6,933,113 B2	8/2005	Case et al.	8,822,663 B2	9/2014	Schrum et al.
6,979,539 B2	12/2005	Cox, III et al.	8,835,148 B2	9/2014	Janulaitis et al.
7,013,219 B2	3/2006	Case et al.	8,846,578 B2	9/2014	McCray et al.
7,045,337 B2	5/2006	Schultz et al.	8,871,445 B2	10/2014	Cong et al.
7,067,650 B1	6/2006	Tanaka	8,889,418 B2	11/2014	Zhang et al.
7,070,928 B2	7/2006	Liu et al.	8,900,814 B2	12/2014	Yasukawa et al.
7,078,208 B2	7/2006	Smith et al.	8,945,839 B2	2/2015	Zhang
7,083,970 B2	8/2006	Schultz et al.	8,975,232 B2	3/2015	Liu et al.
7,163,824 B2	1/2007	Cox, III et al.	8,993,233 B2	3/2015	Zhang et al.
7,192,739 B2	3/2007	Liu et al.	8,999,641 B2	4/2015	Zhang et al.
7,223,545 B2	5/2007	Liu et al.	9,023,594 B2	5/2015	Liu et al.
7,329,807 B2	2/2008	Vadrucci et al.	9,023,649 B2	5/2015	Mali et al.
7,354,761 B2	4/2008	Schultz et al.	9,034,650 B2	5/2015	Padidam
7,368,275 B2	5/2008	Schultz et al.	9,068,179 B1	6/2015	Liu et al.
7,419,669 B2	9/2008	Kosmatopoulos et al.	9,150,626 B2	10/2015	Liu et al.
7,442,160 B2	10/2008	Liu et al.	9,163,271 B2	10/2015	Schultz et al.
7,476,500 B1	1/2009	Liu et al.	9,163,284 B2	10/2015	Liu et al.
7,476,734 B2	1/2009	Liu	9,181,535 B2	11/2015	Liu et al.
7,479,573 B2	1/2009	Chu et al.	9,200,045 B2	12/2015	Liu et al.
7,488,718 B2	2/2009	Scheinberg et al.	9,221,886 B2	12/2015	Liu et al.
7,491,494 B2	2/2009	Liu et al.	9,228,207 B2	1/2016	Liu et al.
7,510,706 B2	3/2009	Yonemitsu et al.	9,234,213 B2	1/2016	Wu
7,541,450 B2	6/2009	Liu et al.	9,243,038 B2	1/2016	Liu et al.
7,556,940 B2	7/2009	Galarza et al.	9,267,127 B2	2/2016	Liu et al.
7,557,068 B2	7/2009	Liu et al.	9,290,773 B2	3/2016	Edgerton

(56)	References Cited					
U.S. PATENT DOCUMENTS						
9,296,790 B2	3/2016 Chatterjee et al.	11,053,481 B2	7/2021 Liu et al.			
9,322,006 B2	4/2016 Liu et al.	11,124,782 B2	9/2021 Liu et al.			
9,322,037 B2	4/2016 Liu et al.	11,214,780 B2 *	1/2022 Liu	A61P 17/00		
9,340,799 B2	5/2016 Liu et al.	11,319,532 B2	5/2022 Liu et al.			
9,340,800 B2	5/2016 Liu et al.	11,421,016 B2	8/2022 Nguyen et al.			
9,359,599 B2	6/2016 Liu et al.	11,447,770 B1	9/2022 Liu et al.			
9,388,430 B2	7/2016 Liu et al.	11,542,496 B2	1/2023 Liu et al.			
9,394,537 B2	7/2016 Liu et al.	11,542,509 B2	1/2023 Maianti et al.			
9,434,774 B2	9/2016 Liu et al.	11,560,566 B2	1/2023 Liu et al.			
9,458,484 B2	10/2016 Ma et al.	11,578,343 B2	2/2023 Liu et al.			
9,512,446 B1	12/2016 Joung et al.	11,643,652 B2	5/2023 Liu et al.			
9,526,784 B2	12/2016 Liu et al.	11,661,590 B2	5/2023 Liu et al.			
9,534,210 B2	1/2017 Park et al.	11,702,651 B2	7/2023 Liu et al.			
9,580,698 B1	2/2017 Xu et al.	11,732,274 B2	8/2023 Liu et al.			
9,593,356 B2	3/2017 Haugwitz et al.	11,795,443 B2	10/2023 Liu et al.			
9,610,322 B2	4/2017 Liu et al.	11,795,452 B2	10/2023 Liu et al.			
9,637,739 B2	5/2017 Siksnys et al.	11,820,969 B2	11/2023 Maianti et al.			
9,663,770 B2	5/2017 Rogers et al.	11,898,179 B2	2/2024 Maianti et al.			
9,695,446 B2	7/2017 Mangeot et al.	11,912,985 B2	2/2024 Liu et al.			
9,737,604 B2	8/2017 Liu et al.	11,920,181 B2	3/2024 Liu et al.			
9,738,693 B2	8/2017 Telford et al.	11,932,884 B2	3/2024 Liu et al.			
9,753,340 B2	9/2017 Saitou	11,999,947 B2	6/2024 Liu et al.			
9,765,304 B2	9/2017 Klatzmann et al.	12,006,520 B2	6/2024 Liu et al.			
9,771,574 B2	9/2017 Liu et al.	12,031,126 B2	7/2024 Liu et al.			
9,777,043 B2	10/2017 Anderson et al.	12,043,852 B2	7/2024 Liu et al.			
9,783,791 B2	10/2017 Hogrefe et al.	12,084,663 B2	9/2024 Maianti et al.			
9,816,093 B1	11/2017 Donohoue et al.	12,157,760 B2	12/2024 Liu et al.			
9,840,538 B2	12/2017 Telford et al.	12,215,365 B2	2/2025 Liu et al.			
9,840,690 B2	12/2017 Karli et al.	2003/0082575 A1	5/2003 Schultz et al.			
9,840,699 B2	12/2017 Liu et al.	2003/0087817 A1	5/2003 Cox et al.			
9,840,702 B2	12/2017 Collingwood et al.	2003/0096337 A1	5/2003 Hillman et al.			
9,850,521 B2	12/2017 Braman et al.	2003/0108885 A1	6/2003 Schultz et al.			
9,873,907 B2	1/2018 Zeiner et al.	2003/0119764 A1	6/2003 Loeb et al.			
9,879,270 B2	1/2018 Hittinger et al.	2003/0167533 A1	9/2003 Yadav et al.			
9,914,939 B2	3/2018 Church et al.	2003/0203480 A1	10/2003 Kovacs et al.			
9,932,567 B1	4/2018 Xu et al.	2004/0003420 A1	1/2004 Kuhn et al.			
9,938,288 B1	4/2018 Kishi et al.	2004/0028687 A1	2/2004 Waelti			
9,944,933 B2	4/2018 Storici et al.	2004/0115184 A1	6/2004 Smith et al.			
9,982,279 B1	5/2018 Gill et al.	2004/0158661 A1	8/2004 Figdor et al.			
9,999,671 B2	6/2018 Liu et al.	2004/0197892 A1	10/2004 Moore et al.			
10,011,868 B2	7/2018 Liu et al.	2004/0203109 A1	10/2004 Lal et al.			
10,040,830 B2	8/2018 Chatterjee et al.	2005/0136429 A1	6/2005 Guarante et al.			
10,053,725 B2	8/2018 Liu et al.	2005/0222030 A1	10/2005 Allison			
10,059,940 B2	8/2018 Zhong	2005/0260626 A1	11/2005 Lorens et al.			
10,077,453 B2	9/2018 Liu et al.	2006/008864 A1	4/2006 Smolke et al.			
10,113,163 B2	10/2018 Liu et al.	2006/0104984 A1	5/2006 Littlefield et al.			
10,150,955 B2	12/2018 Lambowitz et al.	2006/0216702 A1	9/2006 Compans et al.			
10,167,457 B2	1/2019 Liu et al.	2006/0246568 A1	11/2006 Honjo et al.			
10,179,911 B2	1/2019 Liu et al.	2007/0015238 A1	1/2007 Snyder et al.			
10,189,831 B2	1/2019 Arrington et al.	2007/0049533 A1	3/2007 Liu et al.			
10,202,593 B2	2/2019 Liu et al.	2007/0264692 A1	11/2007 Liu et al.			
10,202,658 B2	2/2019 Parkin et al.	2007/0269817 A1	11/2007 Shapero			
10,227,581 B2	3/2019 Liu et al.	2007/0298118 A1	12/2007 Lotvall et al.			
10,323,236 B2	6/2019 Liu et al.	2008/0008697 A1	1/2008 Mintier et al.			
10,336,997 B2	7/2019 Liu et al.	2008/0051317 A1	2/2008 Church et al.			
10,358,670 B2	7/2019 Janulaitis et al.	2008/0124725 A1	5/2008 Barrangou et al.			
10,392,674 B2	8/2019 Liu et al.	2008/0182254 A1	7/2008 Hall et al.			
10,407,474 B2	9/2019 Liu et al.	2008/0220502 A1	9/2008 Schellenberger et al.			
10,407,695 B2	9/2019 Charneau et al.	2008/0241917 A1	10/2008 Akite et al.			
10,407,697 B2	9/2019 Doudna et al.	2008/0268516 A1	10/2008 Perreault et al.			
10,465,176 B2	11/2019 Liu et al.	2009/0111119 A1	4/2009 Doyon et al.			
10,508,298 B2	12/2019 Liu et al.	2009/0130718 A1	5/2009 Short			
10,583,201 B2	3/2020 Chen et al.	2009/0202622 A1	8/2009 Fleury et al.			
10,597,679 B2	3/2020 Liu et al.	2009/0215878 A1	8/2009 Tan et al.			
10,612,011 B2	4/2020 Liu et al.	2009/0234109 A1	9/2009 Han et al.			
10,640,767 B2	5/2020 Maianti et al.	2010/0076057 A1	3/2010 Sontheimer et al.			
10,682,410 B2	6/2020 Liu et al.	2010/0093617 A1	4/2010 Barrangou et al.			
10,704,062 B2	7/2020 Liu et al.	2010/0104690 A1	4/2010 Barrangou et al.			
10,745,677 B2	8/2020 Maianti et al.	2010/0105134 A1	4/2010 Quay et al.			
10,858,639 B2	12/2020 Liu et al.	2010/0273857 A1	10/2010 Thakker et al.			
10,912,833 B2	2/2021 Liu et al.	2010/0305197 A1	12/2010 Che			
10,930,367 B2	2/2021 Zhang et al.	2010/0316643 A1	12/2010 Eckert et al.			
10,947,530 B2	3/2021 Liu et al.	2011/0016540 A1	1/2011 Weinstein et al.			
10,954,548 B2	3/2021 Liu et al.	2011/0059160 A1	3/2011 Essner et al.			
11,046,948 B2	6/2021 Liu et al.	2011/0059502 A1	3/2011 Chalasani			
		2011/0104787 A1	5/2011 Church et al.			
		2011/0123509 A1	5/2011 Jantz et al.			
		2011/0177495 A1	7/2011 Liu et al.			
		2011/0189775 A1	8/2011 Ainley et al.			

(56)	References Cited				
U.S. PATENT DOCUMENTS					
2011/0189776 A1	8/2011 Terns et al.	2015/0071906 A1	3/2015 Liu et al.		
2011/0206672 A1	8/2011 Little	2015/0079680 A1	3/2015 Bradley et al.		
2011/0217739 A1	9/2011 Terns et al.	2015/0079681 A1	3/2015 Zhang		
2011/0301073 A1	12/2011 Gregory et al.	2015/0098954 A1	4/2015 Hyde et al.		
2012/0129759 A1	5/2012 Liu et al.	2015/0118216 A1	4/2015 Liu et al.		
2012/0141523 A1	6/2012 Castado et al.	2015/0128300 A1	5/2015 Warming et al.		
2012/0159653 A1	6/2012 Weinstein et al.	2015/0132269 A1	5/2015 Orkin et al.		
2012/0244601 A1	9/2012 Bertozzi et al.	2015/0140664 A1	5/2015 Byrne et al.		
2012/0270273 A1	10/2012 Zhang et al.	2015/0159172 A1	6/2015 Miller et al.		
2012/0322861 A1	12/2012 Byrne et al.	2015/0165054 A1	6/2015 Liu et al.		
2013/0022980 A1	1/2013 Nelson et al.	2015/0166980 A1	6/2015 Liu et al.		
2013/0053426 A1	2/2013 Seow et al.	2015/0166981 A1	6/2015 Liu et al.		
2013/0059931 A1	3/2013 Petersen-Mahrt et al.	2015/0166982 A1	6/2015 Liu et al.		
2013/0108657 A1	5/2013 Yee et al.	2015/0166983 A1	6/2015 Liu et al.		
2013/0117869 A1	5/2013 Duchateau et al.	2015/0166984 A1	6/2015 Liu et al.		
2013/0130248 A1	5/2013 Haurwitz et al.	2015/0166985 A1	6/2015 Liu et al.		
2013/0158245 A1	6/2013 Russell et al.	2015/0191744 A1	7/2015 Wolfe et al.		
2013/0165389 A1	6/2013 Schellenberger et al.	2015/0197759 A1	7/2015 Xu et al.		
2013/0212725 A1	8/2013 Kuhn et al.	2015/0211058 A1	7/2015 Carstens		
2013/0309720 A1	11/2013 Schultz et al.	2015/0218573 A1	8/2015 Loque et al.		
2013/0344117 A1	12/2013 Mirosevich et al.	2015/0225773 A1	8/2015 Farmer et al.		
2013/0345064 A1	12/2013 Liu et al.	2015/0241440 A1	8/2015 Fasan et al.		
2014/0004280 A1	1/2014 Loomis	2015/0252358 A1	9/2015 Maeder et al.		
2014/0005269 A1	1/2014 Ngwuluka et al.	2015/0275202 A1	10/2015 Liu et al.		
2014/0017214 A1	1/2014 Cost	2015/0291965 A1	10/2015 Zhang et al.		
2014/0018404 A1	1/2014 Chen et al.	2015/0307889 A1	10/2015 Petolino et al.		
2014/0044793 A1	2/2014 Goll et al.	2015/0315252 A1	11/2015 Haugwitz et al.		
2014/0065711 A1	3/2014 Liu et al.	2015/0344549 A1	12/2015 Muir et al.		
2014/0068797 A1	3/2014 Doudna et al.	2016/0017393 A1	1/2016 Jacobson et al.		
2014/0127752 A1	5/2014 Zhou et al.	2016/0017396 A1	1/2016 Cann et al.		
2014/0128449 A1	5/2014 Liu et al.	2016/0032292 A1	2/2016 Storici et al.		
2014/0141094 A1	5/2014 Smyth et al.	2016/0032353 A1	2/2016 Braman et al.		
2014/0141487 A1	5/2014 Feldman et al.	2016/0040155 A1	2/2016 Maizels et al.		
2014/0179770 A1	6/2014 Zhang et al.	2016/0046952 A1	2/2016 Hitninger et al.		
2014/0186843 A1	7/2014 Zhang et al.	2016/0046961 A1	2/2016 Jinek et al.		
2014/0186919 A1	7/2014 Zhang et al.	2016/0046962 A1	2/2016 May et al.		
2014/0186958 A1	7/2014 Zhang et al.	2016/0053272 A1	2/2016 Wurtzel et al.		
2014/0201858 A1	7/2014 Ostertag et al.	2016/0053304 A1	2/2016 Wurtzel et al.		
2014/0234289 A1	8/2014 Liu et al.	2016/0074535 A1	3/2016 Ranganathan et al.		
2014/0248702 A1	9/2014 Zhang et al.	2016/0076093 A1	3/2016 Shendure et al.		
2014/0273037 A1	9/2014 Wu	2016/0090603 A1	3/2016 Carnes et al.		
2014/0273226 A1	9/2014 Wu	2016/0090622 A1	3/2016 Liu et al.		
2014/0273230 A1	9/2014 Chen et al.	2016/0115488 A1	4/2016 Zhang et al.		
2014/0273234 A1	9/2014 Zhang et al.	2016/0137716 A1	5/2016 El Andaloussi et al.		
2014/0283156 A1	9/2014 Zador et al.	2016/0138046 A1	5/2016 Wu		
2014/0295556 A1	10/2014 Joung et al.	2016/0153003 A1	6/2016 Joung et al.		
2014/0295557 A1	10/2014 Joung et al.	2016/0186214 A1	6/2016 Brouns et al.		
2014/0342456 A1	11/2014 Mali et al.	2016/0200779 A1	7/2016 Liu et al.		
2014/0342457 A1	11/2014 Mali et al.	2016/0201040 A1	7/2016 Liu et al.		
2014/0342458 A1	11/2014 Mali et al.	2016/0201089 A1	7/2016 Gersbach et al.		
2014/0349400 A1	11/2014 Jakimo et al.	2016/0206566 A1	7/2016 Lu et al.		
2014/0356867 A1	12/2014 Peter et al.	2016/0208243 A1	7/2016 Zhang et al.		
2014/0356956 A1	12/2014 Church et al.	2016/0208288 A1	7/2016 Liu et al.		
2014/0356958 A1	12/2014 Mali et al.	2016/0215275 A1	7/2016 Zhong		
2014/0356959 A1	12/2014 Church et al.	2016/0215276 A1	7/2016 Liu et al.		
2014/0357523 A1	12/2014 Zeiner et al.	2016/0215300 A1	7/2016 May et al.		
2014/0377868 A1	12/2014 Joung et al.	2016/0244784 A1	8/2016 Jacobson et al.		
2015/0010526 A1	1/2015 Liu et al.	2016/0244829 A1	8/2016 Bang et al.		
2015/0031089 A1	1/2015 Lindstrom	2016/0264934 A1	9/2016 Giallourakis et al.		
2015/0031132 A1	1/2015 Church et al.	2016/0272593 A1	9/2016 Ritter et al.		
2015/0031133 A1	1/2015 Church et al.	2016/0272965 A1	9/2016 Zhang et al.		
2015/0044191 A1	2/2015 Liu et al.	2016/0281072 A1	9/2016 Zhang		
2015/0044192 A1	2/2015 Liu et al.	2016/0298136 A1	10/2016 Chen et al.		
2015/0044472 A1	2/2015 Zhao	2016/0304846 A1	10/2016 Liu et al.		
2015/0050699 A1	2/2015 Siksny et al.	2016/0304855 A1	10/2016 Stark et al.		
2015/0056177 A1	2/2015 Liu et al.	2016/0312304 A1	10/2016 Sorrentino et al.		
2015/0056629 A1	2/2015 Guthrie-Honea	2016/0319262 A1	11/2016 Doudna et al.		
2015/0064138 A1	3/2015 Lu et al.	2016/0333389 A1	11/2016 Liu et al.		
2015/0064789 A1	3/2015 Paschon et al.	2016/0340622 A1	11/2016 Abdou		
2015/0071898 A1	3/2015 Liu et al.	2016/0340661 A1	11/2016 Cong et al.		
2015/0071899 A1	3/2015 Liu et al.	2016/0340662 A1	11/2016 Zhang et al.		
2015/0071900 A1	3/2015 Liu et al.	2016/0345578 A1	12/2016 Barrangou et al.		
2015/0071901 A1	3/2015 Liu et al.	2016/0346360 A1	12/2016 Quake et al.		
2015/0071902 A1	3/2015 Liu et al.	2016/0346362 A1	12/2016 Quake et al.		
2015/0071903 A1	3/2015 Liu et al.	2016/0348074 A1	12/2016 Quake et al.		
2015/0071904 A1	3/2015 Liu et al.	2016/0348096 A1	12/2016 Liu et al.		
2015/0071905 A1	3/2015 Liu et al.	2016/0350476 A1	12/2016 Quake et al.		
2015/0071906 A1	3/2015 Liu et al.	2016/0355796 A1	12/2016 Davidson et al.		

(56)

References Cited**U.S. PATENT DOCUMENTS**

2016/0369262 A1	12/2016	Reik et al.	2018/0179547 A1	6/2018	Zhang et al.
2017/0009224 A1	1/2017	Liu et al.	2018/0201921 A1	7/2018	Malcolm
2017/0009242 A1	1/2017	McKinley et al.	2018/0230464 A1	8/2018	Zhong
2017/0014449 A1	1/2017	Banger et al.	2018/0230471 A1	8/2018	Storici et al.
2017/0020922 A1	1/2017	Wagner et al.	2018/0236081 A1	8/2018	Liu et al.
2017/0022251 A1	1/2017	Rammensee et al.	2018/0237787 A1	8/2018	Maianti et al.
2017/0037432 A1	2/2017	Donohoue et al.	2018/0245066 A1	8/2018	Yao et al.
2017/0044520 A1	2/2017	Liu et al.	2018/0258418 A1	9/2018	Kim
2017/0044592 A1	2/2017	Peter et al.	2018/0265864 A1	9/2018	Li et al.
2017/0053729 A1	2/2017	Kotani et al.	2018/0273935 A1	9/2018	Lane et al.
2017/0058271 A1	3/2017	Joung et al.	2018/0273939 A1	9/2018	Yu et al.
2017/0058272 A1	3/2017	Carter et al.	2018/0273976 A1	9/2018	Ümit et al.
2017/0058298 A1	3/2017	Kennedy et al.	2018/0282722 A1	10/2018	Jakimo et al.
2017/0073663 A1	3/2017	Wang et al.	2018/0298391 A1	10/2018	Jakimo et al.
2017/0073670 A1	3/2017	Nishida et al.	2018/0305688 A1	10/2018	Zhong
2017/0087224 A1	3/2017	Quake	2018/0305704 A1	10/2018	Zhang
2017/0087225 A1	3/2017	Quake	2018/0312822 A1	11/2018	Lee et al.
2017/0088587 A1	3/2017	Quake	2018/0312825 A1	11/2018	Liu et al.
2017/0088828 A1	3/2017	Quake	2018/0312828 A1	11/2018	Liu et al.
2017/0107536 A1	4/2017	Zhang et al.	2018/0312835 A1	11/2018	Yao et al.
2017/0107560 A1	4/2017	Peter et al.	2018/0327756 A1	11/2018	Zhang et al.
2017/0112773 A1	4/2017	Stachowiak et al.	2018/0346927 A1	12/2018	Doudna et al.
2017/0114367 A1	4/2017	Hu et al.	2018/0371497 A1	12/2018	Gill et al.
2017/0121693 A1	5/2017	Liu et al.	2019/0010481 A1	1/2019	Joung et al.
2017/0145394 A1	5/2017	Yeo et al.	2019/0032053 A1	1/2019	Ji et al.
2017/0145405 A1	5/2017	Tang et al.	2019/0055543 A1	2/2019	Tran et al.
2017/0145438 A1	5/2017	Kantor	2019/0055549 A1	2/2019	Capurso et al.
2017/0152528 A1	6/2017	Zhang	2019/0062734 A1	2/2019	Cotta-Ramusino et al.
2017/0152787 A1	6/2017	Kubo et al.	2019/0093099 A1	3/2019	Liu et al.
2017/0159033 A1	6/2017	Kamtekar et al.	2019/0135869 A1	5/2019	Chatterjee et al.
2017/0166928 A1	6/2017	Vyas et al.	2019/0167810 A1	6/2019	Hean et al.
2017/0173113 A1	6/2017	Besner et al.	2019/0185883 A1	6/2019	Liu et al.
2017/0175086 A1	6/2017	Schmitt et al.	2019/0203228 A1	7/2019	Bouille et al.
2017/0175104 A1	6/2017	Doudna et al.	2019/0218547 A1	7/2019	Lee et al.
2017/0175142 A1	6/2017	Zhang et al.	2019/0224331 A1	7/2019	Wiklander
2017/0191047 A1	7/2017	Terns et al.	2019/0225955 A1	7/2019	Liu et al.
2017/0191078 A1	7/2017	Zhang et al.	2019/0233847 A1	8/2019	Savage et al.
2017/0198269 A1	7/2017	Zhang et al.	2019/0241633 A1	8/2019	Fotin-Mleczek et al.
2017/0198277 A1	7/2017	Kmiec et al.	2019/0256842 A1	8/2019	Liu et al.
2017/0198302 A1	7/2017	Feng et al.	2019/0264202 A1	8/2019	Church et al.
2017/0211061 A1	7/2017	Weiss et al.	2019/0276816 A1	9/2019	Liu et al.
2017/0224843 A1	8/2017	Deglon et al.	2019/0309290 A1	10/2019	Neuteboom et al.
2017/0226522 A1	8/2017	Hu et al.	2019/0322992 A1	10/2019	Liu et al.
2017/0233703 A1	8/2017	Xie et al.	2019/0330619 A1	10/2019	Smith et al.
2017/0233708 A1	8/2017	Liu et al.	2019/0352632 A1	11/2019	Liu et al.
2017/0233756 A1	8/2017	Begemann et al.	2019/0367891 A1	12/2019	Liu et al.
2017/0247671 A1	8/2017	Yung et al.	2019/0388347 A1	12/2019	Wiklander et al.
2017/0247703 A1	8/2017	Sloan et al.	2020/0010818 A1	1/2020	Liu et al.
2017/0268022 A1	9/2017	Liu et al.	2020/0010835 A1	1/2020	Maianti et al.
2017/0275648 A1	9/2017	Barrangou et al.	2020/0023012 A1	1/2020	Joseph et al.
2017/0275665 A1	9/2017	Silas et al.	2020/0056206 A1	2/2020	Tremblay et al.
2017/0283797 A1	10/2017	Robb et al.	2020/0060980 A1	2/2020	Von Maltzahn et al.
2017/0283831 A1	10/2017	Zhang et al.	2020/0062813 A1	2/2020	Nordin et al.
2017/0306306 A1	10/2017	Potter et al.	2020/0063127 A1	2/2020	Lu et al.
2017/0314016 A1	11/2017	Kim et al.	2020/0071722 A1	3/2020	Liu et al.
2017/0362635 A1	12/2017	Chamberlain et al.	2020/0109398 A1	4/2020	Rubens et al.
2018/0023062 A1	1/2018	Lamb et al.	2020/0172931 A1	6/2020	Liu et al.
2018/0033787 A1	2/2018	Gao et al.	2020/0181619 A1	6/2020	Tang et al.
2018/0064077 A1	3/2018	Dunham et al.	2020/0190493 A1	6/2020	Liu et al.
2018/0066258 A1	3/2018	Powell	2020/0255868 A1	8/2020	Liu et al.
2018/0068062 A1	3/2018	Zhang et al.	2020/0277587 A1	9/2020	Liu et al.
2018/0073012 A1	3/2018	Liu et al.	2020/0323984 A1	10/2020	Liu et al.
2018/0080051 A1	3/2018	Sheikh et al.	2020/0399619 A1	12/2020	Maianti et al.
2018/0087046 A1	3/2018	Badran et al.	2020/0399626 A1	12/2020	Liu et al.
2018/0100147 A1	4/2018	Yates et al.	2021/0054416 A1	2/2021	Liu et al.
2018/0105867 A1	4/2018	Xiao et al.	2021/0115428 A1	4/2021	Maianti et al.
2018/0119118 A1	5/2018	Lu et al.	2021/0198330 A1	7/2021	Liu et al.
2018/0127759 A1	5/2018	Lu et al.	2021/0214698 A1	7/2021	Liu et al.
2018/0127780 A1	5/2018	Liu et al.	2021/0230577 A1	7/2021	Liu et al.
2018/0155708 A1	6/2018	Church et al.	2021/0254127 A1	8/2021	Liu et al.
2018/0155720 A1	6/2018	Donohoue et al.	2021/0315994 A1	10/2021	Liu et al.
2018/0163213 A1	6/2018	Aneja et al.	2021/0317440 A1	10/2021	Liu et al.
2018/0170984 A1	6/2018	Harris et al.	2022/0033785 A1	2/2022	Liu et al.
2018/0177727 A1	6/2018	Kalluri et al.	2022/0170013 A1	6/2022	Liu et al.
2018/0179503 A1	6/2018	Maianti et al.	2022/0177877 A1	6/2022	Church et al.
			2022/0204975 A1	6/2022	Liu et al.
			2022/0213507 A1	7/2022	Liu et al.
			2022/0220462 A1	7/2022	Liu et al.
			2022/0238182 A1	7/2022	Shen et al.

(56)	References Cited					
U.S. PATENT DOCUMENTS						
2022/0249697 A1	8/2022	Liu et al.	CN	103820454 A	5/2014	
2022/0282275 A1	9/2022	Liu et al.	CN	103911376 A	7/2014	
2022/0290115 A1	9/2022	Liu et al.	CN	103923911 A	7/2014	
2022/0307001 A1	9/2022	Liu et al.	CN	103088008 A	8/2014	
2022/0307003 A1	9/2022	Liu et al.	CN	103981211 A	8/2014	
2022/0315906 A1	10/2022	Liu et al.	CN	103981212 A	8/2014	
2022/0356469 A1	11/2022	Liu et al.	CN	104004778 A	8/2014	
2022/0380740 A1	12/2022	Liu et al.	CN	104004782 A	8/2014	
2022/0389395 A1	12/2022	Liu et al.	CN	104017821 A	9/2014	
2023/0002745 A1	1/2023	Liu et al.	CN	104109687 A	10/2014	
2023/0021641 A1	1/2023	Liu et al.	CN	104178461 A	12/2014	
2023/0056852 A1	2/2023	Liu et al.	CN	104404036 A	3/2015	
2023/0058176 A1	2/2023	Liu et al.	CN	104450774 A	3/2015	
2023/0078265 A1	3/2023	Liu et al.	CN	104480144 A	4/2015	
2023/0086199 A1	3/2023	Liu et al.	CN	104498493 A	4/2015	
2023/0090221 A1	3/2023	Liu et al.	CN	104504304 A	4/2015	
2023/0108687 A1	4/2023	Liu et al.	CN	104531704 A	4/2015	
2023/0123669 A1	4/2023	Liu et al.	CN	104531705 A	4/2015	
2023/0127008 A1	4/2023	Liu et al.	CN	104560864 A	4/2015	
2023/0159913 A1	5/2023	Liu et al.	CN	104561095 A	4/2015	
2023/0193295 A1	6/2023	Maianti et al.	CN	104593418 A	5/2015	
2023/0220374 A1	7/2023	Liu et al.	CN	104593422 A	5/2015	
2023/0272425 A1	8/2023	Liu et al.	CN	104611370 A	5/2015	
2023/0279443 A1	9/2023	Liu et al.	CN	104651392 A	5/2015	
2023/0332144 A1	10/2023	Liu et al.	CN	104651398 A	5/2015	
2023/0340465 A1	10/2023	Liu et al.	CN	104651399 A	5/2015	
2023/0340466 A1	10/2023	Liu et al.	CN	104651401 A	5/2015	
2023/0340467 A1	10/2023	Liu et al.	CN	104673816 A	6/2015	
2023/0348883 A1	11/2023	Liu et al.	CN	104725626 A	6/2015	
2023/0357766 A1	11/2023	Liu et al.	CN	104726449 A	6/2015	
2023/0383289 A1	11/2023	Liu et al.	CN	104726494 A	6/2015	
2024/0035017 A1	2/2024	Liu et al.	CN	104745626 A	7/2015	
2024/0076652 A1	3/2024	Liu et al.	CN	104762321 A	7/2015	
2024/0110166 A1	4/2024	Maianti et al.	CN	104805078 A	7/2015	
2024/0124866 A1	4/2024	Liu et al.	CN	104805099 A	7/2015	
2024/0173430 A1	5/2024	Liu et al.	CN	104805118 A	7/2015	
2024/0209329 A1	6/2024	Liu et al.	CN	104846010 A	8/2015	
2024/0229077 A1	7/2024	Liu et al.	CN	104894068 A	9/2015	
2024/0271116 A1	8/2024	Maianti et al.	CN	104894075 A	9/2015	
2024/0287487 A1	8/2024	Liu et al.	CN	104928321 A	9/2015	
2024/0327872 A1	10/2024	Liu et al.	CN	105039339 A	11/2015	
2024/0401018 A1	12/2024	Liu et al.	CN	105039399 A	11/2015	
2024/0417715 A1	12/2024	Liu et al.	CN	105063061 A	11/2015	
2024/0417719 A1	12/2024	Liu et al.	CN	105121648 A	12/2015	
2024/0417753 A1	12/2024	Liu et al.	CN	105132427 A	12/2015	
2025/0011748 A1	1/2025	Liu et al.	CN	105132451 A	12/2015	
2025/0027114 A1	1/2025	Liu et al.	CN	105177038 A	12/2015	
2025/0034549 A1	1/2025	Liu et al.	CN	105177126 A	12/2015	
2025/0059244 A1	2/2025	Liu et al.	CN	105210981 A	1/2016	
2025/0064979 A1	2/2025	Liu et al.	CN	105219799 A	1/2016	
2025/0064981 A1	2/2025	Liu et al.	CN	105238806 A	1/2016	
FOREIGN PATENT DOCUMENTS						
AU	2015252023 A1	11/2015	CN	105255937 A	1/2016	
AU	2015101792 A4	1/2016	CN	105274144 A	1/2016	
BR	112015013786 A2	7/2017	CN	105296518 A	2/2016	
CA	2480696 A1	10/2003	CN	105296537 A	2/2016	
CA	2894668 A1	6/2014	CN	105316324 A	2/2016	
CA	2894681 A1	6/2014	CN	105316327 A	2/2016	
CA	2894684 A1	6/2014	CN	105316337 A	2/2016	
CA	2852593 A1	11/2015	CN	105331607 A	2/2016	
CN	1069962 A	3/1993	CN	105331608 A	2/2016	
CN	101460619 A	6/2009	CN	105331609 A	2/2016	
CN	101873862 A	10/2010	CN	105331627 A	2/2016	
CN	104342457 A	2/2011	CN	105400773 A	3/2016	
CN	102057039 A	5/2011	CN	105400779 A	3/2016	
CN	102892777 A	1/2013	CN	105400810 A	3/2016	
CN	103224947 A	7/2013	CN	105441451 A	3/2016	
CN	103233028 A	8/2013	CN	105462968 A	4/2016	
CN	103388006 A	11/2013	CN	105463003 A	4/2016	
CN	103614415 A	3/2014	CN	105463027 A	4/2016	
CN	103642836 A	3/2014	CN	105492608 A	4/2016	
CN	103668472 A	3/2014	CN	105492609 A	4/2016	
CN	103820441 A	5/2014	CN	105505976 A	4/2016	
			CN	105505979 A	4/2016	
			CN	105518134 A	4/2016	
			CN	105518135 A	4/2016	

(56)	References Cited	CN	106191057 A	12/2016
	FOREIGN PATENT DOCUMENTS	CN	106191061 A	12/2016
CN	105518137 A 4/2016	CN	106191062 A	12/2016
CN	105518138 A 4/2016	CN	106191064 A	12/2016
CN	105518139 A 4/2016	CN	106191071 A	12/2016
CN	105518140 A 4/2016	CN	106191099 A	12/2016
CN	105543228 A 5/2016	CN	106191107 A	12/2016
CN	105543266 A 5/2016	CN	106191113 A	12/2016
CN	105543270 A 5/2016	CN	106191114 A	12/2016
CN	105567688 A 5/2016	CN	106222177 A	12/2016
CN	105567689 A 5/2016	CN	106222193 A	12/2016
CN	105567734 A 5/2016	CN	106222203 A	12/2016
CN	105567735 A 5/2016	CN	106232823 A	12/2016
CN	105567738 A 5/2016	CN	106244555 A	12/2016
CN	105593367 A 5/2016	CN	106244557 A	12/2016
CN	105594664 A 5/2016	CN	106244591 A	12/2016
CN	105602987 A 5/2016	CN	106244609 A	12/2016
CN	105624146 A 6/2016	CN	106282241 A	1/2017
CN	105624187 A 6/2016	CN	106318934 A	1/2017
CN	105646719 A 6/2016	CN	106318973 A	1/2017
CN	105647922 A 6/2016	CN	106350540 A	1/2017
CN	105647962 A 6/2016	CN	106367435 A	2/2017
CN	105647968 A 6/2016	CN	106399306 A	2/2017
CN	105647969 A 6/2016	CN	106399311 A	2/2017
CN	105671070 A 6/2016	CN	106399360 A	2/2017
CN	105671083 A 6/2016	CN	106399367 A	2/2017
CN	105695485 A 6/2016	CN	106399375 A	2/2017
CN	105779448 A 7/2016	CN	106399377 A	2/2017
CN	105779449 A 7/2016	CN	106434651 A	2/2017
CN	105802980 A 7/2016	CN	106434663 A	2/2017
CN	105821039 A 8/2016	CN	106434688 A	2/2017
CN	105821040 A 8/2016	CN	106434737 A	2/2017
CN	105821049 A 8/2016	CN	106434748 A	2/2017
CN	105821072 A 8/2016	CN	106434752 A	2/2017
CN	105821075 A 8/2016	CN	106434782 A	2/2017
CN	105821116 A 8/2016	CN	106446600 A	2/2017
CN	105838733 A 8/2016	CN	106479985 A	3/2017
CN	105861547 A 8/2016	CN	106480027 A	3/2017
CN	105861552 A 8/2016	CN	106480036 A	3/2017
CN	105861554 A 8/2016	CN	106480067 A	3/2017
CN	105886498 A 8/2016	CN	106480080 A	3/2017
CN	105886534 A 8/2016	CN	106480083 A	3/2017
CN	105886616 A 8/2016	CN	106480097 A	3/2017
CN	105907758 A 8/2016	CN	106544351 A	3/2017
CN	105907785 A 8/2016	CN	106544353 A	3/2017
CN	105925608 A 9/2016	CN	106544357 A	3/2017
CN	105934516 A 9/2016	CN	106554969 A	4/2017
CN	105950560 A 9/2016	CN	106566838 A	4/2017
CN	105950626 A 9/2016	CN	106701763 A	5/2017
CN	105950633 A 9/2016	CN	106701808 A	5/2017
CN	105950639 A 9/2016	CN	106701818 A	5/2017
CN	105985985 A 10/2016	CN	106701823 A	5/2017
CN	106011104 A 10/2016	CN	106701830 A	5/2017
CN	106011150 A 10/2016	CN	106754912 A	5/2017
CN	106011167 A 10/2016	CN	106755026 A	5/2017
CN	106011171 A 10/2016	CN	106755077 A	5/2017
CN	106032540 A 10/2016	CN	106755088 A	5/2017
CN	106047803 A 10/2016	CN	106755091 A	5/2017
CN	106047877 A 10/2016	CN	106755097 A	5/2017
CN	106047930 A 10/2016	CN	106755424 A	5/2017
CN	106086008 A 11/2016	CN	106801056 A	6/2017
CN	106086028 A 11/2016	CN	106834323 A	6/2017
CN	106086061 A 11/2016	CN	106834341 A	6/2017
CN	106086062 A 11/2016	CN	106834347 A	6/2017
CN	106103475 A 11/2016	CN	106845151 A	6/2017
CN	106109417 A 11/2016	CN	106868008 A	6/2017
CN	106119275 A 11/2016	CN	106868031 A	6/2017
CN	106119283 A 11/2016	CN	106906240 A	6/2017
CN	106148286 A 11/2016	CN	106906242 A	6/2017
CN	106148370 A 11/2016	CN	106916820 A	7/2017
CN	106148416 A 11/2016	CN	106916852 A	7/2017
CN	106167525 A 11/2016	CN	106939303 A	7/2017
CN	106167808 A 11/2016	CN	106947750 A	7/2017
CN	106167810 A 11/2016	CN	106947780 A	7/2017
CN	106167821 A 11/2016	CN	106957830 A	7/2017
CN	106172238 A 12/2016	CN	106957831 A	7/2017
CN	106190903 A 12/2016	CN	106957844 A	7/2017

(56)	References Cited	CN	107446954 A	12/2017
	FOREIGN PATENT DOCUMENTS	CN	107460196 A	12/2017
CN	106957855 A 7/2017	CN	107474129 A	12/2017
CN	106957858 A 7/2017	CN	107475300 A	12/2017
CN	106967697 A 7/2017	CN	107488649 A	12/2017
CN	106967726 A 7/2017	CN	107502608 A	12/2017
CN	106978428 A 7/2017	CN	107502618 A	12/2017
CN	106987570 A 7/2017	CN	107513531 A	12/2017
CN	106987757 A 7/2017	CN	107519492 A	12/2017
CN	107012164 A 8/2017	CN	107523567 A	12/2017
CN	107012174 A 8/2017	CN	107541525 A	1/2018
CN	107012213 A 8/2017	CN	107557373 A	1/2018
CN	107012250 A 8/2017	CN	107557378 A	1/2018
CN	107022562 A 8/2017	CN	107557381 A	1/2018
CN	107034188 A 8/2017	CN	107557390 A	1/2018
CN	107034218 A 8/2017	CN	107557393 A	1/2018
CN	107034229 A 8/2017	CN	107557394 A	1/2018
CN	107043775 A 8/2017	CN	107557455 A	1/2018
CN	107043779 A 8/2017	CN	107574179 A	1/2018
CN	107043787 A 8/2017	CN	107586777 A	1/2018
CN	107058320 A 8/2017	CN	107604003 A	1/2018
CN	107058328 A 8/2017	CN	107619829 A	1/2018
CN	107058358 A 8/2017	CN	107619837 A	1/2018
CN	107058372 A 8/2017	CN	107630006 A	1/2018
CN	107083392 A 8/2017	CN	107630041 A	1/2018
CN	107099533 A 8/2017	CN	107630042 A	1/2018
CN	107099850 A 8/2017	CN	107630043 A	1/2018
CN	107119053 A 9/2017	CN	107641631 A	1/2018
CN	107119071 A 9/2017	CN	107653256 A	2/2018
CN	107129999 A 9/2017	CN	107686848 A	2/2018
CN	107130000 A 9/2017	CN	206970581 U	2/2018
CN	107142272 A 9/2017	CN	107760652 A	3/2018
CN	107142282 A 9/2017	CN	107760663 A	3/2018
CN	107177591 A 9/2017	CN	107760684 A	3/2018
CN	107177595 A 9/2017	CN	107760715 A	3/2018
CN	107177625 A 9/2017	CN	107784200 A	3/2018
CN	107177631 A 9/2017	CN	107794272 A	3/2018
CN	107190006 A 9/2017	CN	107794276 A	3/2018
CN	107190008 A 9/2017	CN	107815463 A	3/2018
CN	107217042 A 9/2017	CN	107828738 A	3/2018
CN	107217075 A 9/2017	CN	107828794 A	3/2018
CN	107227307 A 10/2017	CN	107828826 A	3/2018
CN	107227352 A 10/2017	CN	107828874 A	3/2018
CN	107236737 A 10/2017	CN	107858346 A	3/2018
CN	107236739 A 10/2017	CN	107858373 A	3/2018
CN	107236741 A 10/2017	CN	107880132 A	4/2018
CN	107245502 A 10/2017	CN	107881184 A	4/2018
CN	107254485 A 10/2017	CN	107893074 A	4/2018
CN	107266541 A 10/2017	CN	107893075 A	4/2018
CN	107267515 A 10/2017	CN	107893076 A	4/2018
CN	107287245 A 10/2017	CN	107893080 A	4/2018
CN	107298701 A 10/2017	CN	107893086 A	4/2018
CN	107299114 A 10/2017	CN	107904261 A	4/2018
CN	107304435 A 10/2017	CN	107937427 A	4/2018
CN	107312785 A 11/2017	CN	107937432 A	4/2018
CN	107312793 A 11/2017	CN	107937501 A	4/2018
CN	107312795 A 11/2017	CN	107974466 A	5/2018
CN	107312798 A 11/2017	CN	107988229 A	5/2018
CN	107326042 A 11/2017	CN	107988246 A	5/2018
CN	107326046 A 11/2017	CN	107988256 A	5/2018
CN	107354156 A 11/2017	CN	107988268 A	5/2018
CN	107354173 A 11/2017	CN	108018316 A	5/2018
CN	107356793 A 11/2017	CN	108034656 A	5/2018
CN	107362372 A 11/2017	CN	108048466 A	5/2018
CN	107365786 A 11/2017	CN	108102940 A	6/2018
CN	107365804 A 11/2017	CN	108103090 A	6/2018
CN	107384894 A 11/2017	CN	108103092 A	6/2018
CN	107384922 A 11/2017	CN	108103098 A	6/2018
CN	107400677 A 11/2017	CN	108103586 A	6/2018
CN	107418974 A 12/2017	CN	108148835 A	6/2018
CN	107435051 A 12/2017	CN	108148837 A	6/2018
CN	107435069 A 12/2017	CN	108148873 A	6/2018
CN	107446922 A 12/2017	CN	108192956 A	6/2018
CN	107446923 A 12/2017	CN	108243575 A	7/2018
CN	107446924 A 12/2017	CN	108251423 A	7/2018
CN	107446932 A 12/2017	CN	108251451 A	7/2018
CN	107446951 A 12/2017	CN	108251452 A	7/2018

(56)	References Cited	CN	109517841 A	3/2019
	FOREIGN PATENT DOCUMENTS	EP	0264166 A1	4/1988
CN	108342480 A 7/2018	EP	0321201 B2	6/1989
CN	108359691 A 8/2018	EP	0519463 A1	12/1992
CN	108359712 A 8/2018	EP	1085892 A2	3/2001
CN	108384784 A 8/2018	EP	1092770 A2	4/2001
CN	108396027 A 8/2018	EP	2350295 B1	5/2013
CN	108410877 A 8/2018	EP	2604255 A1	6/2013
CN	108410906 A 8/2018	EP	2840140 A1	2/2015
CN	108410907 A 8/2018	EP	2877490 A2	6/2015
CN	108410911 A 8/2018	EP	2966170 A1	1/2016
CN	108424931 A 8/2018	EP	3009511 A2	4/2016
CN	108441519 A 8/2018	EP	3031921 A1	6/2016
CN	108441520 A 8/2018	EP	3045537 A1	7/2016
CN	108472314 A 8/2018	EP	3115457 A	1/2017
CN	108486108 A 9/2018	EP	3144390 A1	3/2017
CN	108486111 A 9/2018	EP	2583974 B1	4/2017
CN	108486145 A 9/2018	EP	3199632 A1	8/2017
CN	108486146 A 9/2018	EP	3216867 A1	9/2017
CN	108486154 A 9/2018	EP	3235828 A1	10/2017
CN	108486159 A 9/2018	EP	3252160 A1	12/2017
CN	108486234 A 9/2018	EP	2498823 B1	8/2018
CN	108504657 A 9/2018	EP	3365437 A1	8/2018
CN	108504685 A 9/2018	EP	3454889 A2	3/2019
CN	108504693 A 9/2018	EP	3008192 B1	7/2019
CN	108513575 A 9/2018	ES	3079725 B1	10/2019
CN	108546712 A 9/2018	GB	3450553 B1	12/2019
CN	108546717 A 9/2018	GB	2740248 T3	2/2020
CN	108546718 A 9/2018	GB	2528177 A	1/2016
CN	108559730 A 9/2018	HK	2531454 A	4/2016
CN	108559732 A 9/2018	JP	2542653 A	3/2017
CN	108559745 A 9/2018	JP	1208045 A1	2/2016
CN	108559760 A 9/2018	JP	2007-501626 A	2/2007
CN	108570479 A 9/2018	JP	2008-515405 A	5/2008
CN	108588071 A 9/2018	JP	2010-033344 A	2/2010
CN	108588123 A 9/2018	JP	2010-535744 A	11/2010
CN	108588128 A 9/2018	JP	2010-539929 A	12/2010
CN	108588182 A 9/2018	JP	2011-081011 A	4/2011
CN	108610399 A 10/2018	JP	2011-523353 A	8/2011
CN	108611364 A 10/2018	JP	2012-525146 A	10/2012
CN	108624622 A 10/2018	JP	2012-210172 A	11/2012
CN	108642053 A 10/2018	JP	2012-531909 A	12/2012
CN	108642055 A 10/2018	JP	2015-523856 A	8/2015
CN	108642077 A 10/2018	JP	2015-532654 A	11/2015
CN	108642078 A 10/2018	JP	2016-525888 A	9/2016
CN	108642090 A 10/2018	JP	2016-534132 A	11/2016
CN	108690844 A 10/2018	JP	2017-500035 A	1/2017
CN	108699542 A 10/2018	JP	2018-521045 A	8/2018
CN	108707604 A 10/2018	JP	2019-506123 A	2/2019
CN	108707620 A 10/2018	JP	6629734 B2	1/2020
CN	108707621 A 10/2018	JP	6633524 B2	1/2020
CN	108707628 A 10/2018	JP	6830517 B2	2/2021
CN	108707629 A 10/2018	KR	7324523 B2	8/2023
CN	108715850 A 10/2018	KR	101584933 B1	1/2016
CN	108728476 A 11/2018	KR	2016-0050069 A	5/2016
CN	108728486 A 11/2018	KR	20160133380 A	11/2016
CN	108753772 A 11/2018	KR	20170037025 A	4/2017
CN	108753783 A 11/2018	KR	20170037028 A	4/2017
CN	108753813 A 11/2018	KR	101748575 B1	6/2017
CN	108753817 A 11/2018	KR	20170128137 A	11/2017
CN	108753832 A 11/2018	KR	2018-0022465 A	3/2018
CN	108753835 A 11/2018	RU	2016104674 A	8/2017
CN	108753836 A 11/2018	RU	2634395 C1	10/2017
CN	108795902 A 11/2018	RU	2652899 C1	5/2018
CN	108822217 A 11/2018	RU	2015128057 A	3/2019
CN	108823248 A 11/2018	RU	2015128098 A	3/2019
CN	108823249 A 11/2018	RU	2687451 C1	5/2019
CN	108823291 A 11/2018	RU	2019112514 A	6/2019
CN	108841845 A 11/2018	RU	2019127300 A	9/2019
CN	108853133 A 11/2018	RU	2701850 C2	10/2019
CN	108866093 A 11/2018	RU	10201707569 Y	10/2017
CN	108893529 A 11/2018	SG	10201710486X	1/2018
CN	108913664 A 11/2018	SG	10201710487V	1/2018
CN	108913691 A 11/2018	TW	10201710488 T	1/2018
CN	108913714 A 11/2018	TW	I608100 B	12/2017
CN	108913717 A 11/2018	TW	201809272 A	3/2018
CN	208034188 U 11/2018	WO	2018-29773 A	8/2018
CN		WO	WO 1990/002809	3/1990
CN		WO	WO 1991/003162 A1	3/1991

(56)	References Cited							
FOREIGN PATENT DOCUMENTS								
WO	WO 1991/016024 A1	10/1991	WO	WO 2013/013105 A1	1/2013			
WO	WO 1991/017271 A1	11/1991	WO	WO 2013/039857 A1	3/2013			
WO	WO 1991/017424 A1	11/1991	WO	WO 2013/039861 A2	3/2013			
WO	WO 1992/006188 A2	4/1992	WO	WO 2013/040093 A2	3/2013			
WO	WO 1992/006200 A1	4/1992	WO	WO 2013/045632 A1	4/2013			
WO	WO 1992/007065 A1	4/1992	WO	WO 2013/047844 A1	4/2013			
WO	WO 1993/015187 A1	8/1993	WO	WO 2013/066438 A2	5/2013			
WO	WO 1993/024641 A2	12/1993	WO	WO 2013/086441 A2	6/2013			
WO	WO 1994/018316 A2	8/1994	WO	WO 2013/086444 A2	6/2013			
WO	WO 1994/026877 A1	11/1994	WO	WO 2013/098244 A1	7/2013			
WO	WO 1996/004403 A1	2/1996	WO	WO 2013/119602 A1	8/2013			
WO	WO 1996/010640 A1	4/1996	WO	WO 2013/120022 A2	8/2013			
WO	WO 1997/025416 A2	7/1997	WO	WO 2013/122617 A1	8/2013			
WO	WO 1998/032845 A1	7/1998	WO	WO 2013/126794 A1	8/2013			
WO	WO 1998/050538 A1	11/1998	WO	WO 2013/130683 A2	9/2013			
WO	WO 2001/036452 A2	5/2001	WO	WO 2013/130824 A1	9/2013			
WO	WO 2001/038547 A2	5/2001	WO	WO 2013/141680 A1	9/2013			
WO	WO 2002/059296 A2	8/2002	WO	WO 2013/142578 A1	9/2013			
WO	WO 2002/068676 A2	9/2002	WO	WO 2013/152359 A1	10/2013			
WO	WO 2002/103028 A2	12/2002	WO	WO 2013/160230 A1	10/2013			
WO	WO 2003/004608 A2	1/2003	WO	WO 2013/166315 A1	11/2013			
WO	WO 2004/007684 A2	1/2004	WO	WO 2013/169398 A2	11/2013			
WO	WO 2005/014791 A2	2/2005	WO	WO 2013/169802 A1	11/2013			
WO	WO 2005/019415 A2	3/2005	WO	WO 2013/176772 A1	11/2013			
WO	WO 2006/002547 A1	1/2006	WO	WO 2013/176915 A1	11/2013			
WO	WO 2006/042112 A2	4/2006	WO	WO 2013/176916 A1	11/2013			
WO	WO 2007/025097 A2	3/2007	WO	WO 2013/181440 A1	12/2013			
WO	WO 2007/037444 A1	4/2007	WO	WO 2013/186754 A2	12/2013			
WO	WO 2007/066923 A1	6/2007	WO	WO 2013/188037 A2	12/2013			
WO	WO 2007/136815 A2	11/2007	WO	WO 2013/188522 A2	12/2013			
WO	WO 2007/143574 A1	12/2007	WO	WO 2013/188638 A2	12/2013			
WO	WO 2008/005529 A2	1/2008	WO	WO 2013/192278 A1	12/2013			
WO	WO 2008/108989 A2	9/2008	WO	WO 2013/142378 A9	1/2014			
WO	WO 2009/002418 A2	12/2008	WO	WO 2014/004336 A2	1/2014			
WO	WO 2009/019317 A1	2/2009	WO	WO 2014/005042 A2	1/2014			
WO	WO 2009/098290 A1	8/2009	WO	WO 2014/011237 A1	1/2014			
WO	WO 2009/134808 A2	11/2009	WO	WO 2014/011901 A2	1/2014			
WO	WO 2010/011961 A2	1/2010	WO	WO 2014/018423 A2	1/2014			
WO	WO 2010/012902 A1	2/2010	WO	WO 2014/020608 A1	2/2014			
WO	WO 2010/028347 A2	3/2010	WO	WO 2014/022120 A1	2/2014			
WO	WO 2010/054108 A2	5/2010	WO	WO 2014/022702 A2	2/2014			
WO	WO 2010/054154 A2	5/2010	WO	WO 2014/036219 A2	3/2014			
WO	WO 2010/068289 A2	6/2010	WO	WO 2014/039513 A2	3/2014			
WO	WO 2010/075424 A2	7/2010	WO	WO 2014/039523 A1	3/2014			
WO	WO 2010/091122 A1	8/2010	WO	WO 2014/039585 A2	3/2014			
WO	WO 2010/102257 A2	9/2010	WO	WO 2014/039684 A1	3/2014			
WO	WO 2010/104749 *	9/2010 C07K 19/00		WO	WO 2014/039692 A2	3/2014	
WO	WO 2010/129019 A2	11/2010			WO	WO 2014/039702 A2	3/2014	
WO	WO 2010/129023 A2	11/2010			WO	WO 2014/039872 A1	3/2014	
WO	WO 2010/132092 A2	11/2010			WO	WO 2014/039970 A1	3/2014	
WO	WO 2010/144150 A2	12/2010			WO	WO 2014/041327 A1	3/2014	
WO	WO 2011/002503 A1	1/2011			WO	WO 2014/043143 A1	3/2014	
WO	WO 2011/017293 A2	2/2011			WO	WO 2014/047103 A2	3/2014	
WO	WO 2011/053868 A1	5/2011			WO	WO 2014/055782 A1	4/2014	
WO	WO 2011/053982 A2	5/2011			WO	WO 2014/059173 A2	4/2014	
WO	WO 2011/068810 A1	6/2011			WO	WO 2014/059255 A1	4/2014	
WO	WO 2011/075627 A1	6/2011			WO	WO 2014/065596 A1	5/2014	
WO	WO 2011/091311 A2	7/2011			WO	WO 2014/066505 A1	5/2014	
WO	WO 2011/091396 A1	7/2011			WO	WO 2014/068346 A2	5/2014	
WO	WO 2011/109031 A1	9/2011			WO	WO 2014/070887 A1	5/2014	
WO	WO 2011/143124 A2	11/2011			WO	WO 2014/071006 A1	5/2014	
WO	WO 2011/147590 A2	12/2011			WO	WO 2014/071219 A1	5/2014	
WO	WO 2011/159369 A1	12/2011			WO	WO 2014/071235 A1	5/2014	
WO	WO 2012/054726 A1	4/2012			WO	WO 2014/072941 A1	5/2014	
WO	WO 2012/061815 A2	5/2012			WO	WO 2014/081729 A1	5/2014	
WO	WO 2012/065043 A2	5/2012			WO	WO 2014/081730 A1	5/2014	
WO	WO 2012/088381 A2	6/2012			WO	WO 2014/081855 A1	5/2014	
WO	WO 2012/125445 A2	9/2012			WO	WO 2014/082644 A1	6/2014	
WO	WO 2012/138927 A2	10/2012			WO	WO 2014/085261 A1	6/2014	
WO	WO 2012/149470 A1	11/2012			WO	WO 2014/085593 A1	6/2014	
WO	WO 2012/158985 A2	11/2012			WO	WO 2014/085830 A2	6/2014	
WO	WO 2012/158986 A2	11/2012			WO	WO 2014/089212 A1	6/2014	
WO	WO 2012/164565 A1	12/2012			WO	WO 2014/089290 A1	6/2014	
WO	WO 2012/170930 A1	12/2012			WO	WO 2014/089348 A1	6/2014	
WO	WO 2013/012674 A1	1/2013			WO	WO 2014/089513 A1	6/2014	
					WO	WO 2014/089533 A2	6/2014	
					WO	WO 2014/089541 A2	6/2014	
					WO	WO 2014/093479 A1	6/2014	
					WO	WO 2014/093595 A1	6/2014	

(56)	References Cited					
FOREIGN PATENT DOCUMENTS						
WO	WO 2014/093622 A2	6/2014	WO	WO 2014/204578 A1	12/2014	
WO	WO 2014/093635 A1	6/2014	WO	WO 2014/204723 A1	12/2014	
WO	WO 2014/093655 A2	6/2014	WO	WO 2014/204724 A1	12/2014	
WO	WO 2014/093661 A2	6/2014	WO	WO 2014/204725 A1	12/2014	
WO	WO 2014/093694 A1	6/2014	WO	WO 2014/204726 A1	12/2014	
WO	WO 2014/093701 A1	6/2014	WO	WO 2014/204727 A1	12/2014	
WO	WO 2014/093709 A1	6/2014	WO	WO 2014/204728 A1	12/2014	
WO	WO 2014/093712 A1	6/2014	WO	WO 2014/204729 A1	12/2014	
WO	WO 2014/093718 A1	6/2014	WO	WO 2014/205192 A2	12/2014	
WO	WO 2014/093736 A1	6/2014	WO	WO 2014/207043 A1	12/2014	
WO	WO 2014/093768 A1	6/2014	WO	WO 2015/002780 A1	1/2015	
WO	WO 2014/093852 A1	6/2014	WO	WO 2015/004241 A2	1/2015	
WO	WO 2014/096972 A2	6/2014	WO	WO 2015/006747 A2	1/2015	
WO	WO 2014/099744 A1	6/2014	WO	WO 2015/007194 A1	1/2015	
WO	WO 2014/099750 A2	6/2014	WO	WO 2015/010114 A1	1/2015	
WO	WO 2014/104878 A1	7/2014	WO	WO 2015/011483 A1	1/2015	
WO	WO 2014/110006 A1	7/2014	WO	WO 2015/013583 A2	1/2015	
WO	WO 2014/110552 A1	7/2014	WO	WO 2015/017866 A1	2/2015	
WO	WO 2014/113493 A1	7/2014	WO	WO 2015/018503 A1	2/2015	
WO	WO 2014/123967 A2	8/2014	WO	WO 2015/021353 A1	2/2015	
WO	WO 2014/124226 A1	8/2014	WO	WO 2015/021426 A1	2/2015	
WO	WO 2014/125668 A1	8/2014	WO	WO 2015/021990 A1	2/2015	
WO	WO 2014/127287 A1	8/2014	WO	WO 2015/024017 A2	2/2015	
WO	WO 2014/128324 A1	8/2014	WO	WO 2015/024986 A1	2/2015	
WO	WO 2014/128659 A1	8/2014	WO	WO 2015/026883 A1	2/2015	
WO	WO 2014/130706 A1	8/2014	WO	WO 2015/026885 A1	2/2015	
WO	WO 2014/130955 A1	8/2014	WO	WO 2015/026886 A1	2/2015	
WO	WO 2014/131833 A1	9/2014	WO	WO 2015/026887 A1	2/2015	
WO	WO 2014/138379 A1	9/2014	WO	WO 2015/027134 A1	2/2015	
WO	WO 2014/143381 A1	9/2014	WO	WO 2015/028969 A2	3/2015	
WO	WO 2014/144094 A1	9/2014	WO	WO 2015/030881 A1	3/2015	
WO	WO 2014/144155 A1	9/2014	WO	WO 2015/031619 A1	3/2015	
WO	WO 2014/144288 A1	9/2014	WO	WO 2015/031775 A1	3/2015	
WO	WO 2014/144592 A2	9/2014	WO	WO 2015/032494 A2	3/2015	
WO	WO 2014/144761 A2	9/2014	WO	WO 2015/033293 A1	3/2015	
WO	WO 2014/144951 A1	9/2014	WO	WO 2015/034872 A2	3/2015	
WO	WO 2014/145599 A2	9/2014	WO	WO 2015/034885 A1	3/2015	
WO	WO 2014/145736 A2	9/2014	WO	WO 2015/035136 A2	3/2015	
WO	WO 2014/150624 A1	9/2014	WO	WO 2015/035139 A2	3/2015	
WO	WO 2014/152432 A2	9/2014	WO	WO 2015/035162 A2	3/2015	
WO	WO 2014/152940 A1	9/2014	WO	WO 2015/040075 A1	3/2015	
WO	WO 2014/153118 A1	9/2014	WO	WO 2015/040402 A1	3/2015	
WO	WO 2014/153470 A2	9/2014	WO	WO 2015/042393 A2	3/2015	
WO	WO 2014/158593 A1	10/2014	WO	WO 2015/042585 A1	3/2015	
WO	WO 2014/161821 A1	10/2014	WO	WO 2015/048577 A2	4/2015	
WO	WO 2014/164466 A1	10/2014	WO	WO 2015/048690 A1	4/2015	
WO	WO 2014/165177 A1	10/2014	WO	WO 2015/048707 A2	4/2015	
WO	WO 2014/165349 A1	10/2014	WO	WO 2015/048801 A2	4/2015	
WO	WO 2014/165612 A2	10/2014	WO	WO 2015/049897 A1	4/2015	
WO	WO 2014/165707 A2	10/2014	WO	WO 2015/051191 A1	4/2015	
WO	WO 2014/165825 A2	10/2014	WO	WO 2015/052133 A1	4/2015	
WO	WO 2014/172458 A1	10/2014	WO	WO 2015/052231 A2	4/2015	
WO	WO 2014/172470 A2	10/2014	WO	WO 2015/052335 A1	4/2015	
WO	WO 2014/172489 A2	10/2014	WO	WO 2015/053995 A1	4/2015	
WO	WO 2014/173955 A1	10/2014	WO	WO 2015/054253 A1	4/2015	
WO	WO 2014/182700 A1	11/2014	WO	WO 2015/054315 A1	4/2015	
WO	WO 2014/183071 A2	11/2014	WO	WO 2015/057671 A1	4/2015	
WO	WO 2014/184143 A1	11/2014	WO	WO 2015/057834 A1	4/2015	
WO	WO 2014/184741 A1	11/2014	WO	WO 2015/057852 A1	4/2015	
WO	WO 2014/184744 A1	11/2014	WO	WO 2015/057976 A1	4/2015	
WO	WO 2014/186585 A2	11/2014	WO	WO 2015/057980 A1	4/2015	
WO	WO 2014/186686 A2	11/2014	WO	WO 2015/059265 A1	4/2015	
WO	WO 2014/190181 A1	11/2014	WO	WO 2015/065964 A1	5/2015	
WO	WO 2014/191128 A1	12/2014	WO	WO 2015/066119 A1	5/2015	
WO	WO 2014/191518 A1	12/2014	WO	WO 2015/066634 A2	5/2015	
WO	WO 2014/191521 A2	12/2014	WO	WO 2015/066636 A2	5/2015	
WO	WO 2014/191525 A1	12/2014	WO	WO 2015/066637 A1	5/2015	
WO	WO 2014/191527 A1	12/2014	WO	WO 2015/066638 A2	5/2015	
WO	WO 2014/193583 A2	12/2014	WO	WO 2015/066643 A1	5/2015	
WO	WO 2014/194190 A1	12/2014	WO	WO 2015/069682 A2	5/2015	
WO	WO 2014/197568 A2	12/2014	WO	WO 2015/070083 A1	5/2015	
WO	WO 2014/197748 A2	12/2014	WO	WO 2015/070193 A1	5/2015	
WO	WO 2014/199358 A1	12/2014	WO	WO 2015/070212 A1	5/2015	
WO	WO 2014/200659 A1	12/2014	WO	WO 2015/071474 A2	5/2015	
WO	WO 2014/201015 A2	12/2014	WO	WO 2015/073683 A2	5/2015	

(56)	References Cited							
FOREIGN PATENT DOCUMENTS								
WO	WO 2015/073867	A1	5/2015	WO	WO 2015/148760	A1	10/2015	
WO	WO 2015/073990	A1	5/2015	WO	WO 2015/148761	A1	10/2015	
WO	WO 2015/075056	A1	5/2015	WO	WO 2015/148860	A1	10/2015	
WO	WO 2015/075154	A2	5/2015	WO	WO 2015/148863	A2	10/2015	
WO	WO 2015/075175	A1	5/2015	WO	WO 2015/153760	A2	10/2015	
WO	WO 2015/075195	A1	5/2015	WO	WO 2015/153780	A1	10/2015	
WO	WO 2015/075557	A2	5/2015	WO	WO 2015/153789	A1	10/2015	
WO	WO 2015/077058	A2	5/2015	WO	WO 2015/153791	A1	10/2015	
WO	WO 2015/077290	A2	5/2015	WO	WO 2015/153889	A2	10/2015	
WO	WO 2015/077318	A1	5/2015	WO	WO 2015/153940	A1	10/2015	
WO	WO 2015/079056	A1	6/2015	WO	WO 2015/159068	A1	10/2015	
WO	WO 2015/079057	A2	6/2015	WO	WO 2015/159086	A1	10/2015	
WO	WO 2015/086795	A1	6/2015	WO	WO 2015/159087	A1	10/2015	
WO	WO 2015/086798	A2	6/2015	WO	WO 2015/160683	A1	10/2015	
WO	WO 2015/088643	A1	6/2015	WO	WO 2015/161276	A2	10/2015	
WO	WO 2015/089046	A1	6/2015	WO	WO 2015/163733	A1	10/2015	
WO	WO 2015/089077	A2	6/2015	WO	WO 2015/164740	A1	10/2015	
WO	WO 2015/089277	A1	6/2015	WO	WO 2015/164748	A1	10/2015	
WO	WO 2015/089351	A1	6/2015	WO	WO 2015/165274	A1	11/2015	
WO	WO 2015/089354	A1	6/2015	WO	WO 2015/165275	A1	11/2015	
WO	WO 2015/089364	A1	6/2015	WO	WO 2015/165276	A1	11/2015	
WO	WO 2015/089419	A2	6/2015	WO	WO 2015/166272	A2	11/2015	
WO	WO 2015/089427	A1	6/2015	WO	WO 2015/167766	A1	11/2015	
WO	WO 2015/089462	A1	6/2015	WO	WO 2015/167956	A1	11/2015	
WO	WO 2015/089465	A1	6/2015	WO	WO 2015/168125	A1	11/2015	
WO	WO 2015/089473	A1	6/2015	WO	WO 2015/168158	A1	11/2015	
WO	WO 2015/089486	A1	6/2015	WO	WO 2015/168404	A1	11/2015	
WO	WO 2015/095804	A1	6/2015	WO	WO 2015/168547	A2	11/2015	
WO	WO-2015089406	A1 *	6/2015 A61K 38/465	WO	WO 2015/168800	A1	11/2015
WO	WO 2015/099850	A1	7/2015	WO	WO 2015/171603	A1	11/2015	
WO	WO 2015/100929	A1	7/2015	WO	WO 2015/171894	A1	11/2015	
WO	WO 2015/103057	A1	7/2015	WO	WO 2015/171932	A1	11/2015	
WO	WO 2015/103153	A1	7/2015	WO	WO 2015/172128	A1	11/2015	
WO	WO 2015/105928	A1	7/2015	WO	WO 2015/173436	A1	11/2015	
WO	WO 2015/108993	A1	7/2015	WO	WO 2015/175642	A2	11/2015	
WO	WO 2015/109752	A1	7/2015	WO	WO 2015/179540	A1	11/2015	
WO	WO 2015/110474	A1	7/2015	WO	WO 2015/183025	A1	12/2015	
WO	WO 2015/112790	A2	7/2015	WO	WO 2015/183026	A1	12/2015	
WO	WO 2015/112896	A2	7/2015	WO	WO 2015/183885	A1	12/2015	
WO	WO 2015/113063	A1	7/2015	WO	WO 2015/184259	A1	12/2015	
WO	WO 2015/114365	A1	8/2015	WO	WO 2015/184262	A1	12/2015	
WO	WO 2015/115903	A1	8/2015	WO	WO 2015/184268	A1	12/2015	
WO	WO 2015/116686	A1	8/2015	WO	WO 2015/188056	A1	12/2015	
WO	WO 2015/116969	A2	8/2015	WO	WO 2015/188065	A1	12/2015	
WO	WO 2015/117021	A1	8/2015	WO	WO 2015/188094	A1	12/2015	
WO	WO 2015/117041	A1	8/2015	WO	WO 2015/188109	A1	12/2015	
WO	WO 2015/117081	A2	8/2015	WO	WO 2015/188132	A1	12/2015	
WO	WO 2015/118156	A1	8/2015	WO	WO 2015/188135	A1	12/2015	
WO	WO 2015/119941	A2	8/2015	WO	WO 2015/188191	A1	12/2015	
WO	WO 2015/121454	A1	8/2015	WO	WO 2015/189693	A1	12/2015	
WO	WO 2015/122967	A1	8/2015	WO	WO 2015/191693	A2	12/2015	
WO	WO 2015/123339	A1	8/2015	WO	WO 2015/191899	A1	12/2015	
WO	WO 2015/124715	A1	8/2015	WO	WO 2015/191911	A2	12/2015	
WO	WO 2015/124718	A1	8/2015	WO	WO 2015/193858	A1	12/2015	
WO	WO 2015/126927	A2	8/2015	WO	WO 2015/195547	A1	12/2015	
WO	WO 2015/127428	A1	8/2015	WO	WO 2015/195621	A1	12/2015	
WO	WO 2015/127439	A1	8/2015	WO	WO 2015/195798	A1	12/2015	
WO	WO 2015/129686	A1	9/2015	WO	WO 2015/198020	A1	12/2015	
WO	WO 2015/131101	A1	9/2015	WO	WO 2015/200334	A1	12/2015	
WO	WO 2015/133554	A1	9/2015	WO	WO 2015/200378	A1	12/2015	
WO	WO 2015/134121	A2	9/2015	WO	WO 2015/200555	A2	12/2015	
WO	WO 2015/134812	A1	9/2015	WO	WO 2015/200805	A2	12/2015	
WO	WO 2015/136001	A1	9/2015	WO	WO 2016/001978	A1	1/2016	
WO	WO 2015/138510	A1	9/2015	WO	WO 2016/004010	A1	1/2016	
WO	WO 2015/138739	A2	9/2015	WO	WO 2016/004318	A1	1/2016	
WO	WO 2015/138855	A1	9/2015	WO	WO 2016/007347	A1	1/2016	
WO	WO 2015/138870	A2	9/2015	WO	WO 2016/007604	A1	1/2016	
WO	WO 2015/139008	A1	9/2015	WO	WO 2016/007948	A1	1/2016	
WO	WO 2015/139139	A1	9/2015	WO	WO 2016/011080	A2	1/2016	
WO	WO 2015/143046	A2	9/2015	WO	WO 2016/011210	A2	1/2016	
WO	WO 2015/143177	A1	9/2015	WO	WO 2016/011428	A1	1/2016	
WO	WO 2015/145417	A1	10/2015	WO	WO 2016/012544	A2	1/2016	
WO	WO 2015/148431	A1	10/2015	WO	WO 2016/012552	A1	1/2016	
WO	WO 2015/148670	A1	10/2015	WO	WO 2016/014409	A1	1/2016	
WO	WO 2015/148680	A1	10/2015	WO	WO 2016/014565	A2	1/2016	

(56)	References Cited					
FOREIGN PATENT DOCUMENTS						
WO	WO 2016/014794 A1	1/2016	WO	WO 2016/084088 A1	6/2016	
WO	WO 2016/014837 A1	1/2016	WO	WO 2016/086177 A2	6/2016	
WO	WO 2016/016119 A1	2/2016	WO	WO 2016/089433 A1	6/2016	
WO	WO 2016/016358 A1	2/2016	WO	WO 2016/094845 A2	6/2016	
WO	WO 2016/019144 A2	2/2016	WO	WO 2016/094867 A1	6/2016	
WO	WO 2016/020399 A1	2/2016	WO	WO 2016/094872 A1	6/2016	
WO	WO 2016/021972 A1	2/2016	WO	WO 2016/094874 A1	6/2016	
WO	WO 2016/021973 A1	2/2016	WO	WO 2016/094880 A1	6/2016	
WO	WO 2016/022363 A2	2/2016	WO	WO 2016/094888 A1	6/2016	
WO	WO 2016/022866 A1	2/2016	WO	WO 2016/097212 A1	6/2016	
WO	WO 2016/022931 A1	2/2016	WO	WO 2016/097231 A2	6/2016	
WO	WO 2016/025131 A1	2/2016	WO	WO 2016/097751 A1	6/2016	
WO	WO 2016/025469 A1	2/2016	WO	WO 2016/099887 A1	6/2016	
WO	WO 2016/025759 A1	2/2016	WO	WO 2016/100272 A1	6/2016	
WO	WO 2016/026444 A1	2/2016	WO	WO 2016/100389 A1	6/2016	
WO	WO 2016/028682 A1	2/2016	WO	WO 2016/100568 A1	6/2016	
WO	WO 2016/028843 A1	2/2016	WO	WO 2016/100571 A1	6/2016	
WO	WO 2016/028887 A1	2/2016	WO	WO 2016/100951 A2	6/2016	
WO	WO 2016/033088 A1	3/2016	WO	WO 2016/100955 A2	6/2016	
WO	WO 2016/033230 A1	3/2016	WO	WO 2016/100974 A1	6/2016	
WO	WO 2016/033246 A1	3/2016	WO	WO 2016/103233 A2	6/2016	
WO	WO 2016/033298 A1	3/2016	WO	WO 2016/104716 A1	6/2016	
WO	WO 2016/035044 A1	3/2016	WO	WO 2016/106236 A1	6/2016	
WO	WO 2016/035918 A1	3/2016	WO	WO 2016/106239 A1	6/2016	
WO	WO 2016/036754 A1	3/2016	WO	WO 2016/106244 A1	6/2016	
WO	WO 2016/037157 A2	3/2016	WO	WO 2016/106338 A2	6/2016	
WO	WO 2016/040030 A1	3/2016	WO	WO 2016/108926 A1	7/2016	
WO	WO 2016/040594 A1	3/2016	WO	WO 2016/109255 A1	7/2016	
WO	WO 2016/044182 A1	3/2016	WO	WO 2016/109840 A2	7/2016	
WO	WO 2016/044416 A1	3/2016	WO	WO 2016/110214 A1	7/2016	
WO	WO 2016/046635 A1	3/2016	WO	WO 2016/110453 A1	7/2016	
WO	WO 2016/049024 A2	3/2016	WO	WO 2016/110511 A1	7/2016	
WO	WO 2016/049163 A2	3/2016	WO	WO 2016/110512 A1	7/2016	
WO	WO 2016/049230 A1	3/2016	WO	WO 2016/111546 A2	7/2016	
WO	WO 2016/049251 A1	3/2016	WO	WO 2016/112242 A1	7/2016	
WO	WO 2016/049258 A2	3/2016	WO	WO 2016/112351 A1	7/2016	
WO	WO 2016/053397 A2	4/2016	WO	WO 2016/112963 A1	7/2016	
WO	WO 2016/054326 A1	4/2016	WO	WO 2016/113357 A1	7/2016	
WO	WO 2016/057061 A2	4/2016	WO	WO 2016/114972 A1	7/2016	
WO	WO 2016/057821 A2	4/2016	WO	WO 2016/115179 A1	7/2016	
WO	WO 2016/057835 A2	4/2016	WO	WO 2016/115326 A1	7/2016	
WO	WO 2016/057850 A1	4/2016	WO	WO 2016/115355 A1	7/2016	
WO	WO 2016/057951 A2	4/2016	WO	WO 2016/116032 A1	7/2016	
WO	WO 2016/057961 A1	4/2016	WO	WO 2016/120480 A1	8/2016	
WO	WO 2016/061073 A1	4/2016	WO	WO 2016/123071 A1	8/2016	
WO	WO 2016/061374 A1	4/2016	WO	WO 2016/123230 A1	8/2016	
WO	WO 2016/061481 A1	4/2016	WO	WO 2016/123243 A1	8/2016	
WO	WO 2016/061523 A1	4/2016	WO	WO 2016/123578 A1	8/2016	
WO	WO 2016/064894 A2	4/2016	WO	WO 2016/126747 A1	8/2016	
WO	WO 2016/065364 A1	4/2016	WO	WO 2016/130600 A2	8/2016	
WO	WO 2016/069282 A1	5/2016	WO	WO 2016/130697 A1	8/2016	
WO	WO 2016/069283 A1	5/2016	WO	WO 2016/131009 A1	8/2016	
WO	WO 2016/069591 A2	5/2016	WO	WO 2016/132122 A1	8/2016	
WO	WO 2016/069774 A1	5/2016	WO	WO 2016/133165 A1	8/2016	
WO	WO 2016/069910 A1	5/2016	WO	WO 2016/135507 A1	9/2016	
WO	WO 2016/069912 A1	5/2016	WO	WO 2016/135557 A2	9/2016	
WO	WO 2016/070037 A2	5/2016	WO	WO 2016/135558 A2	9/2016	
WO	WO 2016/070070 A1	5/2016	WO	WO 2016/135559 A2	9/2016	
WO	WO 2016/070129 A1	5/2016	WO	WO 2016/137774 A1	9/2016	
WO	WO 2016/072399 A1	5/2016	WO	WO 2016/137949 A1	9/2016	
WO	WO 2016/072936 A1	5/2016	WO	WO 2016/141224 A1	9/2016	
WO	WO 2016/073433 A1	5/2016	WO	WO 2016/141893 A1	9/2016	
WO	WO 2016/073559 A1	5/2016	WO	WO 2016/142719 A1	9/2016	
WO	WO 2016/073990 A2	5/2016	WO	WO 2016/145150 A2	9/2016	
WO	WO 2016/075662 A2	5/2016	WO	WO 2016/148994 A1	9/2016	
WO	WO 2016/076672 A1	5/2016	WO	WO 2016/149484 A2	9/2016	
WO	WO 2016/077273 A1	5/2016	WO	WO 2016/149547 A1	9/2016	
WO	WO 2016/077350 A1	5/2016	WO	WO 2016/150336 A1	9/2016	
WO	WO 2016/080097 A1	5/2016	WO	WO 2016/150855 A1	9/2016	
WO	WO 2016/080795 A1	5/2016	WO	WO 2016/154016 A2	9/2016	
WO	WO 2016/081923 A2	5/2016	WO	WO 2016/154579 A2	9/2016	
WO	WO 2016/081924 A1	5/2016	WO	WO 2016/154596 A1	9/2016	
WO	WO 2016/082135 A1	6/2016	WO	WO 2016/155482 A1	10/2016	
WO	WO 2016/083811 A1	6/2016	WO	WO 2016/161004 A1	10/2016	
WO	WO 2016/084084 A1	6/2016	WO	WO 2016/161207 A1	10/2016	

(56)	References Cited						
FOREIGN PATENT DOCUMENTS							
WO	WO 2016/161260 A1	10/2016	WO	WO 2016/205688 A2	12/2016		
WO	WO 2016/161380 A1	10/2016	WO	WO 2016/205703 A1	12/2016		
WO	WO 2016/161446 A1	10/2016	WO	WO 2016/205711 A1	12/2016		
WO	WO 2016/164305 A1	10/2016	WO	WO 2016/205728 A1	12/2016		
WO	WO 2016/164356 A1	10/2016	WO	WO 2016/205745 A2	12/2016		
WO	WO 2016/164797 A1	10/2016	WO	WO 2016/205749 A1	12/2016		
WO	WO 2016/166340 A1	10/2016	WO	WO 2016/205759 A1	12/2016		
WO	WO 2016/167300 A1	10/2016	WO	WO 2016/205764 A1	12/2016		
WO	WO 2016/168631 A1	10/2016	WO	WO 2017/001572 A1	1/2017		
WO	WO 2016/170484 A1	10/2016	WO	WO 2017/001988 A1	1/2017		
WO	WO 2016/172359 A2	10/2016	WO	WO 2017/004261 A1	1/2017		
WO	WO 2016/172727 A1	10/2016	WO	WO 2017/004279 A2	1/2017		
WO	WO 2016/174056 A1	11/2016	WO	WO 2017/004616 A1	1/2017		
WO	WO 2016/174151 A1	11/2016	WO	WO 2017/005807 A1	1/2017		
WO	WO 2016/174250 A1	11/2016	WO	WO 2017/009399 A1	1/2017		
WO	WO 2016/176191 A1	11/2016	WO	WO 2017/010556 A1	1/2017		
WO	WO 2016/176404 A1	11/2016	WO	WO 2017/011519 A1	1/2017		
WO	WO 2016/176690 A2	11/2016	WO	WO 2017/011721 A1	1/2017		
WO	WO 2016/177682 A1	11/2016	WO	WO 2017/011804 A1	1/2017		
WO	WO 2016/178207 A1	11/2016	WO	WO 2017/015015 A1	1/2017		
WO	WO 2016/179038 A1	11/2016	WO	WO 2017/015101 A1	1/2017		
WO	WO 2016/179112 A1	11/2016	WO	WO 2017/019867 A1	2/2017		
WO	WO 2016/181357 A1	11/2016	WO	WO 2017/019895 A1	2/2017		
WO	WO 2016/182893 A1	11/2016	WO	WO 2017/023803 A1	2/2017		
WO	WO 2016/182917 A1	11/2016	WO	WO 2017/023974 A1	2/2017		
WO	WO 2016/182959 A1	11/2016	WO	WO 2017/024047 A1	2/2017		
WO	WO 2016/183236 A1	11/2016	WO	WO 2017/024319 A1	2/2017		
WO	WO 2016/183298 A2	11/2016	WO	WO 2017/024343 A1	2/2017		
WO	WO 2016/183345 A1	11/2016	WO	WO 2017/024602 A1	2/2017		
WO	WO 2016/183402 A2	11/2016	WO	WO 2017/025323 A1	2/2017		
WO	WO 2016/183438 A1	11/2016	WO	WO 2017/027423 A1	2/2017		
WO	WO 2016/183448 A1	11/2016	WO	WO 2017/028768 A1	2/2017		
WO	WO 2016/184955 A2	11/2016	WO	WO 2017/029664 A1	2/2017		
WO	WO 2016/184989 A1	11/2016	WO	WO 2017/031360 A1	2/2017		
WO	WO 2016/185411 A1	11/2016	WO	WO 2017/031483 A1	2/2017		
WO	WO 2016/186745 A1	11/2016	WO	WO 2017/035416 A2	3/2017		
WO	WO 2016/186772 A2	11/2016	WO	WO 2017/040348 A1	3/2017		
WO	WO 2016/186946 A1	11/2016	WO	WO 2017/040511 A1	3/2017		
WO	WO 2016/186953 A1	11/2016	WO	WO 2017/040709 A1	3/2017		
WO	WO 2016/187717 A1	12/2016	WO	WO 2017/040786 A1	3/2017		
WO	WO 2016/187904 A1	12/2016	WO	WO 2017/040793 A1	3/2017		
WO	WO 2016/191684 A1	12/2016	WO	WO 2017/040813 A2	3/2017		
WO	WO 2016/191869 A1	12/2016	WO	WO 2017/043573 A1	3/2017		
WO	WO 2016/196273 A1	12/2016	WO	WO 2017/043656 A1	3/2017		
WO	WO 2016/196282 A1	12/2016	WO	WO 2017/044419 A1	3/2017		
WO	WO 2016/196308 A1	12/2016	WO	WO 2017/044776 A1	3/2017		
WO	WO 2016/196361 A1	12/2016	WO	WO 2017/044857 A2	3/2017		
WO	WO 2016/196499 A1	12/2016	WO	WO 2017/048390 A1	3/2017		
WO	WO 2016/196539 A2	12/2016	WO	WO 2017/049129 A2	3/2017		
WO	WO 2016/196655 A1	12/2016	WO	WO 2017/050963 A1	3/2017		
WO	WO 2016/196805 A1	12/2016	WO	WO 2017/053312 A1	3/2017		
WO	WO 2016/196887 A1	12/2016	WO	WO 2017/053431 A2	3/2017		
WO	WO 2016/197132 A1	12/2016	WO	WO 2017/053713 A1	3/2017		
WO	WO 2016/197133 A1	12/2016	WO	WO 2017/053729 A1	3/2017		
WO	WO 2016/197354 A1	12/2016	WO	WO 2017/053753 A1	3/2017		
WO	WO 2016/197355 A1	12/2016	WO	WO 2017/053762 A1	3/2017		
WO	WO 2016/197356 A1	12/2016	WO	WO 2017/053879 A1	3/2017		
WO	WO 2016/197357 A1	12/2016	WO	WO 2017/054721 A1	4/2017		
WO	WO 2016/197358 A1	12/2016	WO	WO 2017/058658 A2	4/2017		
WO	WO 2016/197359 A1	12/2016	WO	WO 2017/059241 A1	4/2017		
WO	WO 2016/197360 A1	12/2016	WO	WO 2017/062605 A1	4/2017		
WO	WO 2016/197361 A1	12/2016	WO	WO 2017/062723 A1	4/2017		
WO	WO 2016/197362 A1	12/2016	WO	WO 2017/062754 A1	4/2017		
WO	WO 2016/198361 A1	12/2016	WO	WO 2017/062855 A1	4/2017		
WO	WO 2016/198500 A1	12/2016	WO	WO 2017/062886 A1	4/2017		
WO	WO 2016/200263 A1	12/2016	WO	WO 2017/062983 A1	4/2017		
WO	WO 2016/201047 A1	12/2016	WO	WO 2017/064439 A1	4/2017		
WO	WO 2016/201138 A1	12/2016	WO	WO 2017/064546 A1	4/2017		
WO	WO 2016/201152 A1	12/2016	WO	WO 2017/064566 A2	4/2017		
WO	WO 2016/201153 A1	12/2016	WO	WO 2017/066175 A1	4/2017		
WO	WO 2016/201155 A1	12/2016	WO	WO 2017/066497 A2	4/2017		
WO	WO 2016/205276 A1	12/2016	WO	WO 2017/066588 A2	4/2017		
WO	WO 2016/205613 A1	12/2016	WO	WO 2017/066707 A1	4/2017		
WO	WO 2016/205623 A1	12/2016	WO	WO 2017/066781 A1	4/2017		
WO	WO 2016/205680 A1	12/2016	WO	WO 2017/068077 A1	4/2017		

(56)	References Cited					
FOREIGN PATENT DOCUMENTS						
WO	WO 2017/068377 A1	4/2017	WO	WO 2017/142923 A1	8/2017	
WO	WO 2017/069829 A2	4/2017	WO	WO 2017/142999 A2	8/2017	
WO	WO 2017/070029 A1	4/2017	WO	WO 2017/143042 A2	8/2017	
WO	WO 2017/070032 A1	4/2017	WO	WO 2017/147056 A1	8/2017	
WO	WO 2017/070169 A1	4/2017	WO	WO 2017/147278 A1	8/2017	
WO	WO 2017/070284 A1	4/2017	WO	WO 2017/147432 A1	8/2017	
WO	WO 2017/070598 A1	4/2017	WO	WO 2017/147446 A1	8/2017	
WO	WO 2017/070605 A1	4/2017	WO	WO 2017/147555 A1	8/2017	
WO	WO 2017/070632 A2	4/2017	WO	WO 2017/151444 A1	9/2017	
WO	WO 2017/070633 A2	4/2017	WO	WO 2017/151719 A1	9/2017	
WO	WO 2017/072590 A1	5/2017	WO	WO 2017/152015 A1	9/2017	
WO	WO 2017/074526 A1	5/2017	WO	WO 2017/155717 A1	9/2017	
WO	WO 2017/074962 A1	5/2017	WO	WO 2017/157422 A1	9/2017	
WO	WO 2017/075261 A1	5/2017	WO	WO 2017/158153 A1	9/2017	
WO	WO 2017/075335 A1	5/2017	WO	WO 2017/160689 A1	9/2017	
WO	WO 2017/075475 A1	5/2017	WO	WO 2017/160752 A1	9/2017	
WO	WO 2017/077135 A1	5/2017	WO	WO 2017/160890 A1	9/2017	
WO	WO 2017/077329 A2	5/2017	WO	WO 2017/161068 A1	9/2017	
WO	WO 2017/078751 A1	5/2017	WO	WO 2017/165741 A1	9/2017	
WO	WO 2017/079400 A1	5/2017	WO	WO 2017/165826 A1	9/2017	
WO	WO 2017/079428 A1	5/2017	WO	WO 2017/165862 A1	9/2017	
WO	WO 2017/079673 A1	5/2017	WO	WO 2017/167712 A1	10/2017	
WO	WO 2017/079724 A1	5/2017	WO	WO 2017/172644 A2	10/2017	
WO	WO 2017/081097 A1	5/2017	WO	WO 2017/172645 A2	10/2017	
WO	WO 2017/081288 A1	5/2017	WO	WO 2017/172860 A1	10/2017	
WO	WO 2017/083368 A1	5/2017	WO	WO 2017/173004 A1	10/2017	
WO	WO 2017/083722 A1	5/2017	WO	WO 2017/173054 A1	10/2017	
WO	WO 2017/083766 A1	5/2017	WO	WO 2017/173092 A1	10/2017	
WO	WO 2017/087395 A1	5/2017	WO	WO 2017/174329 A1	10/2017	
WO	WO 2017/090724 A1	6/2017	WO	WO 2017/176529 A1	10/2017	
WO	WO 2017/091510 A1	6/2017	WO	WO 2017/176806 A1	10/2017	
WO	WO 2017/091630 A1	6/2017	WO	WO 2017/178590 A1	10/2017	
WO	WO 2017/092201 A1	6/2017	WO	WO 2017/180694 A1	10/2017	
WO	WO 2017/093370 A1	6/2017	WO	WO 2017/180711 A1	10/2017	
WO	WO 2017/093969 A1	6/2017	WO	WO 2017/180915 A2	10/2017	
WO	WO 2017/095111 A1	6/2017	WO	WO 2017/180926 A1	10/2017	
WO	WO 2017/096041 A1	6/2017	WO	WO 2017/181107 A2	10/2017	
WO	WO 2017/096237 A1	6/2017	WO	WO 2017/181735 A2	10/2017	
WO	WO 2017/100158 A1	6/2017	WO	WO 2017/182468 A1	10/2017	
WO	WO 2017/100431 A2	6/2017	WO	WO 2017/182585 A1	10/2017	
WO	WO 2017/104404 A1	6/2017	WO	WO 2017/182607 A1	10/2017	
WO	WO 2017/105251 A1	6/2017	WO	WO 2017/184334 A1	10/2017	
WO	WO 2017/105350 A1	6/2017	WO	WO 2017/184768 A1	10/2017	
WO	WO 2017/105991 A1	6/2017	WO	WO 2017/184786 A1	10/2017	
WO	WO 2017/106414 A1	6/2017	WO	WO 2017/186550 A1	11/2017	
WO	WO 2017/106528 A2	6/2017	WO	WO 2017/189308 A1	11/2017	
WO	WO 2017/106537 A2	6/2017	WO	WO 2017/189336 A1	11/2017	
WO	WO 2017/106569 A1	6/2017	WO	WO 2017/190041 A1	11/2017	
WO	WO 2017/106616 A1	6/2017	WO	WO 2017/190257 A1	11/2017	
WO	WO 2017/106657 A1	6/2017	WO	WO 2017/190664 A1	11/2017	
WO	WO 2017/106767 A1	6/2017	WO	WO 2017/191210 A1	11/2017	
WO	WO 2017/109134 A1	6/2017	WO	WO 2017/191274 A2	11/2017	
WO	WO 2017/109757 A1	6/2017	WO	WO 2017/192172 A1	11/2017	
WO	WO 2017/112620 A1	6/2017	WO	WO 2017/192512 A2	11/2017	
WO	WO 2017/115268 A1	7/2017	WO	WO 2017/192544 A1	11/2017	
WO	WO 2017/117395 A1	7/2017	WO	WO 2017/192573 A1	11/2017	
WO	WO 2017/118598 A1	7/2017	WO	WO 2017/193029 A2	11/2017	
WO	WO 2017/118720 A1	7/2017	WO	WO 2017/193053 A1	11/2017	
WO	WO 2017/123609 A1	7/2017	WO	WO 2017/196768 A1	11/2017	
WO	WO 2017/123910 A1	7/2017	WO	WO 2017/197038 A1	11/2017	
WO	WO 2017/124086 A1	7/2017	WO	WO 2017/197238 A1	11/2017	
WO	WO 2017/124100 A1	7/2017	WO	WO 2017/197301 A1	11/2017	
WO	WO 2017/124652 A1	7/2017	WO	WO 2017/201476 A1	11/2017	
WO	WO 2017/126987 A1	7/2017	WO	WO 2017/205290 A1	11/2017	
WO	WO 2017/127807 A1	7/2017	WO	WO 2017/205423 A1	11/2017	
WO	WO 2017/131237 A1	8/2017	WO	WO 2017/207589 A1	12/2017	
WO	WO 2017/132112 A1	8/2017	WO	WO 2017/208247 A1	12/2017	
WO	WO 2017/132580 A2	8/2017	WO	WO 2017/209809 A1	12/2017	
WO	WO 2017/136520 A1	8/2017	WO	WO 2017/213896 A1	12/2017	
WO	WO 2017/136629 A1	8/2017	WO	WO 2017/213898 A2	12/2017	
WO	WO 2017/136794 A1	8/2017	WO	WO 2017/214460 A1	12/2017	
WO	WO 2017/139264 A1	8/2017	WO	WO 2017/216392 A1	12/2017	
WO	WO 2017/139505 A2	8/2017	WO	WO 2017/216771 A2	12/2017	
WO	WO 2017/141173 A2	8/2017	WO	WO 2017/218185 A1	12/2017	
WO	WO 2017/142835 A1	8/2017	WO	WO 2017/219027 A1	12/2017	
			WO	WO 2017/219033 A1	12/2017	
			WO	WO 2017/220751 A1	12/2017	
			WO	WO 2017/222370 A1	12/2017	
			WO	WO 2017/222773 A1	12/2017	

(56)	References Cited					
FOREIGN PATENT DOCUMENTS						
WO	WO 2017/222834 A1	12/2017	WO	WO 2018/099256 A1	6/2018	
WO	WO 2017/223107 A1	12/2017	WO	WO 2018/103686 A1	6/2018	
WO	WO 2017/223330 A1	12/2017	WO	WO 2018/106268 A1	6/2018	
WO	WO 2018/000657 A1	1/2018	WO	WO 2018/107028 A1	6/2018	
WO	WO 2018/002719 A1	1/2018	WO	WO 2018/107103 A1	6/2018	
WO	WO 2018/005117 A1	1/2018	WO	WO 2018/107129 A1	6/2018	
WO	WO 2018/005289 A2	1/2018	WO	WO 2018/108272 A1	6/2018	
WO	WO 2018/005691 A1	1/2018	WO	WO 2018/109101 A1	6/2018	
WO	WO 2018/005782 A1	1/2018	WO	WO 2018/111946 A1	6/2018	
WO	WO 2018/005873 A1	1/2018	WO	WO 2018/111947 A1	6/2018	
WO	WO 2018/06693 A1	1/2018	WO	WO 2018/112336 A1	6/2018	
WO	WO 2018/009520 A1	1/2018	WO	WO 2018/112446 A2	6/2018	
WO	WO 2018/009562 A1	1/2018	WO	WO 2018/119354 A1	6/2018	
WO	WO 2018/009822 A1	1/2018	WO	WO 2018/119359 A1	6/2018	
WO	WO 2018/013821 A1	1/2018	WO	WO 2018/120283 A1	7/2018	
WO	WO 2018/013932 A1	1/2018	WO	WO 2018/130830 A1	7/2018	
WO	WO 2018/013990 A1	1/2018	WO	WO 2018/135838 A2	7/2018	
WO	WO 2018/014384 A1	1/2018	WO	WO 2018/136396 A2	7/2018	
WO	WO 2018/015444 A1	1/2018	WO	WO 2018/138385 A1	8/2018	
WO	WO 2018/015936 A2	1/2018	WO	WO 2018/142364 A1	8/2018	
WO	WO 2018/017754 A1	1/2018	WO	WO 2018/148246 A1	8/2018	
WO	WO 2018/018979 A1	2/2018	WO	WO 2018/148256 A1	8/2018	
WO	WO 2018/020248 A1	2/2018	WO	WO 2018/148647 A2	8/2018	
WO	WO 2018/021878 A1	2/2018	WO	WO 2018/149418 A1	8/2018	
WO	WO 2018/022480 A1	2/2018	WO	WO 2018/149888 A1	8/2018	
WO	WO 2018/022634 A1	2/2018	WO	WO 2018/149915 A1	8/2018	
WO	WO 2018/025206 A1	2/2018	WO	WO 2018/152197 A1	8/2018	
WO	WO 2018/026723 A1	2/2018	WO	WO 2018/152418 A1	8/2018	
WO	WO 2018/026976 A1	2/2018	WO	WO 2018/154380 A1	8/2018	
WO	WO 2018/027078 A1	2/2018	WO	WO 2018/154387 A1	8/2018	
WO	WO 2018/030608 A1	2/2018	WO	WO 2018/154412 A1	8/2018	
WO	WO 2018/031683 A1	2/2018	WO	WO 2018/154413 A1	8/2018	
WO	WO 2018/035250 A1	2/2018	WO	WO 2018/154418 A1	8/2018	
WO	WO 2018/035300 A1	2/2018	WO	WO 2018/154439 A1	8/2018	
WO	WO 2018/035423 A1	2/2018	WO	WO 2018/154459 A1	8/2018	
WO	WO 2018/035503 A1	2/2018	WO	WO 2018/154462 A2	8/2018	
WO	WO 2018/039145 A1	3/2018	WO	WO 2018/156372 A1	8/2018	
WO	WO 2018/039438 A1	3/2018	WO	WO 2018/156824 A1	8/2018	
WO	WO 2018/039440 A1	3/2018	WO	WO 2018/161009 A1	9/2018	
WO	WO 2018/039448 A1	3/2018	WO	WO 2018/161032 A1	9/2018	
WO	WO 2018/045630 A1	3/2018	WO	WO 2018/165504 A1	9/2018	
WO	WO 2018/048827 A1	3/2018	WO	WO 2018/165629 A1	9/2018	
WO	WO 2018/049073 A1	3/2018	WO	WO 2018/170015 A1	9/2018	
WO	WO 2018/049168 A1	3/2018	WO	WO 2018/170340 A1	9/2018	
WO	WO 2018/051347 A1	3/2018	WO	WO 2018/175502 A2	9/2018	
WO	WO 2018/058064 A1	3/2018	WO	WO 2018/176009 A1	9/2018	
WO	WO 2018/062866 A2	4/2018	WO	WO 2018/177351 A1	10/2018	
WO	WO 2018/064352 A1	4/2018	WO	WO 2018/179578 A1	10/2018	
WO	WO 2018/064371 A1	4/2018	WO	WO 2018/183403 A1	10/2018	
WO	WO 2018/064516 A1	4/2018	WO	WO 2018/189184 A1	10/2018	
WO	WO 2018/067546 A1	4/2018	WO	WO 2018/191388 A1	10/2018	
WO	WO 2018/067846 A1	4/2018	WO	WO 2018/195402 A1	10/2018	
WO	WO 2018/068053 A2	4/2018	WO	WO 2018/195545 A2	10/2018	
WO	WO 2018/069474 A1	4/2018	WO	WO 2018/195555 A1	10/2018	
WO	WO 2018/071623 A2	4/2018	WO	WO 2018/197020 A1	11/2018	
WO	WO 2018/071663 A1	4/2018	WO	WO 2018/197495 A1	11/2018	
WO	WO 2018/071868 A1	4/2018	WO	WO 2018/200597 A1	11/2018	
WO	WO 2018/071892 A1	4/2018	WO	WO 2018/202800 A1	11/2018	
WO	WO 2018/074979 A1	4/2018	WO	WO 2018/204493 A1	11/2018	
WO	WO 2018/079134 A1	5/2018	WO	WO 2018/208755 A1	11/2018	
WO	WO 2018/080573 A1	5/2018	WO	WO 2018/208998 A1	11/2018	
WO	WO 2018/081504 A1	5/2018	WO	WO 2018/209158 A2	11/2018	
WO	WO 2018/081535 A2	5/2018	WO	WO 2018/209320 A1	11/2018	
WO	WO 2018/081728 A1	5/2018	WO	WO 2018/213351 A1	11/2018	
WO	WO 2018/083128 A2	5/2018	WO	WO 2018/213708 A1	11/2018	
WO	WO 2018/083606 A1	5/2018	WO	WO 2018/213726 A1	11/2018	
WO	WO 2018/085288 A1	5/2018	WO	WO 2018/213771 A1	11/2018	
WO	WO 2018/085414 A1	5/2018	WO	WO 2018/213791 A1	11/2018	
WO	WO 2018/085842 A1	5/2018	WO	WO 2018/217852 A1	11/2018	
WO	WO 2018/086623 A1	5/2018	WO	WO 2018/217981 A1	11/2018	
WO	WO 2018/089664 A1	5/2018	WO	WO 2018/218166 A1	11/2018	
WO	WO 2018/093990 A1	5/2018	WO	WO 2018/218188 A2	11/2018	
WO	WO 2018/098383 A1	5/2018	WO	WO 2018/218206 A1	11/2018	
WO	WO 2018/098480 A1	5/2018	WO	WO 2018/226855 A1	12/2018	
WO	WO 2018/098587 A1	6/2018	WO	WO 2019/005884 A1	1/2019	
			WO	WO 2019/005886 A1	1/2019	
			WO	WO 2019/010384 A1	1/2019	
			WO	WO 2019/023680 A1	1/2019	
			WO	WO 2019/042284 A1	3/2019	

(56)	References Cited				
FOREIGN PATENT DOCUMENTS					
WO	WO 2019/051097 A1	3/2019	WO	WO 2023/102550 A2	6/2023
WO	WO 2019/067992 A1	4/2019	WO	WO 2023/173140 A2	9/2023
WO	WO 2019/075357 A1	4/2019	WO	WO 2024/155741 A1	7/2024
WO	WO 2019/079347 A1	4/2019	WO	WO 2024/155745 A1	7/2024
WO	WO 2019/084062 A1	5/2019	WO	WO 2024/215652 A2	10/2024
WO	WO 2019/090169 A1	5/2019	OTHER PUBLICATIONS		
WO	WO 2019/090367 A1	5/2019	Bae et al., "Heteroclitic CD33 Peptide With Enhanced Anti-Acute		
WO	WO 2019/092042 A1	5/2019	Myeloid Leukemic Immunogenicity" 10 Clinical Cancer Research		
WO	WO 2019/118497 A1	6/2019	7043-7052 (Year: 2004).*		
WO	WO 2019/118935 A1	6/2019	Fikes et al., "Design of multi-epitope, analogue-based cancer vac-		
WO	WO 2019/118949 A1	6/2019	cines" 3(6) Expert Opinion On Biological Therapy 985-993 (Year:		
WO	WO 2019/123430 A1	6/2019	2003).*		
WO	WO 2019/126709 A1	6/2019	Houghton et al., "Immunological validation of the EpitOptimizer		
WO	WO 2019/139645 A2	7/2019	program for streamlined design of heteroclitic epitopes" 25 Vaccine		
WO	WO 2019/139951 A1	7/2019	5330-5342 (Year: 2007).*		
WO	WO 2019/147014 A1	8/2019	U.S. Appl. No. 61/716,256, filed Oct. 19, 2012, Jinek et al.		
WO	WO 2019/161251 A1	8/2019	U.S. Appl. No. 61/717,324, filed Oct. 23, 2012, Cho et al.		
WO	WO 2019/168953 A1	9/2019	U.S. Appl. No. 61/734,256, filed Dec. 6, 2012, Chen et al.		
WO	WO 2019/183641 A1	9/2019	U.S. Appl. No. 61/758,624, filed Jan. 30, 2013, Chen et al.		
WO	WO 2019/204369 A1	10/2019	U.S. Appl. No. 61/761,046, filed Feb. 5, 2013, Knight et al.		
WO	WO 2019/213257 A1	11/2019	U.S. Appl. No. 61/794,422, filed Mar. 15, 2013, Knight et al.		
WO	WO 2019/217942 A1	11/2019	U.S. Appl. No. 61/803,599, filed Mar. 20, 2013, Kim et al.		
WO	WO 2019/226593 A1	11/2019	U.S. Appl. No. 61/837,481, filed Jun. 20, 2013, Cho et al.		
WO	WO 2019/226953 A1	11/2019	U.S. Appl. No. 61/838,178, filed Jun. 21, 2013, Joung et al.		
WO	WO 2019/236566 A1	12/2019	U.S. Appl. No. 61/874,682, filed Sep. 6, 2013, Liu et al.		
WO	WO 2019/241649 A1	12/2019	U.S. Appl. No. 61/874,746, filed Sep. 6, 2013, Liu et al.		
WO	WO 2020/014261 A1	1/2020	U.S. Appl. No. 62/288,661, filed Jan. 29, 2016, Muir et al.		
WO	WO 2020/028555 A2	2/2020	U.S. Appl. No. 62/357,332, filed Jun. 30, 2016, Liu et al.		
WO	WO 2020/028823 A1	2/2020	International Search Report for PCT/US2018/021880, mailed Jun.		
WO	WO 2020/041751 A1	2/2020	20, 2018.		
WO	WO 2020/047124 A1	3/2020	International Preliminary Report on Patentability for PCT/US2018/		
WO	WO 2020/051360 A1	3/2020	021880, mailed on Sep. 19, 2019.		
WO	WO 2020/086908 A1	4/2020	[No Author Listed] HyPhy—Hypothesis testing using Phylogenies.		
WO	WO 2020/092453 A1	5/2020	Last modified Apr. 21, 2017. Accessed online via http://hyphy.org/w/index.php/Main_Page on Apr. 28, 2021.		
WO	WO 2020/102659 A1	5/2020	[No Author Listed] NCBI Accession No. XP_015843220.1. C ->U		
WO	WO 2020/154500 A1	7/2020	editing enzyme APOBEC-1 [Peromyscus maniculatus bairdii],		
WO	WO 2020/180975 A1	9/2020	XP002793540. Mar. 21, 2016.		
WO	WO 2020/181178 A1	9/2020	[No Author Listed] NCBI Accession No. XP_021505673.1. C ->U		
WO	WO 2020/181180 A1	9/2020	editing enzyme APOBEC-1 [Meriones unguiculatus], XP002793541.		
WO	WO 2020/181193 A1	9/2020	Jun. 27, 2017.		
WO	WO 2020/181195 A1	9/2020	[No Author Listed] NCBI Reference Sequence: WP_00087959824.		
WO	WO 2020/181202 A1	9/2020	1. Oct. 9, 2019. 2 pages.		
WO	WO 2020/191153 A1	9/2020	[No Author Listed] Score result for SEQ 355 to W02017032580.		
WO	WO 2020/191171 A1	9/2020	Muir et al. 2016.		
WO	WO 2020/191233 A1	9/2020	[No Author Listed], "Human genome." Encyclopedia Britannica.		
WO	WO 2020/191234 A1	9/2020	Encyclopedia Britannica, Inc. Published Feb. 15, 2019. Last accessed		
WO	WO 2020/191239 A1	9/2020	online via https://www.britannica.com/science/human-genome on		
WO	WO 2020/191241 A1	9/2020	Mar. 19, 2021. 2 pages.		
WO	WO 2020/191242 A1	9/2020	[No Author Listed], EMBL Accession No. Q99ZW2. Nov. 2012. 2		
WO	WO 2020/191243 A1	9/2020	pages.		
WO	WO 2020/191245 A1	9/2020	[No Author Listed], Invitrogen Lipofectamine™ 2000 product		
WO	WO 2020/191246 A1	9/2020	sheets, 2002. 2 pages.		
WO	WO 2020/191248 A1	9/2020	[No Author Listed], Invitrogen Lipofectamine™ 2000 product		
WO	WO 2020/191249 A1	9/2020	sheets, 2005. 3 pages.		
WO	WO 2020/210751 A1	10/2020	[No Author Listed], Invitrogen Lipofectamine™ LTX product sheets,		
WO	WO 2020/214842 A1	10/2020	2011. 4 pages.		
WO	WO 2020/236982 A1	11/2020	[No Author Listed], <i>Mus musculus</i> (Mouse). UniProtKB Accession		
WO	WO 2021/025750 A1	2/2021	No. P51908 (ABEC1_MOUSE). Oct. 1, 1996. 10 pages.		
WO	WO 2021/030666 A1	2/2021	[No Author Listed], Thermo Fisher Scientific—How Cationic Lipid		
WO	WO 2021/072328 A1	4/2021	Mediated Transfection Works, retrieved from the internet Aug. 27,		
WO	WO 2021/108717 A2	6/2021	2015. 2 pages.		
WO	WO 2021/138469 A1	7/2021	Abremski et al., Bacteriophage P1 site-specific recombination.		
WO	WO 2021/155065 A1	8/2021	Purification and properties of the Cre recombinase protein. J Biol		
WO	WO 2021/158921 A2	8/2021	Chem. Feb. 10, 1984;259(3):1509-14.		
WO	WO 2021/158995 A1	8/2021	Abudayyeh et al., C2c2 is a single-component programmable RNA-		
WO	WO 2021/158999 A1	8/2021	guided RNA-targeting CRISPR effector. Science Aug.		
WO	WO 2021/222318 A1	11/2021	2016;353(6299):aaf5573. DOI: 10.1126/science.aaf5573.		
WO	WO 2021/226558 A1	11/2021			
WO	WO 2021/252924 A1	12/2021			
WO	WO 2022/067130 A2	3/2022			
WO	WO 2022/150790 A2	7/2022			
WO	WO 2022/165262 A1	8/2022			
WO	WO 2023/015309 A2	2/2023			
WO	WO 2023/102537 A2	6/2023			
WO	WO 2023/102538 A1	6/2023			

(56)

References Cited**OTHER PUBLICATIONS**

- Abudayyeh et al., A cytosine deaminase for programmable single-base RNA editing. *Science*. Jul. 26, 2019;365(6451):382-386. doi: 10.1126/science.aax7063. Epub Jul. 11, 2019.
- Abudayyeh et al., RNA targeting with CRISPR-Cas13. *Nature*. Oct. 12, 2017;550(7675):280-284. doi: 10.1038/nature24049. Epub Oct. 4, 2017.
- Ada et al., Carbohydrate-protein conjugate vaccines. *Clin Microbiol Infect*. Feb. 2003;9(2):79-85. doi: 10.1046/j.1469-0691.2003.00530.x.
- Adamala et al., Programmable RNA-binding protein composed of repeats of a single modular unit. *Proc Natl Acad Sci U S A*. May 10, 2016;113(19):E2579-88. doi: 10.1073/pnas.1519368113. Epub Apr. 26, 2016.
- Adams et al., New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J Am Chem Soc*. May 29, 2002;124(21):6063-76. doi: 10.1021/ja017687n.
- Addgene Plasmid # 44246. pdCas9-humanized. 2017, Stanley Qi.
- Addgene Plasmid # 73021. PCMV-BE3, 2017, David Liu.
- Addgene Plasmid # 79620. pcDNA3.1_pCMV-nCas-PmCDA1-ugi pH1-gRNA(HPRT), 2017, Akihiko Kondo.
- Adli, The CRISPR tool kit for genome editing and beyond. *Nat Commun*. May 15, 2018;9(1):1911. doi: 10.1038/s41467-018-04252-2.
- Aguilo et al., Coordination of m(6)A mRNA Methylation and Gene Transcription by ZFP217 Regulates Pluripotency and Reprogramming. *Cell Stem Cell*. Dec. 3, 2015;17(6):689-704. doi: 10.1016/j.stem.2015.09.005. Epub Oct. 29, 2015.
- Ahmad et al., Antibody-mediated specific binding and cytotoxicity of liposome-entrapped doxorubicin to lung cancer cells in vitro. *Cancer Res*. Sep. 1, 1992;52(17):4817-20.
- Aihara et al., A conformational switch controls the DNA cleavage activity of lambda integrase. *Mol Cell*. Jul. 2003;12(1):187-98.
- Aik et al., Structure of human RNA ?-methyladenine demethylase ALKBH5 provides insights into its mechanisms of nucleic acid recognition and demethylation. *Nucleic Acids Res*. Apr. 2014;42(7):4741-54. doi: 10.1093/nar/gku085. Epub Jan. 30, 2014.
- Aird et al., Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun Biol*. May 31, 2018;1:54. doi: 10.1038/s42003-018-0054-2.
- Akcakaya et al., In vivo CRISPR editing with no detectable genome-wide off-target mutations. *Nature*. Sep. 2018;561(7723):416-419. doi: 10.1038/s41586-018-0500-9. Epub Sep. 12, 2018. PMID: 30209390; PMCID: PMC6194229.
- Akins et al., Mitochondrial plasmids of *Neurospora*: integration into mitochondrial DNA and evidence for reverse transcription in mitochondria. *Cell*. Nov. 21, 1986;47(4):505-16. doi: 10.1016/0092-8674(86)90615-x.
- Akinsheye et al., Fetal hemoglobin in sickle cell anemia. *Blood*. Jul. 7, 2011;118(1):19-27. doi: 10.1182/blood-2011-03-325258. Epub Apr. 13, 2011.
- Akopian et al., Chimeric recombinases with designed DNA sequence recognition. *Proc Natl Acad Sci U S A*. Jul. 22, 2003;100(15):8688-91. Epub Jul. 1, 2003.
- Alarcón et al., HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events. *Cell*. Sep. 10, 2015;162(6):1299-308. doi: 10.1016/j.cell.2015.08.011. Epub Aug. 27, 2015.
- Alarcón et al., N6-methyladenosine marks primary microRNAs for processing. *Nature*. Mar. 26, 2015;519(7544):482-5. doi: 10.1038/nature14281. Epub Mar. 18, 2015.
- Alexander, HFE-associated hereditary hemochromatosis. *Genet Med*. May 2009;11(5):307-13. doi: 10.1097/GIM.0b013e31819d30f2.
- Alexandrov et al., Signatures of mutational processes in human cancer. *Nature*. Aug. 22, 2013;500(7463):415-21. doi: 10.1038/nature12477. Epub Aug. 14, 2013.
- Ali et al., Novel genetic abnormalities in Bernard-Soulier syndrome in India. *Ann Hematol*. Mar. 2014;93(3):381-4. doi: 10.1007/s00277-013-1895-x. Epub Sep. 1, 2013.
- Altschul et al., Basic local alignment search tool. *J Mol Biol*. Oct. 5, 1990;215(3):403-10. doi: 10.1016/S0022-2836(05)80360-2.
- Amato et al., Interpreting elevated fetal hemoglobin in pathology and health at the basic laboratory level: new and known γ-gene mutations associated with hereditary persistence of fetal hemoglobin. *Int J Lab Hematol*. Feb. 2014;36(1):13-9. doi: 10.1111/ijlh.12094. Epub Apr. 29, 2013.
- Ames et al., A eubacterial riboswitch class that senses the coenzyme tetrahydrofolate. *Chem Biol*. Jul. 30, 2010;17(7):681-5. doi: 10.1016/j.chembiol.2010.05.020.
- Amrann et al., Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene*. Sep. 30, 1988;69(2):301-15.
- Anders et al., Chapter One: In Vitro Enzymology of Cas9. in Methods in Enzymology, eds Doudna et al. 2014: 546:1-20.
- Anders et al., Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*. Sep. 25, 2014;513(7519):569-73. doi: 10.1038/nature13579. Epub Jul. 27, 2014.
- Anderson, Human gene therapy. *Science*. May 8, 1992;256(5058):808-13. doi: 10.1126/science.1589762.
- André et al., Axotomy-induced expression of calcium-activated chloride current in subpopulations of mouse dorsal root ganglion neurons. *J Neurophysiol*. Dec. 2003;90(6):3764-73. doi: 10.1152/jn.00449.2003. Epub Aug. 27, 2003.
- Anzalone et al., Reprogramming eukaryotic translation with ligand-responsive synthetic RNA switches. *Nat Methods*. May 2016;13(5):453-8. doi: 10.1038/nmeth.3807. Epub Mar. 21, 2016.
- Anzalone et al., Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. Dec. 2019;576(7785):149-157. doi: 10.1038/s41586-019-1711-4. Epub Oct. 21, 2019.
- Aplan, Causes of oncogenic chromosomal translocation. *Trends Genet*. Jan. 2006;22(1):46-55. doi: 10.1016/j.tig.2005.10.002. Epub Oct. 28, 2005.
- Arakawa et al., A method to convert mRNA into a gRNA library for CRISPR/Cas9 editing of any organism. *Sci Adv*. Aug. 24, 2016;2(8):e1600699. doi: 10.1126/sciadv.1600699.
- Araki et al., Comparative analysis of right element mutant lox sites on recombination efficiency in embryonic stem cells. *BMC Biotechnol*. Mar. 31, 2010;10:29. doi: 10.1186/1472-6750-10-29.
- Araki et al., Site-specific recombinase, R, encoded by yeast plasmid pSR1. *J Mol Biol*. May 5, 1992;225(1):25-37. doi: 10.1016/0022-2836(92)91023-i.
- Araki et al., Targeted integration of DNA using mutant lox sites in embryonic stem cells. *Nucleic Acids Res*. Feb. 15, 1997;25(4):868-72. doi: 10.1093/nar/25.4.868.
- Arambula et al., Surface display of a massively variable lipoprotein by a *Legionella* diversity-generating retroelement. *Proc Natl Acad Sci U S A*. May 14, 2013;110(20):8212-7. doi: 10.1073/pnas.1301366110. Epub Apr. 30, 2013.
- Arazoe et al., Targeted Nucleotide Editing Technologies for Microbial Metabolic Engineering. *Biotechnol J*. Sep. 2018;13(9):e1700596. doi: 10.1002/biot.201700596. Epub Jun. 19, 2018.
- Arbab et al., Cloning-free CRISPR. *Stem Cell Reports*. Nov. 10, 2015;5(5):908-917. doi: 10.1016/j.stemcr.2015.09.022. Epub Oct. 29, 2015.
- Arezi et al., Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic Acids Res*. Feb. 2009;37(2):473-81. doi: 10.1093/nar/gkn952. Epub Dec. 4, 2008.
- Arnold et al., Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. *EMBO J*. Mar. 1, 1999;18(5):1407-14.
- Asante et al., A naturally occurring variant of the human prion protein completely prevents prion disease. *Nature*. Jun. 25, 2015;522(7557):478-81. doi: 10.1038/nature14510. Epub Jun. 10, 2015.
- Asokan et al., The AAV vector toolkit: poised at the clinical crossroads. *Mol Ther*. Apr. 2012;20(4):699-708. doi: 10.1038/mt.2011.287. Epub Jan. 24, 2012.

(56)

References Cited

OTHER PUBLICATIONS

- Atkins et al., Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Res.* Sep. 6, 2016;44(15):7007-78. doi: 10.1093/nar/gkw530. Epub Jul. 19, 2016.
- Auer et al., Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.* Jan. 2014;24(1):142-53. doi: 10.1101/gr.161638.113. Epub Oct. 31, 2013.
- Auricchio et al., Exchange of surface proteins impacts on viral vector cellular specificity and transduction characteristics: the retina as a model. *Hum Mol Genet.* Dec. 15, 2001;10(26):3075-81. doi: 10.1093/hmg/10.26.3075.
- Autieri et al., IRT-1, a novel interferon-gamma-responsive transcript encoding a growth-suppressing basic leucine zipper protein. *J Biol Chem.* Jun. 12, 1998;273(24):14731-7. doi: 10.1074/jbc.273.24.14731.
- Avidan et al., The processivity and fidelity of DNA synthesis exhibited by the reverse transcriptase of bovine leukemia virus. *Eur J Biochem.* Feb. 2002;269(3):859-67. doi: 10.1046/j.0014-2956.2001.02719.x.
- Babacic et al., CRISPR-cas gene-editing as plausible treatment of neuromuscular and nucleotide-repeat-expansion diseases: A systematic review. *PLoS One.* Feb. 22, 2019;14(2):e0212198. doi: 10.1371/journal.pone.0212198.
- Bacman et al., Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat Med.* Sep. 2013;19(9):1111-3. doi: 10.1038/nm.3261. Epub Aug. 4, 2013.
- Badran et al., Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. *Nature.* May 5, 2016;533(7601):58-63. doi: 10.1038/nature17938. Epub Apr. 27, 2016.
- Badran et al., Development of potent *in vivo* mutagenesis plasmids with broad mutational spectra. *Nat Commun.* Oct. 7, 2015;6:8425. doi: 10.1038/ncomms9425.
- Badran et al., *In vivo* continuous directed evolution. *Curr Opin Chem Biol.* Feb. 2015;24:1-10. doi: 10.1016/j.cbpa.2014.09.040. Epub Nov. 7, 2014.
- Bae et al., Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics.* May 15, 2014;30(10):1473-5. doi: 10.1093/bioinformatics/btu048. Epub Jan. 24, 2014.
- Bae et al., Microhomology-based choice of Cas9 nuclease target sites. *Nat Methods.* Jul. 2014;11(7):705-6. doi: 10.1038/nmeth.3015.
- Bagal et al., Recent progress in sodium channel modulators for pain. *Bioorg Med Chem Lett.* Aug. 15, 2014;24(16):3690-9. doi: 10.1016/j.bmcl.2014.06.038. Epub Jun. 21, 2014.
- Bagyinszky et al., Characterization of mutations in PRNP (prion) gene and their possible roles in neurodegenerative diseases. *Neuropsychiatr Dis Treat.* Aug. 14, 2018;14:2067-2085. doi: 10.2147/NDT.S165445.
- Balakrishnan et al., Flap endonuclease 1. *Annu Rev Biochem.* 2013;82:119-38. doi: 10.1146/annurev-biochem-072511-122603. Epub Feb. 28, 2013.
- Baldari et al., A novel leader peptide which allows efficient secretion of a fragment of human interleukin 1 beta in *Saccharomyces cerevisiae*. *EMBO J.* Jan. 1987;6(1):229-34.
- Banerjee et al., Cadmium inhibits mismatch repair by blocking the ATPase activity of the MSH2-MSH6 complex [published correction appears in Nucleic Acids Res. 2005;33(5):1738]. *Nucleic Acids Res.* 2005;33(4):1410-1419. Published Mar. 3, 2005. doi:10.1093/nar/gki291.
- Banerji et al., A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell.* Jul. 1983;33(3):729-40. doi: 10.1016/0092-8674(83)90015-6.
- Bannert et al., Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci U S A.* Oct. 5, 2004;101 Suppl 2(Suppl 2):14572-9. doi: 10.1073/pnas.0404838101. Epub Aug. 13, 2004.
- Banno et al., Deaminase-mediated multiplex genome editing in *Escherichia coli*. *Nat Microbiol.* Apr. 2018;3(4):423-429. doi: 10.1038/s41564-017-0102-6. Epub Feb. 5, 2018.
- Barmania et al., C-C chemokine receptor type five (CCR5): An emerging target for the control of HIV infection. *Appl Transl Genom.* May 26, 2013;2:3-16. doi: 10.1016/j.atg.2013.05.004.
- Barnes et al., Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet.* 2004;35:445-76.
- Barnes et al., The fidelity of Taq polymerase catalyzing PCR is improved by an -terminal deletion. *Gene.* Mar. 1, 1992;112(1):29-35. doi: 10.1016/0378-1119(92)90299-5.
- Barrangou et al., CRISPR provides acquired resistance against viruses in prokaryotes. *Science.* Mar. 23, 2007;315(5819):1709-12.
- Barrangou, RNA-mediated programmable DNA cleavage. *Nat Biotechnol.* Sep. 2012;30(9):836-8. doi: 10.1038/nbt.2357.
- Bartlett et al., Efficient expression of protein coding genes from the murine U1 small nuclear RNA promoters. *Proc Natl Acad Sci U S A.* Aug. 20, 1996;93(17):8852-7. doi: 10.1073/pnas.93.17.8852.
- Bartosovic et al., N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3'-end processing. *Nucleic Acids Res.* Nov. 2, 2017;45(19):11356-11370. doi: 10.1093/nar/gkx778.
- Basha et al., Influence of cationic lipid composition on gene silencing properties of lipid nanoparticle formulations of siRNA in antigen-presenting cells. *Mol Ther.* Dec. 2011;19(12):2186-200. doi: 10.1038/mt.2011.190. Epub Oct. 4, 2011.
- Basturea et al., Substrate specificity and properties of the *Escherichia coli* 16S rRNA methyltransferase, RsmE. *RNA.* Nov. 2007;13(11):1969-76. doi: 10.1261/rna.700507. Epub Sep. 13, 2007.
- Batey et al., Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature.* Nov. 18, 2004;432(7015):411-5.
- Beale et al., Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra *in vivo*. *J Mol Biol.* Mar. 26, 2004;337(3):585-96.
- Beaudry et al., Directed evolution of an RNA enzyme. *Science.* Jul. 31, 1992;257(5070):635-41. doi: 10.1126/science.1496376.
- Bebenek et al., Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. *J Biol Chem.* May 15, 1993;268(14):10324-34.
- Bedell et al., *In vivo* genome editing using a high-efficiency TALEN system. *Nature.* Nov. 1, 2012;491(7422):114-8. doi: 10.1038/nature11537. Epub Sep. 23, 2012.
- Begley, Scientists unveil the 'most clever CRISPR gadget' so far. *STAT.* Apr. 20, 2016. <https://www.statnews.com/2016/04/20/clever-crispr-advance-unveiled/>.
- Behr, Gene transfer with synthetic cationic amphiphiles: prospects for gene therapy. *Bioconjug Chem.* Sep.-Oct. 1994;5(5):382-9. doi: 10.1021/bc00029a002.
- Bell et al., Ribozyme-catalyzed excision of targeted sequences from within RNAs. *Biochemistry.* Dec. 24, 2002;41(51):15327-33. doi: 10.1021/bi0267386.
- Belshaw et al., Controlling programmed cell death with a cyclophilin-cyclosporin-based chemical inducer of dimerization. *Chem Biol.* Sep. 1996;3(9):731-8. doi: 10.1016/s1074-5521(96)90249-5.
- Belshaw et al., Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *Proc Natl Acad Sci U S A.* May 14, 1996;93(10):4604-7. doi: 10.1073/pnas.93.10.4604.
- Benarroch, HCN channels: function and clinical implications. *Neurology.* Jan. 15, 2013;80(3):304-10. doi: 10.1212/WNL.0b013e31827dec42.
- Bennett et al., Painful and painless channelopathies. *Lancet Neurol.* Jun. 2014;13(6):587-99. doi: 10.1016/S1474-4422(14)70024-9. Epub May 6, 2014.
- Bentin, T, A ribozyme transcribed by a ribozyme. *Artif DNA PNA XNA.* Apr. 2011;2(2):40-42. doi: 10.4161/adna.2.2.16852.
- Berger et al., Reverse transcriptase and its associated ribonuclease H: interplay of two enzyme activities controls the yield of single-stranded complementary deoxyribonucleic acid. *Biochemistry.* May 10, 1983;22(10):2365-72. doi: 10.1021/bi00279a010.

(56)

References Cited**OTHER PUBLICATIONS**

- Berges et al., Transduction of brain by herpes simplex virus vectors. *Mol Ther.* Jan. 2007;15(1):20-9. doi: 10.1038/sj.mt.6300018.
- Berkhout et al., Identification of an active reverse transcriptase enzyme encoded by a human endogenous HERV-K retrovirus. *J Virol.* Mar. 1999;73(3):2365-75. doi: 10.1128/JVI.73.3.2365-2375.1999.
- Bernhart et al., Local RNA base pairing probabilities in large sequences. *Bioinformatics.* Mar. 1, 2006;22(5):614-5. doi: 10.1093/bioinformatics/btk014. Epub Dec. 20, 2005.
- Bernstein et al., Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature.* Jan. 18, 2001;409(6818):363-6. doi: 10.1038/35053110.
- Bershtein et al., Advances in laboratory evolution of enzymes. *Curr Opin Chem Biol.* Apr. 2008;12(2):151-8. doi: 10.1016/j.cbpa.2008.01.027. Epub Mar. 7, 2008. Review.
- Bertolotti et al., Toward genosafe endonuclease-boosted gene targeting using breakthrough CRISP/Cas9 for next generation stem cell gene therapy culminating in efficient ex VIVO in VIVO gene repair/genomic editing. *Molecular Therapy.* May 2015;23(Suppl1):S139. Abstract 350. 18th Ann Meeting of the American Society of Gene and Cell Therapy. ASGCT 2015. New Orleans, LA. May 13, 2015-May 16, 2015.
- Bertrand et al., Localization of ASHI mRNA particles in living yeast. *Mol Cell.* Oct. 1998;2(4):437-45. doi: 10.1016/s1097-2765(00)80143-4.
- Beumer et al., Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics.* Apr. 2006;172(4):2391-403. Epub Feb. 1, 2006.
- Bhagwat, DNA-cytosine deaminases: from antibody maturation to antiviral defense. *DNA Repair (Amst).* Jan. 5, 2004;3(1):85-9.
- Bi et al., Pseudo attP sites in favor of transgene integration and expression in cultured porcine cells identified by Streptomyces phage phiC31 integrase. *BMC Mol Biol.* Sep. 8, 2013;14:20. doi: 10.1186/1471-2199-14-20.
- Bibb et al., Integration and excision by the large serine recombinase phiRv1 integrase. *Mol Microbiol.* Mar. 2005;55(6):1896-910. doi: 10.1111/j.1365-2958.2005.04517.x.
- Biehs et al., DNA Double-Strand Break Resection Occurs during Non-homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination. *Mol Cell.* Feb. 16, 2017;65(4):671-684.e5. doi: 10.1016/j.molcel.2016.12.016. Epub Jan. 26, 2017.
- Billon et al., CRISPR-Mediated Base Editing Enables Efficient Disruption of Eukaryotic Genes through Induction of STOP Codons. *Mol Cell.* Sep. 21, 2017;67(6):1068-1079.e4. doi: 10.1016/j.molcel.2017.08.008. Epub Sep. 7, 2017.
- Birling et al., Site-specific recombinases for manipulation of the mouse genome. *Methods Mol Biol.* 2009;561:245-63. doi: 10.1007/978-1-60327-019-9_16.
- Biswas et al., A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature.* Jun. 23, 2005;435(7045):1059-66. doi: 10.1038/nature03657.
- Bitinaite et al., FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci U S A.* Sep. 1, 1998;95(18):10570-5.
- Blaeze et al., Vectors in cancer therapy: how will they deliver? *Cancer Gene Ther.* Dec. 1995;2(4):291-7.
- Blain et al., Nuclease activities of Moloney murine leukemia virus reverse transcriptase. Mutants with altered substrate specificities. *J Biol Chem.* Nov. 5, 1993;268(31):23585-92.
- Blaisonneau et al., A circular plasmid from the yeast *Torulaspora delbrueckii*. *Plasmid.* 1997;38(3):202-9. doi: 10.1006/plas.1997.1315.
- Blau et al., A proliferation switch for genetically modified cells. *PNAS* Apr. 1, 1997 94 (7) 3076-3081; https://doi.org/10.1073/pnas.94.7.3076.
- Bloom et al., Evolving strategies for enzyme engineering. *Curr Opin Struct Biol.* Aug. 2005;15(4):447-52.
- Boch, TALEs of genome targeting. *Nat Biotechnol.* Feb. 2011;29(2):135-6. Doi: 10.1038/nbt.1767.
- Bodi et al., Yeast m6A Methylated mRNAs Are Enriched on Translating Ribosomes during Meiosis, and under Rapamycin Treatment. *PLoS One.* Jul. 17, 2015;10(7):e0132090. doi: 10.1371/journal.pone.0132090.
- Boeckle et al., Melittin analogs with high lytic activity at endosomal pH enhance transfection with purified targeted PEI polyplexes. *J Control Release.* May 15, 2006;112(2):240-8. Epub Mar. 20, 2006.
- Boersma et al., Selection strategies for improved biocatalysts. *FEBS J.* May 2007;274(9):2181-95.
- Bogdanov et al., Engineering altered protein-DNA recognition specificity. *Nucleic Acids Res.* Jun. 1, 2018;46(10):4845-4871. doi: 10.1093/nar/gky289.
- Bogdanov et al., TAL effectors: customizable proteins for DNA targeting. *Science.* Sep. 30, 2011;333(6051):1843-6. doi: 10.1126/science.1204094.
- Bohlke et al., Sense codon emancipation for proteome-wide incorporation of noncanonical amino acids: rare isoleucine codon AUA as a target for genetic code expansion. *FEMS Microbiol Lett.* Feb. 2014;351(2):133-44. doi: 10.1111/1574-6968.12371. Epub Jan. 27, 2014.
- Bolotin et al., Clustered regularly interspaced short palindromic repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology.* Aug. 2005;151(Pt 8):2551-61.
- Bolusani et al., Evolution of variants of yeast site-specific recombinase Flp that utilize native genomic sequences as recombination target sites. *Nucleic Acids Res.* 2006;34(18):5259-69. Epub Sep. 26, 2006.
- Bondeson et al., Inversion of the IDS gene resulting from recombination with IDS-related sequences is a common cause of the Hunter syndrome. *Hum Mol Genet.* Apr. 1995;4(4):615-21. doi: 10.1093/hmg/4.4.615.
- Borchardt et al., Controlling mRNA stability and translation with the CRISPR endoribonuclease Csy4. *RNA.* Nov. 2015;21(11):1921-30. doi: 10.1261/rna.051227.115. Epub Sep. 9, 2015.
- Borman, Improved route to single-base genome editing. *Chemical & Engineering News.* Apr. 25, 2016;94(17)p5. http://cen.acs.org/articles/94/i17/Improved-route-single-base-genome.html.
- Bourinet et al., Silencing of the Cav3.2 T-type calcium channel gene in sensory neurons demonstrates its major role in nociception. *EMBO J.* Jan. 26, 2005;24(2):315-24. doi: 10.1038/sj.emboj.7600515. Epub Dec. 16, 2004.
- Boutabout et al., DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1. *Nucleic Acids Res.* Jun. 1, 2001;29(11):2217-22. doi: 10.1093/nar/29.11.2217.
- Box et al., A multi-domain protein system based on the HC fragment of tetanus toxin for targeting DNA to neuronal cells. *J Drug Target.* Jul. 2003;11(6):333-43. doi: 10.1080/1061186310001634667.
- Branden and Tooze, *Introduction to Protein Structure.* 1999; 2nd edition. Garland Science Publisher: 3-12.
- Braun et al., Immunogenic duplex nucleic acids are nuclease resistant. *J Immunol.* Sep. 15, 1988;141(6):2084-9.
- Brierley et al., Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nat Rev Microbiol.* Aug. 2007;5(8):598-610. doi: 10.1038/nrmicro1704.
- Briner et al., Guide RNA functional modules direct Cas9 activity and orthogonality. *Mol Cell.* Oct. 23, 2014;56(2):333-339. doi: 10.1016/j.molcel.2014.09.019.
- Britt et al., Re-engineering plant gene targeting. *Trends Plant Sci.* Feb. 2003;8(2):90-5.
- Brouns et al., Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science.* Aug. 15, 2008;321(5891):960-4. doi: 10.1126/science.1159689.
- Brown et al., A mammalian protein targeted by G1-arresting rapamycin-receptor complex. *Nature.* Jun. 30, 1994;369(6483):756-8. doi: 10.1038/369756a0.
- Brown et al., Characterization of the genetic elements required for site-specific integration of plasmid pSE211 in *Saccharopolyspora erythraea*. *J Bacteriol.* Apr. 1990;172(4):1877-88. doi: 10.1128/jb.172.4.1877-1888.1990.
- Brown et al., Serine recombinases as tools for genome engineering. *Methods.* Apr. 2011;53(4):372-9. doi: 10.1016/j.ymeth.2010.12.031. Epub Dec. 30, 2010.

(56)

References Cited**OTHER PUBLICATIONS**

- Brown et al., Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat Struct Mol Biol.* Jul. 2014;21(7):633-40. doi: 10.1038/nsmb.2844. Epub Jun. 22, 2014.
- Brusse et al., Spinocerebellar ataxia associated with a mutation in the fibroblast growth factor 14 gene (SCA27): A new phenotype. *Mov Disord.* Mar. 2006;21(3):396-401.
- Brzezicha et al., Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic Acids Res.* 2006;34(20):6034-43. doi: 10.1093/nar/gk1765. Epub Oct. 27, 2006.
- Buchholz et al., Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nat Biotechnol.* Nov. 2001;19(11):1047-52.
- Buchsacher et al., Human immunodeficiency virus vectors for inducible expression of foreign genes. *J Virol.* May 1992;66(5):2731-9. doi: 10.1128/JVI.66.5.2731-2739.1992.
- Buchwald et al., Long-term, continuous intravenous heparin administration by an implantable infusion pump in ambulatory patients with recurrent venous thrombosis. *Surgery.* Oct. 1980;88(4):507-16.
- Buckley et al., Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1 α interaction. *J Am Chem Soc.* Mar. 14, 2012;134(10):4465-8. doi: 10.1021/ja209924v. Epub Feb. 27, 2012.
- Budisa et al., Residue-specific bioincorporation of non-natural, biologically active amino acids into proteins as possible drug carriers: structure and stability of the per-thiaproline mutant of annexin V. *Proc Natl Acad Sci U S A.* Jan. 20, 1998;95(2):455-9.
- Budker et al., Protein/amphiphatic polyamine complexes enable highly efficient transfection with minimal toxicity. *Biotechniques.* Jul. 1997;23(1):139, 142-7. doi: 10.2144/97231rr02.
- Budworth et al., A brief history of triplet repeat diseases. *Methods Mol Biol.* 2013;1010:3-17. doi: 10.1007/978-1-62703-411-1_1.
- Bulow et al., Multienzyme systems obtained by gene fusion. *Trends Biotechnol.* Jul. 1991;9(7):226-31.
- Burke et al., Activating mutations of Tn3 resolvase marking interfaces important in recombination catalysis and its regulation. *Mol Microbiol.* Feb. 2004;51(4):937-48.
- Burke et al., RNA Aptamers to the Adenosine Moiety of S-adenosyl Methionine: Structural Inferences From Variations on a Theme and the Reproducibility of SELEX. *Nucleic Acids Res.* May 15, 1997;25(10):2020-4. doi: 10.1093/nar/25.10.2020.
- Burstein et al., New CRISPR-Cas systems from uncultivated microbes. *Nature* Feb. 2017;542(7640):237-240.
- Burton et al., Gene delivery using herpes simplex virus vectors. *DNA Cell Biol.* Dec. 2002;21(12):915-36. doi: 10.1089/104454902762053864.
- Buskirk et al., Directed evolution of ligand dependence: small-molecule-activated protein splicing. *Proc Natl Acad Sci U S A.* Jul. 20, 2004;101(29):10505-10. Epub Jul. 9, 2004.
- Buskirk et al., In vivo evolution of an RNA-based transcriptional activator. *Chem Biol.* Jun. 2003;10(6):533-40. doi: 10.1016/s1074-5521(03)00109-1.
- Butt et al., Efficient CRISPR/Cas9-Mediated Genome Editing Using a Chimeric Single-Guide RNA Molecule. *Front Plant Sci.* Aug. 24, 2017;8:1441(1-8). doi: 10.3389/fpls.2017.01441.
- Byrne et al., Multiplex gene regulation: a two-tiered approach to transgene regulation in transgenic mice. *Proc Natl Acad Sci U S A.* Jul. 1989;86(14):5473-7. doi: 10.1073/pnas.86.14.5473.
- Böck et al., Selenocysteine: the 21st amino acid. *Mol Microbiol.* Mar. 1991;5(3):515-20.
- Cade et al., Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res.* Sep. 2012;40(16):8001-10. doi: 10.1093/nar/gks518. Epub Jun. 7, 2012.
- Cadwell et al., Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* Aug. 1992;2(1):28-33. doi: 10.1101/gr.2.1.28.
- Cai et al., Reconstruction of ancestral protein sequences and its applications. *BMC Evol Biol.* Sep. 17, 2004;4:33. doi: 10.1186/1471-2148-4-33.
- Calame et al., Transcriptional controlling elements in the immunoglobulin and T cell receptor loci. *Adv Immunol.* 1988;43:235-75. doi: 10.1016/s0065-2776(08)60367-3.
- Caldecott et al., Single-strand break repair and genetic disease. *Nat Rev Genet.* Aug. 2008;9(8):619-31. doi: 10.1038/nrg2380.
- Camarero et al., Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity. *J. Am. Chem. Soc.* May 29, 1999; 121(23):5597-5598. https://doi.org/10.1021/ja990929n.
- Cameron, Recent advances in transgenic technology. *Mol Biotechnol.* Jun. 1997;7(3):253-65.
- Camper et al., Postnatal repression of the alpha-fetoprotein gene is enhancer independent. *Genes Dev.* Apr. 1989;3(4):537-46. doi: 10.1101/gad.3.4.537.
- Camps et al., Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. *Proc Natl Acad Sci U S A.* Aug. 19, 2003;100(17):9727-32. Epub Aug. 8, 2003.
- Canchaya et al., Genome analysis of an inducible prophage and prophage remnants integrated in the *Streptococcus pyogenes* strain SF370. *Virology.* Oct. 25, 2002;302(2):245-58. doi: 10.1006/viro.2002.1570.
- Canver et al., Customizing the genome as therapy for the β -hemoglobinopathies. *Blood.* May 26, 2016;127(21):2536-45. doi: 10.1182/blood-2016-01-678128. Epub Apr. 6, 2016.
- Cargill et al., Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet.* Jul. 1999;22(3):231-8.
- Carlier et al., Burkholderia cenocepacia H111 Rhy-family protein. Apr. 16, 2015. Retrieved from the Internet via https://www.ebi.ac.uk/ena/browser/api/emb/CDN65395.1?lineLimit=1000. Last retrieved Apr. 26, 2021.
- Carlson et al., Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat Chem Biol.* Mar. 2014;10(3):216-22. doi: 10.1038/nchembio.1453. Epub Feb. 2, 2014. With Supplementary Results.
- Caron et al., Intracellular delivery of a Tat-eGFP fusion protein into muscle cells. *Mol Ther.* Mar. 2001;3(3):310-8.
- Carr et al., Genome engineering. *Nat Biotechnol.* Dec. 2009;27(12):1151-62. doi: 10.1038/nbt.1590.
- Carroll et al., Gene targeting in *Drosophila* and *Caenorhabditis elegans* with zinc-finger nucleases. *Methods Mol Biol.* 2008;435:63-77. doi: 10.1007/978-1-59745-232-8_5.
- Carroll et al., Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther.* Nov. 2008;15(22):1463-8. doi: 10.1038/gt.2008.145. Epub Sep. 11, 2008.
- Carroll, A CRISPR approach to gene targeting. *Mol Ther.* Sep. 2012;20(9):1658-60. doi: 10.1038/mt.2012.171.
- Carroll, Genome engineering with zinc-finger nucleases. *Genetics.* Aug. 2011;188(4):773-82. doi: 10.1534/genetics.111.131433. Review.
- Carvalho et al., Evolution in health and medicine Sackler colloquium: Genomic disorders: a window into human gene and genome evolution. *Proc Natl Acad Sci U S A.* Jan. 26, 2010;107 Suppl 1(Suppl 1):1765-71. doi: 10.1073/pnas.0906222107. Epub Jan. 13, 2010.
- Caspi et al., Distribution of split DnaE inteins in cyanobacteria. *Mol Microbiol.* Dec. 2003;50(5):1569-77. doi: 10.1046/j.1365-2958.2003.03825.x.
- Cattaneo et al., SEL1L affects human pancreatic cancer cell cycle and invasiveness through modulation of PTEN and genes related to cell-matrix interactions. *Neoplasia.* 2005;7(11):1030-1038.
- Ceccaldi et al., Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol.* Jan. 2016;26(1):52-64. doi: 10.1016/j.tcb.2015.07.009. Epub Oct. 1, 2015.
- Cermak et al., Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* Jul. 2011;39(12):e82. doi: 10.1093/nar/gkr218. Epub Apr. 14, 2011.
- Chadalavada et al., Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions. *RNA.* Dec. 2007;13(12):2189-201. doi: 10.1261/rna.778107. Epub Oct. 23, 2007.
- Chadwick et al., In Vivo Base Editing of PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9) as a Therapeutic Alternative to

(56)

References Cited**OTHER PUBLICATIONS**

- Genome Editing. *Arterioscler Thromb Vasc Biol.* Sep. 2017;37(9):1741-1747. doi: 10.1161/ATVBAHA.117.309881. Epub Jul. 27, 2017.
- Chaikind et al., A programmable Cas9-serine recombinase fusion protein that operates on DNA sequences in mammalian cells. *Nucleic Acids Res.* Nov. 16, 2016;44(20):9758-9770. Epub Aug. 11, 2016.
- Chalberg et al., Integration specificity of phage phiC31 integrase in the human genome. *J Mol Biol.* Mar. 17, 2006;357(1):28-48. doi: 10.1016/j.jmb.2005.11.098. Epub Dec. 22, 2005.
- Chalberg et al., phiC31 integrase confers genomic integration and long-term transgene expression in rat retina. *Invest Ophthalmol Vis Sci.* Jun. 2005;46(6):2140-6. doi: 10.1167/iovs.04-1252.
- Chan et al., Molecular recording of mammalian embryogenesis. *Nature.* Jun. 2019;570(7759):77-82. doi: 10.1038/s41586-019-1184-5. Epub May 13, 2019.
- Chan et al., Novel selection methods for DNA-encoded chemical libraries. *Curr Opin Chem Biol.* 2015;26:55-61. doi:10.1016/j.cbpa.2015.02.010.
- Chan et al., The choice of nucleotide inserted opposite abasic sites formed within chromosomal DNA reveals the polymerase activities participating in translesion DNA synthesis. *DNA Repair (Amst).* Nov. 2013;12(11):878-89. doi: 10.1016/j.dnarep.2013.07.008. Epub Aug. 26, 2013.
- Chapman et al., Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell.* Aug. 24, 2012;47(4):497-510. doi: 10.1016/j.molcel.2012.07.029.
- Chari et al., Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat Methods.* Sep. 2015;12(9):823-6. doi: 10.1038/nmeth.3473. Epub Jul. 13, 2015.
- Charpentier et al., Biotechnology: Rewriting a genome. *Nature.* Mar. 7, 2013;495(7439):50-1. doi: 10.1038/495050a.
- Chaturvedi et al., Stabilization of triple-stranded oligonucleotide complexes: use of probes containing alternating phosphodiester and stereo-uniform cationic phosphoramidate linkages. *Nucleic Acids Res.* Jun. 15, 1996;24(12):2318-23.
- Chavez et al., Highly efficient Cas9-mediated transcriptional programming. *Nat Methods.* Apr. 2015;12(4):326-8. doi: 10.1038/nmeth.3312. Epub Mar. 2, 2015.
- Chavez et al., Precise Cas9 targeting enables genomic mutation prevention. *bioRxiv.* Jun. 14, 2016; <http://dx.doi.org/10.1101/058974>. 6 pages. bioRxiv preprint first posted online Jun. 14, 2016.
- Chavez et al., Therapeutic applications of the PhiC31 integrase system. *Curr Gene Ther.* Oct. 2011;11(5):375-81. Review.
- Chawla et al., An atlas of RNA base pairs involving modified nucleobases with optimal geometries and accurate energies. *Nucleic Acids Res.* Aug. 18, 2015;43(14):6714-29. doi: 10.1093/nar/gkv606. Epub Jun. 27, 2015.
- Chelico et al., Biochemical basis of immunological and retroviral responses to DNA-targeted cytosine deamination by activation-induced cytidine deaminase and APOBEC3G. *J Biol Chem.* Oct. 9, 2009;284(41):27761-5. doi: 10.1074/jbc.R109.052449. Epub Aug. 13, 2009.
- Chelico et al., Stochastic properties of processive cytidine DNA deaminases AID and APOBEC3G. *Philos Trans R Soc Lond B Biol Sci.* Mar. 12, 2009;(1517):583-93. doi: 10.1098/rstb.2008.0195.
- Chen et al., Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature.* Oct. 19, 2017;550(7676):407-410. doi: 10.1038/nature24268. Epub Sep. 20, 2017.
- Chen et al., A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc Natl Acad Sci U S A.* Jul. 12, 2011;108(28):11399-404. doi: 10.1073/pnas.1101046108. Epub Jun. 22, 2011.
- Chen et al., Alterations in PMS2, MSH2 and MLH1 expression in human prostate cancer. *Int J Oncol.* May 2003;22(5):1033-43.
- Chen et al., Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev.* Oct. 2013;65(10):1357-69. doi: 10.1016/j.addr.2012.09.039. Epub Sep. 29, 2012.
- Chen et al., Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell.* Mar. 12, 2015;160(6):1246-60. doi: 10.1016/j.cell.2015.02.038. Epub Mar. 5, 2015.
- Chen et al., Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. *J Biol Chem.* Jul. 8, 2016;291(28):14457-67. doi: 10.1074/jbc.M116.733154. Epub May 5, 2016.
- Chen et al., m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell.* Mar. 5, 2015;16(3):289-301. doi: 10.1016/j.stem.2015.01.016. Epub Feb. 12, 2015.
- Chen et al., Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature.* Mar. 6, 2008;452(7183):116-9. doi: 10.1038/nature06638. Epub Feb. 20, 2008.
- Chen et al., Targeting genomic rearrangements in tumor cells through Cas9-mediated insertion of a suicide gene. *Nat Biotechnol.* Jun. 2017;35(6):543-550. doi: 10.1038/nbt.3843. Epub May 1, 2017.
- Cheng et al., Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* Oct. 2013;23(10):1163-71. doi: 10.1038/cr.2013.122. Epub Aug. 27, 2013.
- Chesnoy et al., Structure and function of lipid-DNA complexes for gene delivery. *Annu Rev Biophys Biomol Struct.* 2000;29:27-47.
- Chester et al., The apolipoprotein B mRNA editing complex performs a multifunctional cycle and suppresses nonsense-mediated decay. *EMBO J.* Aug. 1, 2003;22(15):3971-82. doi: 10.1093/emboj/cdg369.
- Chew et al., A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat Methods.* Oct. 2016;13(10):868-74. doi: 10.1038/nmeth.3993. Epub Sep. 5, 2016.
- Chew et al., A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat Methods.* Oct. 2016;13(10):868-74. doi: 10.1038/nmeth.3993. Epub Sep. 5, 2016. Supplementary Information.
- Chichili et al., Linkers in the structural biology of protein-protein interactions. *Protein Science.* 2013;22:153-67.
- Chin, Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Biochem.* 2014;83:379-408. doi: 10.1146/annurev-biochem-060713-035737. Epub Feb. 10, 2014.
- Chipev et al., A leucine---proline mutation in the H1 subdomain of keratin 1 causes epidermolytic hyperkeratosis. *Cell.* Sep. 4, 1992;70(5):821-8.
- Cho et al., Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* Jan. 2014;24(1):132-41. doi: 10.1101/gr.162339.113. Epub Nov. 19, 2013.
- Cho et al., Site-specific recombination of bacteriophage P22 does not require integration host factor. *J Bacteriol.* Jul. 1999;181(14):4245-9. doi: 10.1128/JB.181.14.4245-4249.1999.
- Cho et al., Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol.* Mar. 2013;31(3):230-2. doi: 10.1038/nbt.2507. Epub Jan. 29, 2013.
- Cho et al., The calcium-activated chloride channel anoctamin 1 acts as a heat sensor in nociceptive neurons. *Nat Neurosci.* May 27, 2012;15(7):1015-21. doi: 10.1038/nn.3111.
- Choe et al., Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Mol Cell.* Feb. 2, 2017;65(3):380-392. doi: 10.1016/j.molcel.2016.12.020.
- Choi et al., (6)-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat Struct Mol Biol.* Feb. 2016;23(2):110-5. doi: 10.1038/nsmb.3148. Epub Jan. 11, 2016.
- Choi et al., Protein trans-splicing and characterization of a split family B-type DNA polymerase from the hyperthermophilic archaeal parasite Nanoarchaeum equitans. *J Mol Biol.* Mar. 10, 2006;356(5):1093-106. doi: 10.1016/j.jmb.2005.12.036. Epub Dec. 27, 2005.
- Choi et al., Translesion synthesis across abasic lesions by human B-family and Y-family DNA polymerases ?, ?, ?, ?, ?, and REV1. *J Mol Biol.* Nov. 19, 2010;404(1):34-44. doi: 10.1016/j.jmb.2010.09.015. Epub Oct. 1, 2010.
- Chong et al., Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein. *J Biol Chem.* Apr. 24, 1998;273(17):10567-77. doi: 10.1074/jbc.273.17.10567.

(56)

References Cited**OTHER PUBLICATIONS**

- Chong et al., Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res.* Nov. 15, 1998;26(22):5109-15. doi: 10.1093/nar/26.22.5109.
- Chong et al., Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an *in vitro* splicing system. *J Biol Chem.* Sep. 6, 1996;271(36):22159-68. doi: 10.1074/jbc.271.36.22159.
- Chong et al., Protein splicing of the *Saccharomyces cerevisiae* VMA intein without the endonuclease motifs. *J Biol Chem.* Jun. 20, 1997;272(25):15587-90. doi: 10.1074/jbc.272.25.15587.
- Chong et al., Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene.* Jun. 19, 1997;192(2):271-81. doi: 10.1016/s0378-1119(97)00105-4.
- Choudhury et al., CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. *Oncotarget.* Jul. 19, 2016;7(29):46545-46556. doi: 10.18633/oncotarget.10234.
- Choudhury et al., CRISPR/Cas9 recombineering-mediated deep mutational scanning of essential genes in *Escherichia coli*. *Mol Syst Biol.* Mar. 2020;16(3):e9265. doi: 10.1525/msb.2019265.
- Choudhury et al., Engineering RNA endonucleases with customized sequence specificities. *Nat Commun.* 2012;3:1147. doi: 10.1038/ncomms2154.
- Choulika et al., Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol.* Apr. 1995;15(4):1968-73. doi: 10.1128/MCB.15.4.1968.
- Christian et al., Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS One.* 2012;7(9):e45383. doi: 10.1371/journal.pone.0045383. Epub Sep. 24, 2012.
- Christian et al., Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics.* Oct. 2010;186(2):757-61. doi: 10.1534/genetics.110.120717. Epub Jul. 26, 2010.
- Christiansen et al., Characterization of the lactococcal temperate phage TP901-1 and its site-specific integration. *J Bacteriol.* Feb. 1994;176(4):1069-76. doi: 10.1128/jb.176.4.1069-1076.1994.
- Chu et al., Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotech.* Feb. 13, 2015;33:543-8. doi: 10.1038/nbt.3198. Epub Mar. 24, 2015.
- Chuai et al., DeepCRISPR: optimized CRISPR guide RNA design by deep learning. *Genome Biol.* Jun. 26, 2018;19(1):80. doi: 10.1186/s13059-018-1459-4.
- Chuai et al., In Silico Meets In Vivo: Towards Computational CRISPR-Based sgRNA Design. *Trends Biotechnol.* Jan. 2017;35(1):12-21. doi: 10.1016/j.tibtech.2016.06.008. Epub Jul. 11, 2016.
- Chuang et al., Novel Heterotypic Rox Sites for Combinatorial Dre Recombination Strategies. *G3 (Bethesda).* Dec. 29, 2015;6(3):559-71. doi: 10.1534/g3.115.025841.
- Chujo et al., Trmt61B is a methyltransferase responsible for 1-methyladenosine at position 58 of human mitochondrial tRNAs. *RNA.* Dec. 2012;18(12):2269-76. doi: 10.1261/rna.035600.112. Epub Oct. 24, 2012.
- Chung-II et al., Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer-small molecule interaction. *RNA.* May 2006;12(5):710-6. Epub Apr. 10, 2006.
- Chylinski et al., The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* May 2013;10(5):726-37. doi: 10.4161/rna.24321. Epub Apr. 5, 2013.
- Clackson et al., Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc Natl Acad Sci U S A.* Sep. 1, 1998;95(18):10437-42. doi: 10.1073/pnas.95.18.10437.
- Clement et al., CRISPRESSo2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol.* Mar. 2019;37(3):224-226. doi: 10.1038/s41587-019-0032-3.
- Cobb et al., Directed evolution as a powerful synthetic biology tool. *Methods.* Mar. 15, 2013;60(1):81-90. doi: 10.1016/j.ymeth.2012.03.009. Epub Mar. 23, 2012.
- Coffey et al., The Economic Impact of BSE on the U.S. Beef Industry: Product Value Losses, Regulatory Costs, and Consumer Reactions. Kansas State University Agricultural Experiment Station and Cooperative Extension Service. MF-2678. May 2005. 68 pages. Accessed via <https://bookstore.ksre.ksu.edu/pubs/MF2678.pdf>.
- Cokol et al., Finding nuclear localization signals. *EMBO Rep.* Nov. 2000;1(5):411-5. doi: 10.1093/embo-reports/kvd092.
- Cole et al., Reconstructing evolutionary adaptive paths for protein engineering. *Methods Mol Biol.* 2013;978:115-25. doi: 10.1007/978-1-62703-293-3_8.
- Cole-Strauss et al., Correction of the mutation responsible for sickle cell anemia by an RNA-DNA oligonucleotide. *Science.* Sep. 6, 1996;273(5280):1386-9.
- Collinge, Prion diseases of humans and animals: their causes and molecular basis. *Annu Rev Neurosci.* 2001;24:519-50. doi: 10.1146/annurev.neuro.24.1.519.
- Cong et al., Multiplex genome engineering using CRISPR/Cas systems. *Science.* Feb. 15, 2013;339(6121):819-23. doi: 10.1126/science.1231143. Epub Jan. 3, 2013.
- Conrad et al., A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts. *EMBO J.* May 18, 2005;24(10):1831-41. doi: 10.1038/sj.emboj.7600662. Epub Apr. 28, 2005.
- Conticello, The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* 2008;9(6):229. doi: 10.1186/GB-2008-9-6-229. Epub Jun. 17, 2008.
- Cornu et al., Refining strategies to translate genome editing to the clinic. *Nat Med.* Apr. 3, 2017;23(4):415-423. doi: 10.1038/nm.4313.
- Costa et al., Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* Mar. 15, 1995;14(6):1276-85.
- Cotton et al., Insertion of a Synthetic Peptide into a Recombinant Protein Framework: A Protein Biosensor. *J Am. Chem. Soc.* Jan. 22, 1999; 121(5):1100-1. <https://doi.org/10.1021/ja983804b>.
- Covino et al., The CCL2/CCR2 Axis in the Pathogenesis of HIV-1 Infection: A New Cellular Target for Therapy? *Current Drug Targets* Dec. 2016;17(1):76-110. DOI: 10.2174/138945011701151217110917.
- Cox et al., An SCN9A channelopathy causes congenital inability to experience pain. *Nature.* Dec. 14, 2006;444(7121):894-8. doi: 10.1038/nature05413.
- Cox et al., Conditional gene expression in the mouse inner ear using Cre-loxP. *J Assoc Res Otolaryngol.* Jun. 2012;13(3):295-322. doi: 10.1007/s10162-012-0324-5. Epub Apr. 24, 2012.
- Cox et al., Congenital insensitivity to pain: novel SCN9A missense and in-frame deletion mutations. *Hum Mutat.* Sep. 2010;31(9):E1670-86. doi: 10.1002/humu.21325.
- Cox et al., RNA editing with CRISPR-Cas13. *Science.* Nov. 24, 2017;358(6366):1019-1027. doi: 10.1126/science.aaq0180. Epub Oct. 25, 2017.
- Cox et al., Therapeutic genome editing: prospects and challenges. *Nat Med.* Feb. 2015;21(2):121-31. doi: 10.1038/nm.3793.
- Cox, Proteins pinpoint double strand breaks. *Elife.* Oct. 29, 2013;2:e01561. doi: 10.7554/elife.01561.
- Crabtree et al., Three-part inventions: intracellular signaling and induced proximity. *Trends Biochem Sci.* Nov. 1996;21(11):418-22. doi: 10.1016/s0968-0004(96)20027-1.
- Cradick et al., CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* Nov. 1, 2013;41(20):9584-92. doi: 10.1093/nar/gkt714. Epub Aug. 11, 2013.
- Cradick et al., ZFN-site searches genomes for zinc finger nuclease target sites and off-target sites. *BMC Bioinformatics.* May 13, 2011;12:152. doi: 10.1186/1471-2105-12-152.
- Cradick et al., Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. *Mol Ther.* May 2010;18(5):947-54. doi: 10.1038/mt.2010.20. Epub Feb. 16, 2010.
- Crick, On protein synthesis. *Symp Soc Exp Biol.* 1958;12:138-63.
- Cronican et al., A class of human proteins that deliver functional proteins into mammalian cells *in vitro* and *in vivo*. *Chem Biol.* Jul. 29, 2011;18(7):833-8. doi: 10.1016/j.chembiol.2011.07.003.

(56)

References Cited**OTHER PUBLICATIONS**

- Cronican et al., Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chem Biol.* Aug. 20, 2010;5(8):747-52. doi: 10.1021/cb1001153.
- Crystal, Transfer of genes to humans: early lessons and obstacles to success. *Science.* Oct. 20, 1995;270(5235):404-10. doi: 10.1126/science.270.5235.404.
- Cui et al., Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res.* May 19, 2016;44(9):4243-51. doi: 10.1093/nar/gkw223. Epub Apr. 8, 2016.
- Cui et al., m6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep.* Mar. 14, 2017;18(11):2622-2634. doi: 10.1016/j.celrep.2017.02.059.
- Cui et al., Review of CRISPR/Cas9 sgRNA Design Tools. *Interdiscip Sci.* Jun. 2018;10(2):455-465. doi: 10.1007/s12539-018-0298-z. Epub Apr. 11, 2018.
- Cui et al., Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat Biotechnol.* Jan. 2011;29(1):64-7. doi: 10.1038/nbt.1731. Epub Dec. 12, 2010.
- Cunningham et al., Ensembl 2015. *Nucleic Acids Res.* Jan. 2015;43(Database issue):D662-9. doi: 10.1093/nar/gku1010. Epub Oct. 28, 2014.
- Cupples et al., A set of lacZ mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc Natl Acad Sci U S A.* Jul. 1989;86(14):5345-9.
- D'Adda di Fagagna et al., The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep.* Jan. 2003;4(1):47-52.
- Dahlem et al., Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.* 2012;8(8):e1002861. doi: 10.1371/journal.pgen.1002861. Epub Aug. 16, 2012.
- Dahlgren et al., A novel mutation in ribosomal protein S4 that affects the function of a mutated RF1. *Biochimie.* Aug. 2000;82(8):683-91.
- Dahlman et al., Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat Biotechnol.* Nov. 2015;33(11):1159-61. doi: 10.1038/nbt.3390.
- Dandage et al., beditor: A Computational Workflow for Designing Libraries of Guide RNAs for CRISPR-Mediated Base Editing. *Genetics.* Jun. 2019;212(2):377-385. doi: 10.1534/genetics.119.302089. Epub Apr. 1, 2019.
- Dang et al., Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* Dec. 15, 2015;16:280. doi: 10.1186/s13059-015-0846-3.
- Das et al., The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure.* May 2004;12(5):819-29. doi: 10.1016/j.str.2004.02.032.
- Dassa et al., Fractured genes: a novel genomic arrangement involving new split inteins and a new homing endonuclease family. *Nucleic Acids Res.* May 2009;37(8):2560-73. doi: 10.1093/nar/gkp095. Epub Mar. 5, 2009.
- Dassa et al., Trans protein splicing of cyanobacterial split inteins in endogenous and exogenous combinations. *Biochemistry.* Jan. 9, 2007;46(1):322-30. doi: 10.1021/bi0611762.
- Database EBI Accession No. ADE34233 Jan. 29, 2004.
- Database EBI Accession No. BFF09785. May 31, 2018. 2 pages.
- Database EBI Accession No. BGE38086. Jul. 25, 2019. 2 pages.
- Database UniProt Accession No. G8I3E0. Jan. 14, 2012.
- Datsenko et al., One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* Jun. 6, 2000;97(12):6640-5.
- Davidson et al., Viral vectors for gene delivery to the nervous system. *Nat Rev Neurosci.* May 2003;4(5):353-64. doi: 10.1038/nrn1104.
- Davis et al., DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res.* Jun. 2013;2(3):130-143.
- Davis et al., Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat Chem Biol.* May 2015;11(5):316-8. doi: 10.1038/nchembio.1793. Epub Apr. 6, 2015.
- De Felipe et al., Co-translational, intraribosomal cleavage of poly-peptides by the foot-and-mouth disease virus 2A peptide. *J Biol Chem.* Mar. 28, 2003;278(13):11441-8. doi: 10.1074/jbc.M211644200. Epub Jan. 8, 2003.
- De La Peña et al., The Hammerhead Ribozyme: A Long History for a Short RNA. *Molecules.* Jan. 4, 2017;22(1):78. doi: 10.3390/molecules22010078.
- De Souza, Primer: genome editing with engineered nucleases. *Nat Methods.* Jan. 2012;9(1):27.
- De Wit et al., The Human CD4+ T Cell Response against Mumps Virus Targets a Broadly Recognized Nucleoprotein Epitope. *J Virol.* Mar. 5, 2019;93(6):e01883-18. doi: 10.1128/JVI.01883-18.
- Dean et al., Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study.* *Science.* Sep. 27, 1996;273(5283):1856-62. doi: 10.1126/science.273.5283.1856.
- DeKosky et al., Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. *Proc Natl Acad Sci U S A.* May 10, 2016;113(19):E2636-45. doi: 10.1073/pnas.1525510113. Epub Apr. 25, 2016.
- Delebecque et al., Organization of intracellular reactions with rationally designed RNA assemblies. *Science.* Jul. 22, 2011;333(6041):470-4. doi: 10.1126/science.1206938. Epub Jun. 23, 2011.
- Deltcheva et al., CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature.* Mar. 31, 2011;471(7340):602-7. doi: 10.1038/nature09886.
- Deng et al., Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res.* Jul. 27, 2015;43(13):6557-67. doi: 10.1093/nar/gkv596. Epub Jun. 11, 2015.
- Denizio et al., Harnessing natural DNA modifying activities for editing of the genome and epigenome. *Curr Opin Chem Biol.* Aug. 2018;45:10-17. doi: 10.1016/j.cbpa.2018.01.016. Epub Feb. 13, 2018.
- Deriano et al., Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu Rev Genet.* 2013;47:433-55. doi: 10.1146/annurev-genet-110711-155540. Epub Sep. 11, 2013.
- Deussing, Targeted mutagenesis tools for modelling psychiatric disorders. *Cell Tissue Res.* Oct. 2013;354(1):9-25. doi: 10.1007/s00441-013-1708-5. Epub Sep. 10, 2013.
- Dever et al., CRISPR/Cas9 β -globin gene targeting in human haematopoietic stem cells. *Nature.* Nov. 17, 2016;539(7629):384-389. doi: 10.1038/nature20134. Epub Nov. 7, 2016.
- Deverman et al., Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat Biotechnol.* Feb. 2016;34(2):204-9. doi: 10.1038/nbt.3440. Epub Feb. 1, 2016.
- Devigili et al., Paroxysmal itch caused by gain-of-function Nav1.7 mutation. *Pain.* Sep. 2014;155(9):1702-1707. doi: 10.1016/j.pain.2014.05.006. Epub May 10, 2014.
- Dianov et al., Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res.* Apr. 1, 2013;41(6):3483-90. doi: 10.1093/nar/gkt076. Epub Feb. 13, 2013.
- Dicarlo et al., Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Research.* Apr. 2013;41(7):4336-43. doi: 10.1093/nar/gkt135. Epub Mar. 4, 2013.
- Dicarlo et al., Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat Biotechnol.* Dec. 2015;33(12):1250-1255. doi: 10.1038/nbt.3412. Epub Nov. 16, 2015.
- Dickey et al., Single-stranded DNA-binding proteins: multiple domains for multiple functions. *Structure.* Jul. 2, 2013;21(7):1074-84. doi: 10.1016/j.str.2013.05.013.
- Dickinson et al., Experimental interrogation of the path dependence and stochasticity of protein evolution using phage-assisted continuous evolution. *Proc Natl Acad Sci USA.* May 2013;110(22):9007-12.
- Dillon, Regulating gene expression in gene therapy. *Trends Biotechnol.* May 1993;11(5):167-73. doi: 10.1016/0167-7799(93)90109-M.

(56)

References Cited

OTHER PUBLICATIONS

- Ding et al., A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*. Feb. 7, 2013;12(2):238-51. doi: 10.1016/j.stem.2012.11.011. Epub Dec. 13, 2012.
- Ding et al., Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ Res*. Aug. 15, 2014;115(5):488-92. doi: 10.1161/CIRCRESAHA.115.304351. Epub Jun. 10, 2014.
- Dingwall et al., Nuclear targeting sequences—a consensus? *Trends Biochem Sci*. Dec. 1991;16(12):478-81. doi: 10.1016/0968-0004(91)90184-w.
- Diver et al., Single-Step Synthesis of Cell-Permeable Protein Dimerizers That Activate Signal Transduction and Gene Expression. *J Am Chem Soc*. Jun. 4, 1997;119(22):5106-5109. https://doi.org/10.1021/ja963891c.
- Dixon et al., Reengineering orthogonally selective riboswitches. *Proc Natl Acad Sci U S A*. Feb. 16, 2010;107(7):2830-5. doi: 10.1073/pnas.0911209107. Epub Jan. 26, 2010.
- Doench et al., Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*. Feb. 2016;34(2):184-191. doi: 10.1038/nbt.3437.
- Doench et al., Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol*. Dec. 2014;32(12):1262-7. doi: 10.1038/nbt.3026. Epub Sep. 3, 2014.
- Dolan et al., Trans-splicing with the group I intron ribozyme from Azoarcus. *RNA*. Feb. 2014;20(2):202-13. doi: 10.1261/rna.041012.113. Epub Dec. 16, 2013.
- Doman et al., Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat Biotechnol*. May 2020;38(5):620-628. doi: 10.1038/s41587-020-0414-6. Epub Feb. 10, 2020.
- Dominissini et al., Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*. Apr. 29, 2012;485(7397):201-6. doi: 10.1038/nature11112.
- Dorgan et al., An enzyme-coupled continuous spectrophotometric assay for S-adenosylmethionine-dependent methyltransferases. *Anal Biochem*. Mar. 15, 2006;350(2):249-55. doi: 10.1016/j.ab.2006.01.004. Epub Feb. 7, 2006.
- Dormiani et al., Long-term and efficient expression of human β-globin gene in a hematopoietic cell line using a new site-specific integrating non-viral system. *Gene Ther*. Aug. 2015;22(8):663-74. doi: 10.1038/gt.2015.30. Epub Apr. 1, 2015.
- Dorr et al., Reprogramming the specificity of sortase enzymes. *Proc Natl Acad Sci U S A*. Sep. 16, 2014;111(37):13343-8. doi: 10.1073/pnas.1411179111. Epub Sep. 3, 2014.
- Doudna et al., Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*. Nov. 28, 2014;346(6213):1258096. doi: 10.1126/science.1258096.
- Doudna, The promise and challenge of therapeutic genome editing. *Nature*. Feb. 2020;578(7794):229-236. doi: 10.1038/s41586-020-1978-5. Epub Feb. 12, 2020.
- Dove et al., Conversion of the omega subunit of *Escherichia coli* RNA polymerase into a transcriptional activator or an activation target. *Genes Dev*. Mar. 1, 1998;12(5):745-54.
- Doyon et al., Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc*. Feb. 22, 2006;128(7):2477-84.
- Doyon et al., Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol*. Jun. 2008;26(6):702-8. doi: 10.1038/nbt1409. Epub May 25, 2008.
- Drake, A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA*. Aug. 15, 1991;88(16):7160-4.
- Drost et al., Inactivation of DNA mismatch repair by variants of uncertain significance in the PMS2 gene. *Hum Mutat*. Nov. 2013;34(11):1477-80. doi: 10.1002/humu.22426. Epub Sep. 11, 2013.
- Duan et al., Enhancement of muscle gene delivery with pseudotyped adeno-associated virus type 5 correlates with myoblast differentiation. *J Virol*. Aug. 2001;75(16):7662-71. doi: 10.1128/JVI.75.16.7662-7671.2001.
- Dubois et al., Retroviral RNA Dimerization: From Structure to Functions. *Front Microbiol*. Mar. 22, 2018;9:527. doi: 10.3389/fmicb.2018.00527.
- Dumas et al., Designing logical codon reassignment—Expanding the chemistry in biology. *Chem Sci*. Jan. 1, 2015;6(1):50-69. doi: 10.1039/c4sc01534g. Epub Jul. 14, 2014. Review.
- Dunaime, Breakthrough method means CRISPR just got a lot more relevant to human health. *The Verge*. Apr. 20, 2016. http://www.theverge.com/2016/4/20/11450262/crispr-base-editing-single-nucleotides-dna-gene-liu-harvard.
- Dunbar et al., Gene therapy comes of age. *Science*. Jan. 12, 2018;359(6372):eaan4672. doi: 10.1126/science.aan4672.
- Dupuy et al., Le syndrome de De La Chapelle [De La Chapelle syndrome]. *Presse Med*. Mar. 3, 2001;30(8):369-72. French.
- Durai et al., A bacterial one-hybrid selection system for interrogating zinc finger-DNA interactions. *Comb Chem High Throughput Screen*. May 2006;9(4):301-11.
- Durai et al., Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res*. Oct. 26, 2005;33(18):5978-90. doi: 10.1093/nar/gki912.
- During et al., Controlled release of dopamine from a polymeric brain implant: in vivo characterization. *Ann Neurol*. Apr. 1989;25(4):351-6.
- East-Seletsky et al., Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* Oct. 2016;538(7624):270-3.
- Edlund et al., Cell-specific expression of the rat insulin gene: evidence for role of two distinct 5' flanking elements. *Science*. Nov. 22, 1985;230(4728):912-6. doi: 10.1126/science.3904002.
- Edwards et al., An *Escherichia coli* tyrosine transfer RNA is a leucine-specific transfer RNA in the yeast *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*. Feb. 15, 1991;88(4):1153-6.
- Edwards et al., Crystal structures of the thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. *Structure*. Sep. 2006;14(9):1459-68.
- Eick et al., Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty. *Mol Biol Evol*. Feb. 1, 2017;34(2):247-261. doi: 10.1093/molbev/msw223.
- Eiler et al., Structural Basis for the Fast Self-Cleavage Reaction Catalyzed by the Twister Ribozyme. *Proc Natl Acad Sci U S A*. Sep. 9, 2014;111(36):13028-33. doi: 10.1073/pnas.1414571111. Epub Aug. 25, 2014.
- Eltoukhy et al., Nucleic acid-mediated intracellular protein delivery by lipid-like nanoparticles. *Biomaterials*. Aug. 2014;35(24):6454-61. doi: 10.1016/j.biomaterials.2014.04.014. Epub May 13, 2014.
- Emery et al., HCN2 ion channels play a central role in inflammatory and neuropathic pain. *Science*. Sep. 9, 2011;333(6048):1462-6. doi: 10.1126/science.1206243.
- Endo et al., Toward establishing an efficient and versatile gene targeting system in higher plants. *Biocatalysis and Agricultural Biotechnology* 2014;3,(1):2-6.
- Engel et al., The emerging role of mRNA methylation in normal and pathological behavior. *Genes Brain Behav*. Mar. 2018;17(3):e12428. doi: 10.1111/gbb.12428. Epub Nov. 17, 2017.
- Engelward et al., Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proc Natl Acad Sci U S A*. Nov. 25, 1997;94(24):13087-92.
- England, Unnatural amino acid mutagenesis: a precise tool for probing protein structure and function. *Biochemistry*. Sep. 21, 2004;43(37):11623-9.
- Enyeart et al., Biotechnological applications of mobile group II introns and their reverse transcriptases: gene targeting, RNA-seq, and non-coding RNA analysis. *Mobile DNA* 5, 2 (2014). https://doi.org/10.1186/1759-8753-5-2. https://doi.org/10.1186/1759-8753-5-2.
- Epstein, HSV-1-based amplicon vectors: design and applications. *Gene Ther*. Oct. 2005;12 Suppl 1:S154-8. doi: 10.1038/sj.gt.3302617.
- Eriksson et al., Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature*. May 15, 2003;423(6937):293-8. doi: 10.1038/nature01629. Epub Apr. 25, 2003. PMID: 12714972.

(56)

References Cited**OTHER PUBLICATIONS**

- Estacion et al., A sodium channel gene SCN9A polymorphism that increases nociceptor excitability. *Ann Neurol.* Dec. 2009;66(6):862-6. doi: 10.1002/ana.21895.
- Esveld et al., A system for the continuous directed evolution of biomolecules. *Nature.* Apr. 28, 2011;472(7344):499-503. doi: 10.1038/nature09929. *Epib.* Apr. 10, 2011.
- Esveld et al., Genome-scale engineering for systems and synthetic biology. *Mol Syst Biol.* 2013;9:641. doi: 10.1038/msb.2012.66.
- Esveld et al., Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods.* Nov. 2013;10(11):1116-21. doi: 10.1038/nmeth.2681. *Epib.* Sep. 29, 2013.
- Evans et al., Protein trans-splicing and cyclization by a naturally split intein from the dnaE gene of *Synechocystis* species PCC6803. *J Biol Chem.* Mar. 31, 2000;275(13):9091-4. doi: 10.1074/jbc.275.13.9091.
- Evans et al., Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* Nov. 1998;7(11):2256-64. doi: 10.1002/pro.5560071103.
- Evans et al., The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. *J Biol Chem.* Jun. 25, 1999;274(26):18359-63. doi: 10.1074/jbc.274.26.18359.
- Evans et al., The in vitro ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. *J Biol Chem.* Feb. 12, 1999;274(7):3923-6. doi: 10.1074/jbc.274.7.3923.
- Evers et al., CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol.* Jun. 2016;34(6):631-3. doi: 10.1038/nbt.3536. *Epib.* Apr. 25, 2016.
- Fagerlund et al., The Cpf1 CRISPR-Cas protein expands genome-editing tools. *Genome Biology.* Nov. 17, 2015;16:251. <https://doi.org/10.1186/s13059-015-0824-9>.
- Falnes et al., DNA repair by bacterial AlkB proteins. *Res Microbiol.* Oct. 2003;154(8):531-8. doi: 10.1016/S0923-2508(03)00150-5.
- Falnes et al., Repair of methyl lesions in DNA and RNA by oxidative demethylation. *Neuroscience.* Apr. 14, 2007;145(4):1222-32. doi: 10.1016/j.neuroscience.2006.11.018. *Epib.* Dec. 18, 2006.
- Fang et al., Synthetic Studies Towards Halichondrins: Synthesis of the Left Halves of Norhalichondrins and Homohalichondrins. *Tetrahedron Letters* 1992;33(12):1557-1560.
- Farboud et al., Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. *Genetics.* Apr. 2015;199(4):959-71. doi: 10.1534/genetics.115.175166. *Epib.* Feb. 18, 2015.
- Farhood et al., Codelivery to mammalian cells of a transcriptional factor with cis-acting element using cationic liposomes. *Anal Biochem.* Feb. 10, 1995;225(1):89-93.
- Fawcett et al., Transposable elements controlling I-R hybrid dysgenesis in *D. melanogaster* are similar to mammalian LINEs. *Cell.* Dec. 26, 1986;47(6):1007-15. doi: 10.1016/0092-8674(86)90815-9.
- Feldstein et al., Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA. *Gene.* Oct. 15, 1989;82(1):53-61. doi: 10.1016/0378-1119(89)90029-2.
- Felletti et al., Twister Ribozymes as Highly Versatile Expression Platforms for Artificial Riboswitches. *Nat Commun.* Sep. 27, 2016;7:12834. doi: 10.1038/ncomms12834.
- Feng et al., Crystal structures of the human RNA demethylase Alkbh5 reveal basis for substrate recognition. *J Biol Chem.* Apr. 25, 2014;289(17):11571-11583. doi: 10.1074/jbc.M113.546168. *Epib.* Mar. 10, 2014.
- Feng et al., Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell.* Nov. 29, 1996;87(5):905-16. doi: 10.1016/s0092-8674(00)81997-2.
- Ferretti et al., Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A.* Apr. 10, 2001;98(8):4658-63.
- Ferry et al., Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. *Nat Commun.* Mar. 3, 2017;8:14633. doi: 10.1038/ncomms14633.
- Feuk, Inversion variants in the human genome: role in disease and genome architecture. *Genome Med.* Feb. 12, 2010;2(2):11. doi: 10.1186/gm132.
- Filippov et al., A novel type of RNase III family proteins in eukaryotes. *Gene.* Mar. 7, 2000;245(1):213-21. doi: 10.1016/s0378-1119(99)00571-5.
- Filippova et al., Guide RNA modification as a way to improve CRISPR/Cas9-based genome-editing systems. *Biochimie.* Dec. 2019;167:49-60. doi: 10.1016/j.biochi.2019.09.003. *Epib.* Sep. 4, 2019.
- Fine et al., Trans-spliced Cas9 allows cleavage of HBB and CCR5 genes in human cells using compact expression cassettes. *Scientific Reports* 2015;5(1):Article No. 10777. doi:10.1038/srep10777. With Supplementary Information.
- Fire et al., Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature.* Feb. 19, 1998;391(6669):806-11. doi: 10.1038/35888.
- Fischbach et al., Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. *Proc Natl Acad Sci U S A.* Jul. 17, 2007;104(29):11951-6. doi: 10.1073/pnas.0705348104. *Epib.* Jul. 9, 2007.
- Fischer et al., Cryptic epitopes induce high-titer humoral immune response in patients with cancer. *J Immunol.* Sep. 1, 2010;185(5):3095-102. doi: 10.4049/jimmunol.0902166. *Epib.* Jul. 26, 2010.
- Fitzjohn, Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution.* Dec. 2012;3(6):1084-92 .doi: 10.1111/j.2041-210X.2012.00234.x.
- Flajole et al., Woodchuck hepatitis virus enhancer I and enhancer II are both involved in -myc2 activation in woodchuck liver tumors. *J Virol.* Jul. 1998;72(7):6175-80. doi: 10.1128/JVI.72.7.6175-6180. 1998.
- Flaman et al., A rapid PCR fidelity assay. *Nucleic Acids Res.* Aug. 11, 1994;22(15):3259-60. doi: 10.1093/nar/22.15.3259.
- Flynn et al., CRISPR-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human iPS cells. *Exp Hematol.* Oct. 2015;43(10):838-848.e3. doi: 10.1016/j.exphem.2015.06.002. *Epib.* Jun. 19, 2015. Including supplementary figures and data.
- Fogg et al., New applications for phage integrases. *J Mol Biol.* Jul. 29, 2014;426(15):2703-16. doi: 10.1016/j.jmb.2014.05.014. *Epib.* May 22, 2014.
- Fogg et al., Genome Integration and Excision by a New Streptomyces Bacteriophage, ?Joe. *Appl Environ Microbiol.* Feb. 15, 2017;83(5):e02767-16. doi: 10.1128/AEM.02767-16.
- Fonfara et al., Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* Feb. 2014;42(4):2577-90. doi: 10.1093/nar/gkt1074. *Epib.* Nov. 22, 2013. Including Supplementary Information.
- Forster et al., Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site. *Cell.* Jul. 3, 1987;50(1):9-16. doi: 10.1016/0092-8674(87)90657-x.
- Fortini et al., Different DNA polymerases are involved in the short-and long-patch base excision repair in mammalian cells. *Biochemistry.* Mar. 17, 1998;37(11):3575-80. doi: 10.1021/bi972999h.
- Fouts et al., Sequencing *Bacillus anthracis* typing phages gamma and cherry reveals a common ancestry. *J Bacteriol.* May 2006;188(9):3402-8. doi: 10.1128/JB.188.9.3402-3408.2006.
- Freitas et al., Mechanisms and signals for the nuclear import of proteins. *Curr Genomics.* Dec. 2009;10(8):550-7. doi: 10.2174/1389209789503941.
- Freshney, Culture of Animal Cells. A Manual of Basic Technique. Alan R. Liss, Inc. New York. 1983;4.
- Fu et al., Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol.* Mar. 2014;32(3):279-84. doi: 10.1038/nbt.2808. *Epib.* Jan. 26, 2014.
- Fu et al., High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol.* Sep. 2013;31(9):822-6. doi: 10.1038/nbt.2623. *Epib.* Jun. 23, 2013.

(56)

References Cited**OTHER PUBLICATIONS**

- Fu et al., Promises and Pitfalls of Intracellular Delivery of Proteins. *Bioconjugate Chemistry*. Aug. 2014;25:1602-8.
- Fuchs et al., Polyarginine as a multifunctional fusion tag. *Protein Sci.* Jun. 2005;14(6):1538-44.
- Fujisawa et al., Disease-associated mutations in CIAS1 induce cathepsin B-dependent rapid cell death of human THP-1 monocytic cells. *Blood*. Apr. 1, 2007;109(7):2903-11.
- Fukui et al., DNA Mismatch Repair in Eukaryotes and Bacteria. *J Nucleic Acids*. Jul. 27, 2010;2010. pii: 260512. doi: 10.4061/2010/260512.
- Fung et al., Repair at single targeted DNA double-strand breaks in pluripotent and differentiated human cells. *PLoS One*. 2011;6(5):e20514. doi: 10.1371/journal.pone.0020514. Epub May 25, 2011.
- Furukawa et al., In vitro selection of allosteric ribozymes that sense the bacterial second messenger c-di-GMP. *Methods Mol Biol.* 2014;1111:209-20. doi: 10.1007/978-1-62703-755-6_15.
- Fusi et al., In Silico Predictive Modeling of CRISPR/Cas9 guide efficiency. Jun. 26, 2015; bioRxiv. <http://dx.doi.org/10.1101/021568>.
- Gaj et al., 3rd. Genome engineering with custom recombinases. *Methods Enzymol.* 2014;546:79-91. doi: 10.1016/B978-0-12-801185-0.00004-0.
- Gaj et al., A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. *Nucleic Acids Res.* Feb. 6, 2013;41(6):3937-46.
- Gaj et al., Enhancing the specificity of recombinase-mediated genome engineering through dimer interface redesign. *J Am Chem Soc.* Apr. 2, 2014;136(13):5047-56. doi: 10.1021/ja4130059. Epub Mar. 20, 2014.
- Gaj et al., Expanding the scope of site-specific recombinases for genetic and metabolic engineering. *Biotechnol Bioeng.* Jan. 2014;111(1):1-15. doi: 10.1002/bit.25096. Epub Sep. 13, 2013.
- Gaj et al., Structure-guided reprogramming of serine recombinase DNA sequence specificity. *Proc Natl Acad Sci U S A.* Jan. 11, 2011;108(2):498-503. doi: 10.1073/pnas.1014214108. Epub Dec. 27, 2010.
- Gaj et al., ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* Jul. 2013;31(7):397-405. doi: 10.1016/j.tibtech.2013.04.004. Epub May 9, 2013.
- Gajula, Designing an Elusive C•G?G•C Crispr Base Editor. *Trends Biochem Sci.* Feb. 2019;44(2):91-94. doi: 10.1016/j.tibs.2018.10.004. Epub Nov. 13, 2018.
- Gallo et al., A novel pathogenic PSEN1 mutation in a family with Alzheimer's disease: phenotypical and neuropathological features. *J Alzheimers Dis.* 2011;25(3):425-31. doi: 10.3233/JAD-2011-110185.
- Gangopadhyay et al., Precision Control of CRISPR-Cas9 Using Small Molecules and Light. *Biochemistry*. Jan. 29, 2019;58(4):234-244. doi: 10.1021/acs.biochem.8b01202. Epub Jan. 22, 2019.
- Gao et al., Cationic liposome-mediated gene transfer. *Gene Ther.* Dec. 1995;2(10):710-22.
- Gao et al., DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute. *Nat Biotechnol.* Jul. 2016;34(7):768-73. doi: 10.1038/nbt.3547. Epub May 2, 2016.
- Gao et al., Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J Integr Plant Biol.* Apr. 2014;56(4):343-9. doi: 10.1111/jipb.12152. Epub Mar. 6, 2014.
- Gao et al., Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature*. Jan. 11, 2018;553(7687):217-221. doi: 10.1038/nature25164. Epub Dec. 20, 2017.
- Gapinske et al., CRISPR-SKIP: programmable gene splicing with single base editors. *Genome Biol.* Aug. 15, 2018;19(1):107. doi: 10.1186/s13059-018-1482-5.
- Garcia et al., Transglycosylation: a mechanism for RNA modification (and editing?). *Bioorg Chem.* Jun. 2005;33(3):229-51. doi: 10.1016/j.bioorg.2005.01.001. Epub Feb. 23, 2005.
- Gardlik et al., Vectors and delivery systems in gene therapy. *Med Sci Monit.* Apr. 2005;11(4):RA110-21. Epub Mar. 24, 2005.
- Garibyan et al., Use of the rpoB gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst)*. May 13, 2003;2(5):593-608.
- Garneau et al., The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. Nov. 4, 2010;468(7320):67-71. doi: 10.1038/nature09523.
- Gasiunas et al., Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A.* Sep. 25, 2012;109(39):E2579-86. Epub Sep. 4, 2012. Supplementary materials included.
- Gasiunas et al., RNA-dependent DNA endonuclease Cas9 of the CRISPR system: Holy Grail of genome editing? *Trends Microbiol.* Nov. 2013;21(11):562-7. doi: 10.1016/j.tim.2013.09.001. Epub Oct. 1, 2013.
- Gaudelli et al., Programmable base editing of AoT to GoC in genomic DNA without DNA cleavage. *Nature*. Nov. 23, 2017;551(7681):464-471. doi: 10.1038/nature24644. Epub Oct. 25, 2017. Erratum in: *Nature*. May 2, 2018.
- Gearing, Addgene blog. CRISPR 101: Cas9 nickase design and homology directed repair. 2018. pp. 1-12. <https://blog.addgene.org/crispr-101-cas9-nickase-design-and-homology-directed-repair>. Last retrieved online Jun. 25, 2021.
- Gehrke et al., An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat Biotechnol.* Nov. 2018;36(10):977-982. doi: 10.1038/nbt.4199. Epub Jul. 30, 2018. GenBank Accession No. J01600.1. Brooks et al., *E.coli* dam gene coding for DNA adenine methylase. Apr. 26, 1993.
- GenBank Accession No. U07651.1. Lu, *Escherichia coli* K12 negative regulator of replication initiation (seqA) gene, complete cds. Jul. 19, 1994.
- GenBank Submission; NIH/NCBI Accession No. 4UN5_B. Anders et al., Jul. 23, 2014. 5 pages.
- GenBank Submission; NIH/NCBI, Accession No. AAA66622.1. Martinelli et al., May 18, 1995. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. AGT42196. Farzadfar et al., Nov. 2, 2013. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. AIT42264.1. Hyun et al., Oct. 15, 2014. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. AKA60242.1. Tong et al., Apr. 5, 2015. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. AKQ21048.1. Gilles et al., Jul. 19, 2015. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. AKS40380.1. Nodvig et al., Aug. 2, 2015. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. APG80656.1. Burstein et al., Dec. 10, 2016. 1 pages.
- GenBank Submission; NIH/NCBI, Accession No. AYD60528.1. Ram et al., Oct. 2, 2018. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. BDB43378. Zhang et al., Aug. 11, 2016. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. J04623. Kita et al., Apr. 26, 1993. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. KR710351.1. Sahni et al., Jun. 1, 2015. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. NC_002737.1. Ferretti et al., Jun. 27, 2013. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_015683.1. Trost et al., Jul. 6, 2013. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_016782.1. Trost et al., Jun. 11, 2013. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_016786.1. Trost et al., Aug. 28, 2013. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_017053.1. Fittipaldi et al., Jul. 6, 2013. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_017317.1. Trost et al., Jun. 11, 2013. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_017861.1. Heidelberg et al., Jun. 11, 2013. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_018010.1. Lucas et al., Jun. 11, 2013. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. NC_018721.1. Feng et al., Jun. 11, 2013. 1 pages.

- (56) **References Cited**
- OTHER PUBLICATIONS
- GenBank Submission; NIH/NCBI, Accession No. NC_021284.1. 1. No Author Listed, Oct. 4, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NC_021314.1. 1. No Author Listed, Dec. 26, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_038431314. 1. No Author Listed, Dec. 26, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_038432938. 1. No Author Listed, Dec. 26, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_038434062. 1. No Author Listed, Dec. 26, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_044924278. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_047338501. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_048327215. 1. No Author Listed, Jun. 26, 2015. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_049519324. 1. No Author Listed, Jul. 20, 2015. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_060798984. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_062913273. 1. Haft et al., Oct. 9, 2019, 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_072754838. 1. No Author Listed, Sep. 23, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_095142515. 1. No Author Listed, Sep. 23, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_118538418. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_119223642. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_119227726. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_119623382. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_132221894. 1. No Author Listed, Sep. 23, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_133478044. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_002342100. 1. Bernardini et al., Jun. 10, 2013. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_002344900. 1. Gundogdu et al., Mar. 19, 2014. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_009137104. 1. Davison, Aug. 13, 2018. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_009283008. 1. Bernardini et al., Sep. 23, 2016. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_820832.1. Makarova et al., Aug. 27, 2013. 2 pages.
- George et al., Adenosine deaminases acting on RNA, RNA editing, and interferon action. *J Interferon Cytokine Res.* Jan. 2011;31(1):99-117. doi: 10.1089/jir.2010.0097. Epub Dec. 23, 2010. PMID: 21182352; PMCID: PMC3034097.
- Gerard et al., Influence on stability in *Escherichia coli* of the carboxy-terminal structure of cloned Moloney murine leukemia virus reverse transcriptase. *DNA*. Aug. 1986;5(4):271-9. doi: 10.1089/dna.1986.5.271.
- Gerard et al., Purification and characterization of the DNA polymerase and RNase H activities in Moloney murine sarcoma-leukemia virus. *J Virol.* Apr. 1975;15(4):785-97. doi: 10.1128/JVI.15.4.785-797.1975.
- Gerard et al., The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res.* Jul. 15, 2002;30(14):3118-29. doi: 10.1093/nar/gkf417.
- Gerber et al., An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science*. Nov. 5, 1999;286(5442):1146-9. doi: 10.1126/science.286.5442.1146.
- Gerber et al., RNA editing by base deamination: more enzymes, more targets, new mysteries. *Trends Biochem Sci.* Jun. 2001;26(6):376-84.
- Gersbach et al., Directed evolution of recombinase specificity by split gene reassembly. *Nucleic Acids Res.* Jul. 2010;38(12):4198-206. doi: 10.1093/nar/gkq125. Epub Mar. 1, 2010.
- GenBank Submission; NIH/NCBI, Accession No. WP_032460140. 1. No Author Listed, Oct. 4, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_032461047. 1. No Author Listed, Oct. 4, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_032462016. 1. Haft et al., Oct. 4, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_032462936. 1. No Author Listed, Oct. 4, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_032464890. 1. No Author Listed, Oct. 4, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_038431314. 1. No Author Listed, Dec. 26, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_038432938. 1. No Author Listed, Dec. 26, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_038434062. 1. No Author Listed, Dec. 26, 2014. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_044924278. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_047338501. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_048327215. 1. No Author Listed, Jun. 26, 2015. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_049519324. 1. No Author Listed, Jul. 20, 2015. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_060798984. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_062913273. 1. Haft et al., Oct. 9, 2019, 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_072754838. 1. No Author Listed, Sep. 23, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_095142515. 1. No Author Listed, Sep. 23, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_118538418. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_119223642. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_119227726. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_119623382. 1. No Author Listed, Oct. 13, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_132221894. 1. No Author Listed, Sep. 23, 2019. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. WP_133478044. 1. Haft et al., Oct. 9, 2019. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_002342100. 1. Bernardini et al., Jun. 10, 2013. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_002344900. 1. Gundogdu et al., Mar. 19, 2014. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_009137104. 1. Davison, Aug. 13, 2018. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_009283008. 1. Bernardini et al., Sep. 23, 2016. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. YP_820832.1. Makarova et al., Aug. 27, 2013. 2 pages.
- George et al., Adenosine deaminases acting on RNA, RNA editing, and interferon action. *J Interferon Cytokine Res.* Jan. 2011;31(1):99-117. doi: 10.1089/jir.2010.0097. Epub Dec. 23, 2010. PMID: 21182352; PMCID: PMC3034097.
- Gerard et al., Influence on stability in *Escherichia coli* of the carboxy-terminal structure of cloned Moloney murine leukemia virus reverse transcriptase. *DNA*. Aug. 1986;5(4):271-9. doi: 10.1089/dna.1986.5.271.
- Gerard et al., Purification and characterization of the DNA polymerase and RNase H activities in Moloney murine sarcoma-leukemia virus. *J Virol.* Apr. 1975;15(4):785-97. doi: 10.1128/JVI.15.4.785-797.1975.
- Gerard et al., The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res.* Jul. 15, 2002;30(14):3118-29. doi: 10.1093/nar/gkf417.
- Gerber et al., An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science*. Nov. 5, 1999;286(5442):1146-9. doi: 10.1126/science.286.5442.1146.
- Gerber et al., RNA editing by base deamination: more enzymes, more targets, new mysteries. *Trends Biochem Sci.* Jun. 2001;26(6):376-84.
- Gersbach et al., Directed evolution of recombinase specificity by split gene reassembly. *Nucleic Acids Res.* Jul. 2010;38(12):4198-206. doi: 10.1093/nar/gkq125. Epub Mar. 1, 2010.

(56)

References Cited**OTHER PUBLICATIONS**

- Gersbach et al., Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. *Nucleic Acids Res.* Sep. 1, 2011;39(17):7868-78. doi: 10.1093/nar/gkr421. Epub Jun. 7, 2011.
- Gahafarokhi et al., Blastocyst Formation Rate and Transgene Expression are Associated with Gene Insertion into Safe and Non-Safe Harbors in the Cattle Genome. *Sci Rep.* Nov. 13, 2017;7(1):15432. doi: 10.1038/s41598-017-15648-3.
- Gibson et al., Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* May 2009;6(5):343-5. doi: 10.1038/nmeth.1318. Epub Apr. 12, 2009.
- Gil, Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation. *Cell.* May 8, 1987;49(3):399-406. doi: 10.1016/0092-8674(87)90292-3.
- Gilbert et al., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell.* 2013;154(2):442-51.
- Gilleron et al., Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nat Biotechnol.* Jul. 2013;31(7):638-46. doi: 10.1038/nbt.2612. Epub Jun. 23, 2013.
- Glasgow et al., DNA-binding properties of the Hin recombinase. *J Biol Chem.* Jun. 15, 1989;264(17):10072-82.
- Glassner et al., Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proc Natl Acad Sci U S A.* Aug. 18, 1998;95(17):9997-10002.
- Goldberg et al., Epigenetics: a landscape takes shape. *Cell.* Feb. 23, 2007;128(4):635-8. doi: 10.1016/j.cell.2007.02.006.
- Goldberg et al., Loss-of-function mutations in the Nav1.7 gene underlie congenital indifference to pain in multiple human populations. *Clin Genet.* Apr. 2007;71(4):311-9. doi: 10.1111/j.1399-0004.2007.00790.x.
- Gong et al., Active DNA demethylation by oxidation and repair. *Cell Res.* Dec. 2011;21(12):1649-51. doi: 10.1038/cr.2011.140. Epub Aug. 23, 2011.
- Gonzalez et al., An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell.* Aug. 7, 2014;15(2):215-26. doi: 10.1016/j.stem.2014.05.018. Epub Jun. 12, 2014.
- Goodnough et al., Development of a delivery vehicle for intracellular transport of botulinum neurotoxin antagonists. *FEBS Lett.* Feb. 27, 2002;513(2-3):163-8.
- Gordley et al., Evolution of programmable zinc finger-recombinases with activity in human cells. *J Mol Biol.* Mar. 30, 2007;367(3):802-13. Epub Jan. 12, 2007.
- Gordley et al., Synthesis of programmable integrases. *Proc Natl Acad Sci U S A.* Mar. 31, 2009;106(13):5053-8. doi: 10.1073/pnas.0812502106. Epub Mar. 12, 2009.
- Gou et al., Designing single guide RNA for CIRSPKR-Cas9 base editor by deep learning. Peer reviewed Thesis/Dissertation. UCLA Electronic Theses and Dissertations. Jan. 1, 2019. Retrieved from the Internet via <https://escholarship.org/uc/item/7vf9z54t>. Last accessed on Apr. 29, 2021.
- Grainge et al., The integrase family of recombinase: organization and function of the active site. *Mol Microbiol.* Aug. 1999;33(3):449-56.
- Gregory et al., Integration site for Streptomyces phage phiBT1 and development of site-specific integrating vectors. *J Bacteriol.* Sep. 2003;185(17):5320-3. doi: 10.1128/jb.185.17.5320-5323.2003.
- Griffiths, Endogenous retroviruses in the human genome sequence. *Genome Biol.* 2001;2(6):Reviews1017. doi: 10.1186/GB-2001-2-6-reviews1017. Epub Jun. 5, 2001.
- Grindley et al., Mechanisms of site-specific recombination. *Annu Rev Biochem.* 2006;75:567-605. doi: 10.1146/annurev.biochem.73.011303.073908.
- Grishok et al., Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control *C. elegans* Developmental Timing. *J. Biol.* Jul. 13, 2001;106(1):P23-4.
- Groher et al., Synthetic riboswitches—A tool comes of age. *Biochim Biophys Acta.* Oct. 2014;1839(10):964-973. doi: 10.1016/j.bbaram.2014.05.005. Epub May 17, 2014.
- Groth et al., Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics.* Apr. 2004;166(4):1775-82. doi: 10.1534/genetics.166.4.1775.
- Groth et al., Phage integrases: biology and applications. *J Mol Biol.* Jan. 16, 2004;335(3):667-78.
- Gruber et al., Strategies for measuring evolutionary conservation of RNA secondary structures. *BMC Bioinformatics.* Feb. 26, 2008;9:122. doi: 10.1186/1471-2105-9-122.
- Gruber et al., The Vienna RNA websuite. *Nucleic Acids Res.* Jul. 1, 2008;36(Web Server issue):W70-4. doi: 10.1093/nar/gkn188. Epub Apr. 19, 2008.
- Grunebaum et al., Recent advances in understanding and managing adenosine deaminase and purine nucleoside phosphorylase deficiencies. *Curr Opin Allergy Clin Immunol.* Dec. 2013;13(6):630-8. doi: 10.1097/ACI.0000000000000006.
- Grünewald et al., Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature.* May 2019;569(7756):433-437. doi: 10.1038/s41586-019-1161-z. Epub Apr. 17, 2019.
- Guedon et al., Current gene therapy using viral vectors for chronic pain. *Mol Pain.* May 13, 2015;11:27. doi: 10.1186/s12990-015-0018-1.
- Guilinger et al., Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat Methods.* Apr. 2014;11(4):429-35. doi: 10.1038/nmeth.2845. Epub Feb. 16, 2014.
- Guilinger et al., Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol.* Jun. 2014;32(6):577-82. doi: 10.1038/nbt.2909. Epub Apr. 25, 2014.
- Gumulya et al., Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering. *Biochem J.* Jan. 1, 2017;474(1):1-19. doi: 10.1042/BCJ20160507.
- Guo et al., Evolution of Tetrahymena ribozyme mutants with increased structural stability. *Nat Struct Biol.* Nov. 2002;9(11):855-61. doi: 10.1038/nsb850.
- Guo et al., Facile functionalization of FK506 for biological studies by the thiol-ene 'click' reaction. *RSC Advances.* 2014;22:11400-3.
- Guo et al., Protein tolerance to random amino acid change. *Proc Natl Acad Sci U S A.* Jun. 22, 2004;101(25):9205-10. Epub Jun. 14, 2004.
- Guo et al., Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature.* Sep. 4, 1997;389(6646):40-6.
- Gupta et al., Cross-talk between cognate and noncognate RpoE sigma factors and Zn(2+)-binding anti-sigma factors regulates photooxidative stress response in *Azospirillum brasilense*. *Antioxid Redox Signal.* Jan. 1, 2014;20(1):42-59. doi: 10.1089/ars.2013.5314. Epub Jul. 19, 2013.
- Gupta et al., Sequences in attB that affect the ability of phiC31 integrase to synapse and to activate DNA cleavage. *Nucleic Acids Res.* 2007;35(10):3407-19. doi: 10.1093/nar/gkm206. Epub May 3, 2007.
- Guzman et al., Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol.* 1995;177(14):4121-4130.
- Haapaniemi et al., CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat Med.* Jul. 2018;24(7):927-930. doi: 10.1038/s41591-018-0049-z. Epub Jun. 11, 2018.
- Haddada et al., Gene therapy using adenovirus vectors. *Curr Top Microbiol Immunol.* 1995;199 (Pt 3):297-306. doi: 10.1007/978-3-642-79586-2_14.
- Haeussler et al., Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* Jul. 5, 2016;17(1):148. doi: 10.1186/s13059-016-1012-2.
- Halbert et al., Repeat transduction in the mouse lung by using adeno-associated virus vectors with different serotypes. *J Virol.* Feb. 2000;74(3):1524-32. doi: 10.1128/jvi.74.3.1524-1532.2000.

(56)

References Cited**OTHER PUBLICATIONS**

- Hale et al., RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell.* Nov. 25, 2009;139(5):945-56. doi: 10.1016/j.cell.2009.07.040.
- Halmi et al., Targeted CRISPR/dCas9-mediated reactivation of epigenetically silenced genes suggests limited escape from the inactive X chromosome. 2nd Intl Conf on Epigenetics and Bioengineering. Oct. 4, 2018; Retrieved from the Internet: <https://aiche.confex.com/aiche/epibio18/webprogram/paper544785.html>. Retrieved Jun. 29, 2020.
- Halperin et al., CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature.* Aug. 2018;560(7717):248-252. doi: 10.1038/s41586-018-0384-8. Epub Aug. 1, 2018.
- Halvas et al., Role of murine leukemia virus reverse transcriptase deoxyribonucleoside triphosphate-binding site in retroviral replication and in vivo fidelity. *J Virol.* Nov. 2000;74(22):10349-58. doi: 10.1128/jvi.74.22.10349-10358.2000.
- Hamano-Takaku et al., A mutant *Escherichia coli* tyrosyl-tRNA synthetase utilizes the unnatural amino acid azatyrosine more efficiently than tyrosine. *J Biol Chem.* Dec. 22, 2000;275(51):40324-8.
- Han, New CRISPR/Cas9-based Tech Edits Single Nucleotides Without Breaking DNA. Genome Web, Apr. 20, 2016. <https://www.genomeweb.com/gene-silencing/gene-editing/new-crisprcas9-based-tech-edits-single-nucleotides-without-breaking-dna>.
- Handa et al., Template-assisted synthesis of adenine-mutagenized cDNA by a retroelement protein complex. *Nucleic Acids Res.* Oct. 12, 2018;46(18):9711-9725. doi: 10.1093/nar/gky620.
- Hanson et al., Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol.* Jan. 2018;19(1):20-30. doi: 10.1038/nrm.2017.91. Epub Oct. 11, 2017.
- Hardt et al., Missense variants in hMLH1 identified in patients from the German HNPCC consortium and functional studies. *Fam Cancer.* Jun. 2011;10(2):273-84. doi: 10.1007/s10689-011-9431-4.
- Harms et al., Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat Rev Genet.* Aug. 2013;14(8):559-71. doi: 10.1038/nrg3540.
- Harmsen et al., DNA mismatch repair and oligonucleotide end-protection promote base-pair substitution distal from a CRISPR/Cas9-induced DNA break. *Nucleic Acids Res.* Apr. 6, 2018;46(6):2945-2955. doi: 10.1093/nar/gky076.
- Harrington et al., A thermostable Cas9 with increased lifetime in human plasma. *Nat Commun.* Nov. 10, 2017;8(1):1424. doi: 10.1038/s41467-017-01408-4. Posted May 16, 2017 as bioRxiv preprint. [DOI.org/10.1101/138867](https://doi.org/10.1101/138867).
- Harris et al., RNA Editing Enzyme APOBEC1 and Some of Its Homologs Can Act as DNA Mutators. *Mol Cell.* Nov. 2002;10(5):1247-53.
- Hartung et al., Correction of metabolic, craniofacial, and neurologic abnormalities in MPS I mice treated at birth with adeno-associated virus vector transducing the human alpha-L-iduronidase gene. *Mol Ther.* Jun. 2004;9(6):866-75.
- Hartung et al., Cre mutants with altered DNA binding properties. *J Biol Chem.* Sep. 4, 1998;273(36):22884-91.
- Hasadsri et al., Functional protein delivery into neurons using polymeric nanoparticles. *J Biol Chem.* Mar. 13, 2009;284(11):6972-81. doi: 10.1074/jbc.M805956200. Epub Jan. 7, 2009.
- Hasegawa et al., Spontaneous mutagenesis associated with nucleotide excision repair in *Escherichia coli*. *Genes Cells.* May 2008;13(5):459-69. doi: 10.1111/j.1365-2443.2008.01185.x.
- Hector et al., CDKL5 variants: Improving our understanding of a rare neurologic disorder. *Neurol Genet.* Dec. 15, 2017;3(6):e200. doi: 10.1212/NXG.0000000000000200.
- Heidenreich et al., Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *EMBO J.* May 1, 2003;22(9):2274-83. doi: 10.1093/emboj/cdg203.
- Held et al., In vivo correction of murine hereditary tyrosinemia type I by phiC31 integrase-mediated gene delivery. *Mol Ther.* Mar. 2005;11(3):399-408. doi: 10.1016/j.ymthe.2004.11.001.
- Heller et al., Replisome assembly and the direct restart of stalled replication forks. *Nat Rev Mol Cell Biol.* Dec. 2006;7(12):932-43. Epub Nov. 8, 2006.
- Hendricks et al., The *S. cerevisiae* Mag1 3-methyladenine DNA glycosylase modulates susceptibility to homologous recombination. *DNA Repair (Amst).* 2002;1(8):645-659.
- Hermonat et al., Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. *Proc Natl Acad Sci U S A.* Oct. 1984;81(20):6466-70. doi: 10.1073/pnas.81.20.6466.
- Herschhorn et al., Retroviral reverse transcriptases. *Cell Mol Life Sci.* Aug. 2010;67(16):2717-47. doi: 10.1007/s00018-010-0346-2. Epub Apr. 1, 2010.
- Herzig et al., A Novel Leu92 Mutant of HIV-1 Reverse Transcriptase with a Selective Deficiency in Strand Transfer Causes a Loss of Viral Replication. *J Virol.* Aug. 2015;89(16):8119-29. doi: 10.1128/JVI.00809-15. Epub May 20, 2015.
- Hess et al., Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat Methods.* Dec. 2016; 13(12):1036-1042. doi: 10.1038/nmeth.4038. Epub Oct. 31, 2016.
- Hickford et al., Antitumour polyether macrolides: four new halichondrins from the New Zealand deep-water marine sponge *Lissodendoryx* sp. *Bioorg Med Chem.* Mar. 15, 2009;17(6):2199-203. doi: 10.1016/j.bmc.2008.10.093. Epub Nov. 19, 2008.
- Hida et al., Directed evolution for drug and nucleic acid; delivery. *Adv Drug Deliv Rev.* Dec. 22, 2007;59(15):1562-78. Epub Aug. 28, 2007.; Review.
- Higgs et al., Genetic complexity in sickle cell disease. *Proc Natl Acad Sci U S A.* Aug. 19, 2008;105(33):11595-6. doi: 10.1073/pnas.0806633105. Epub Aug. 11, 2008.
- Hilbers et al., New developments in structure determination of pseudoknots. *Biopolymers.* 1998;48(2-3):137-53. doi: 10.1002/(SICI)1097-0282(1998)48:2<137::AID-BIP4>3.0.CO;2-H.
- Hill et al., Functional analysis of conserved histidines in ADP-glucose pyrophosphorylase from *Escherichia coli*. *Biochem Biophys Res Commun.* Mar. 17, 1998;244(2):573-7.
- Hille et al., The Biology of CRISPR-Cas: Backward and Forward. *Cell.* Mar. 8, 2018;172(6):1239-1259. doi: 10.1016/j.cell.2017.11.032.
- Hilton et al., Enabling functional genomics with genome engineering. *Genome Res.* Oct. 2015;25(10):1442-55. doi: 10.1101/gr.190124.115.
- Hirano et al., Site-specific recombinases as tools for heterologous gene integration. *Appl Microbiol Biotechnol.* Oct. 2011;92(2):227-39. doi: 10.1007/s00253-011-3519-5. Epub Aug. 7, 2011. Review.
- Hirano et al., Structural Basis for the Altered PAM Specificities of Engineered CRISPR-Cas9. *Mol Cell.* Mar. 17, 2016;61(6):886-94. doi: 10.1016/j.molcel.2016.02.018.
- Hoang et al., UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol.* Feb. 1, 2018;35(2):518-522. doi: 10.1093/molbev/msx281.
- Hockemeyer et al., Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol.* Sep. 2009;27(9):851-7. doi: 10.1038/nbt.1562. Epub Aug. 13, 2009.
- Hockemeyer et al., Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol.* Jul. 7, 2011;29(8):731-4. doi: 10.1038/nbt.1927.
- Hoernes et al., Translating the epitranscriptome. *Wiley Interdiscip Rev RNA.* Jan. 2017;8(1):e1375. doi: 10.1002/wrna.1375. Epub Jun. 27, 2016.
- Hoess et al., DNA specificity of the Cre recombinase resides in the 25 kDa carboxyl domain of the protein. *J Mol Biol.* Dec. 20, 1990;216(4):873-82. doi: 10.1016/S0022-2836(99)80007-2.
- Holden et al., Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature.* Nov. 6, 2008;456(7218):121-4. doi: 10.1038/nature07357. Epub Oct. 12, 2008.
- Hollis et al., Phage integrases for the construction and manipulation of transgenic mammals. *Reprod Biol Endocrinol.* Nov. 7, 2003;1:79. doi: 10.1186/1477-7827-1-79.
- Holsinger et al., Signal transduction in T lymphocytes using a conditional allele of Sos. *Proc Natl Acad Sci U S A.* Oct. 10, 1995;92(21):9810-4. doi: 10.1073/pnas.92.21.9810.

(56)

References Cited**OTHER PUBLICATIONS**

- Holt et al., Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol.* Aug. 2010;28(8):839-47. doi: 10.1038/nbt.1663. *Epub Jul. 2, 2010.*
- Hondares et al., Peroxisome Proliferator-activated Receptor ? (PPAR?) Induces PPAR? Coactivator 1? (PGC-1?) Gene Expression and Contributes to Thermogenic Activation of Brown Fat. *J Biol. Chem.* Oct. 2011; 286(50):43112-22. doi: 10.1074/jbc.M111.252775.
- Hoogenboom et al., Natural and designer binding sites made by phage display technology. *Immunol Today.* Aug. 2000;21(8):371-8.
- Horvath et al., CRISPR/Cas, the immune system of bacteria and archaea. *Science.* Jan. 8, 2010;327(5962):167-70. doi: 10.1126/science.1179555.
- Horvath et al., Diversity, Activity, and Evolution of CRISPR Loci in *Streptococcus thermophilus*. *J Bacteriol.* Feb. 2008;190(4):1401-12. doi: 10.1128/JB.01415-07. *Epub Dec. 7, 2007.*
- Hotta et al., [Neurotropic viruses—classification, structure and characteristics]. *Nihon Rinsho.* Apr. 1997;55(4):777-82. Japanese.
- Hou et al., Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci U S A.* Sep. 24, 2013;110(39):15644-9. doi: 10.1073/pnas.1313587110. *Epub Aug. 12, 2013.*
- Houdebine, The methods to generate transgenic animals and to control transgene expression. *J Biotechnol.* Sep. 25, 2002;98(2-3):145-60.
- Housden et al., Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. *Sci Signal.* Sep. 8, 2015;8(393):rs9. doi: 10.1126/scisignal.aab3729.
- Howard et al., Intracerebral drug delivery in rats with lesion-induced memory deficits. *J Neurosurg.* Jul. 1989;71(1):105-12.
- Hower et al., Shape-based peak identification for ChIP-Seq. *BMC Bioinformatics.* Jan. 12, 2011;12:15. doi: 10.1186/1471-2105-12-15.
- Hsu et al., DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol.* Sep. 2013;31(9):827-32. doi: 10.1038/nbt.2647. *Epub Jul. 21, 2013.*
- Hsu et al., DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol.* Sep. 2013;31(9):827-32. doi: 10.1038/nbt.2647. *Epub Jul. 21, 2013. Supplementary Information. 27 pages.*
- Hu et al., Chemical Biology Approaches to Genome Editing: Understanding, Controlling, and Delivering Programmable Nucleases. *Cell Chem Biol.* Jan. 21, 2016;23(1):57-73. doi: 10.1016/j.chembiol.2015.12.009.
- Hu et al., Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature.* Apr. 5, 2018;556(7699):57-63 and Extended/Supplementary Data. doi: 10.1038/nature26155. *Epub Feb. 28, 2018. 21 pages.*
- Hu et al., Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature.* Apr. 5, 2018;556(7699):57-63. doi: 10.1038/nature26155. *Epub Feb. 28, 2018.*
- Hua et al., Expanding the base editing scope in rice by using Cas9 variants. *Plant Biotechnol J.* Feb. 2019;17(2):499-504. doi: 10.1111/pbi.12993. *Epub Oct. 5, 2018.*
- Hua et al., Precise A*T to G*C Base Editing in the Rice Genome. *Mol Plant.* Apr. 2, 2018;11(4):627-630. doi: 10.1016/j.molp.2018.02.007. *Epub Feb. 21, 2018.*
- Huang et al., Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat Biotechnol.* Jun. 2019;37(6):626-631. doi: 10.1038/s41587-019-0134-y. *Epub May 20, 2019. Including Supplementary Information.*
- Huang et al., Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol.* Aug. 5, 2011;29(8):699-700. doi: 10.1038/nbt.1939.
- Huggins et al., Flap endonuclease 1 efficiently cleaves base excision repair and DNA replication intermediates assembled into nucleosomes. *Mol Cell.* Nov. 2002;10(5):1201-11. doi: 10.1016/s1097-2765(02)00736-0.
- Humbert et al., Targeted gene therapies: tools, applications, optimization. *Crit Rev Biochem Mol Biol.* May-Jun. 2012;47(3):264-81. doi: 10.3109/10409238.2012.658112.
- Hung et al., Protein localization in disease and therapy. *J Cell Sci.* Oct. 15, 2011;124(Pt 20):3381-92. doi: 10.1242/jcs.089110.
- Hurt et al., Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proc Natl Acad Sci U S A.* Oct. 14, 2003;100(21):12271-6. *Epub Oct. 3, 2003.*
- Husimi, Selection and evolution of bacteriophages in cellstat. *Adv Biophys.* ; 1989;25:1-43. Review.
- Hwang et al., Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol.* Mar. 2013;31(3):227-9. doi: 10.1038/nbt.2501. *Epub Jan. 29, 2013.*
- Hwang et al., Efficient In Vivo Genome Editing Using RNA-Guided Nucleases. *Nat Biotechnol.* Mar. 2013; 31(3): 227-229. doi: 10.1038/nbt.2501. *Epub Jan. 29, 2013.*
- Hwang et al., Web-based design and analysis tools for CRISPR base editing. *BMC Bioinformatics.* Dec. 27, 2018;19(1):542. doi: 10.1186/s12859-018-2585-4.
- Ibba et al., Relaxing the substrate specificity of an aminoacyl-tRNA synthetase allows in vitro and in vivo synthesis of proteins containing unnatural amino acids. *FEBS Lett.* May 15, 1995;364(3):272-5.
- Ibba et al., Substrate specificity is determined by amino acid binding pocket size in *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry.* Jun. 14, 1994;33(23):7107-12.
- Ihry et al., p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med.* Jul. 2018;24(7):939-946. doi: 10.1038/s41591-018-0050-6. *Epub Jun. 11, 2018.*
- Iida et al., A site-specific, conservative recombination system carried by bacteriophage P1. Mapping the recombinase gene cin and the cross-over sites cix for the inversion of the C segment. *EMBO J.* 1982;1(11):1445-53.
- Iida et al., The Min DNA inversion enzyme of plasmid p15B of *Escherichia coli* 15T: a new member of the Din family of site-specific recombinases. *Mol Microbiol.* Jun. 1990;4(6):991-7. doi: 10.1111/j.1365-2958.1990.tb00671.x.
- Ikediobi et al., Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. *Mol Cancer Ther.* Nov. 2006;5(11):2606-12. *Epub Nov. 6, 2006.*
- Imanishi et al., Detection of N6-methyladenosine based on the methyl-sensitivity of MazF RNA endonuclease. *Chem Commun (Camb).* Nov. 30, 2017;53(96):12930-12933. doi: 10.1039/c7cc07699a.
- Imburgio et al., Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants. *Biochemistry.* Aug. 29, 2000;39(34):10419-30.
- Ingram, A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature.* Oct. 13, 1956;178(4537):792-4. doi: 10.1038/178792a0.
- Irion et al., Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat Biotechnol.* Dec. 2007;25(12):1477-82. doi: 10.1038/nbt1362. *Epub Nov. 25, 2007.*
- Irrthum et al., Congenital hereditary lymphedema caused by a mutation that inactivates VEGFR3 tyrosine kinase. *Am J Hum Genet.* Aug. 2000;67(2):295-301. *Epub Jun. 9, 2000.*
- Isaacs et al., Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol.* Jul. 2004;22(7):841-7. doi: 10.1038/nbt986. *Epub Jun. 20, 2004.*
- Ishino et al., Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol.* Dec. 1987;169(12):5429-33.
- Iwai et al., Circular beta-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* Oct. 8, 1999;459(2):166-72. doi: 10.1016/s0014-5793(99)01220-x.
- Iwai et al., Highly efficient protein trans-splicing by a naturally split DnaE intein from *Nostoc punctiforme*. *FEBS Lett.* Mar. 20, 2006;580(7):1853-8. doi: 10.1016/j.febslet.2006.02.045. *Epub Feb. 24, 2006.*

(56)

References Cited**OTHER PUBLICATIONS**

- Jaffrey et al., Emerging links between m6A and misregulated mRNA methylation in cancer. *Genome Med.* Jan. 12, 2017;9(1):2. doi: 10.1186/s13073-016-0395-8.
- Jamieson et al., Drug discovery with engineered zinc-finger proteins. *Nat Rev Drug Discov.* May 2003;2(5):361-8.
- Jansen et al., Backbone and nucleobase contacts to glucosamine-6-phosphate in the glmS ribozyme. *Nat Struct Mol Biol.* Jun. 2006;13(6):517-23. Epub May 14, 2006.
- Jansen et al., Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol.* Mar. 2002;43(6):1565-75.
- Jardine et al., HIV-1 Vaccines. Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science.* Jul. 10, 2015;349(6244):156-61. doi: 10.1126/science.aac5894. Epub Jun. 18, 2015.
- Jasin et al., Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol.* Nov. 1, 2013;5(11):a012740. doi: 10.1101/cshperspect.a012740.
- Jeggo, DNA breakage and repair. *Adv Genet.* 1998;38:185-218. doi: 10.1016/s0065-2660(08)60144-3.
- Jemielity et al., Novel "anti-reverse" cap analogs with superior translational properties. *RNA.* Sep. 2003;9(9):1108-22. doi: 10.1261/rna.5430403.
- Jenkins et al., Comparison of a preQ1 riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J Biol Chem.* Jul. 15, 2011;286(28):24626-37. doi: 10.1074/jbc.M111.230375. Epub May 18, 2011.
- Jeong et al., Measurement of deoxyinosine adduct: Can it be a reliable tool to assess oxidative or nitrosative DNA damage? *Toxicol Lett.* Oct. 17, 2012;214(2):226-33. doi: 10.1016/j.toxlet.2012.08.013. Epub Aug. 23, 2012.
- Jia et al., The MLH1 ATPase domain is needed for suppressing aberrant formation of interstitial telomeric sequences. *DNA Repair (Amst).* May 2018;65:20-25. doi: 10.1016/j.dnarep.2018.03.002. Epub Mar. 7, 2018.
- Jiang et al., CRISPR-Cas9 Structures and Mechanisms. *Annu Rev Biophys.* May 22, 2017;46:505-529. doi: 10.1146/annurev-biophys-062215-010822. Epub Mar. 30, 2017.
- Jiang et al., RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol.* Mar. 2013;31(3):233-9. doi: 10.1038/nbt.2508. Epub Jan. 29, 2013.
- Jiang et al., Structural Biology. A Cas9-guide RNA Complex Preorganized for Target DNA Recognition. *Science.* Jun. 26, 2015;348(6242):1477-81. doi: 10.1126/science.aab1452.
- Jiang et al., Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science.* Feb. 19, 2016;351(6275):867-71. doi: 10.1126/science.aad8282. Epub Jan. 14, 2016.
- Jin et al., Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science.* Apr. 19, 2019;364(6437):292-295. doi: 10.1126/science.aaw7166. Epub Feb. 28, 2019.
- Jinek et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* Aug. 17, 2012;337(6096):816-21. doi: 10.1126/science.1225829. Epub Jun. 28, 2012.
- Jinek et al., RNA-programmed genome editing in human cells. *Elife.* Jan. 29, 2013;2:e00471. doi: 10.7554/elife.00471.
- Jinek et al., Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science.* Mar. 14, 2014;343(6176):1247997. doi: 10.1126/science.1247997. Epub Feb. 6, 2014.
- Jiricny, The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol.* May 2006;7(5):335-46. doi: 10.1038/nrm1907.
- Johann et al., GLVR1, a receptor for gibbon ape leukemia virus, is homologous to a phosphate permease of *Neurospora crassa* and is expressed at high levels in the brain and thymus. *J Virol.* Mar. 1992;66(3):1635-40. doi: 10.1128/JVI.66.3.1635-1640.1992.
- Johansson et al., RNA Recognition by the MS2 Phage Coat Protein. *Seminars in Virology.* 1997;8(3):176-85. <https://doi.org/10.1006/smvy.1997.0120>.
- Johansson et al., Selenocysteine in proteins-properties and biotechnological use. *Biochim Biophys Acta.* Oct. 30, 2005;1726(1):1-13. Epub Jun. 1, 2005.
- Johns et al., The promise and peril of continuous in vitro evolution. *J Mol Evol.* Aug. 2005;61(2):253-63. Epub Jun. 27, 2005.
- Johnson et al., Trans insertion-splicing: ribozyme-catalyzed insertion of targeted sequences into RNAs. *Biochemistry.* Aug. 9, 2005;44(31):10702-10. doi: 10.1021/bi0504815.
- Joho et al., Identification of a region of the bacteriophage T3 and T7 RNA polymerases that determines promoter specificity. *J Mol Biol.* Sep. 5, 1990;215(1):31-9.
- Jore et al., Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol.* May 2011;18(5):529-36. doi: 10.1038/nsmb.2019. Epub Apr. 3, 2011.
- Joung et al., TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol.* Jan. 2013;14(1):49-55. doi: 10.1038/nrm3486. Epub Nov. 21, 2012.
- Joyce et al., Amplification, mutation and selection of catalytic RNA. *Gene.* Oct. 15, 1989;82(1):83-7. doi: 10.1016/0378-1119(89)90033-4.
- Jusiak et al., Comparison of Integrases Identifies Bxb1-GA Mutant as the Most Efficient Site-Specific Integrase System in Mammalian Cells. *ACS Synth Biol.* Jan. 18, 2019;8(1):16-24. doi: 10.1021/acssynbio.8b00089. Epub Jan. 9, 2019.
- Jyothi et al., Translocation Down syndrome. *Indian J Med Sci.* Mar. 2002;56(3):122-6.
- Kacian et al., Purification of the DNA polymerase of avian myeloblastosis virus. *Biochim Biophys Acta.* Sep. 24, 1971;246(3):365-83. doi: 10.1016/0005-2787(71)90773-8.
- Kaczmarczyk et al., Manipulating the Prion Protein Gene Sequence and Expression Levels with CRISPR/Cas9. *PLoS One.* Apr. 29, 2016;11(4):e0154604. doi: 10.1371/journal.pone.0154604.
- Kadoch et al., Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell.* Mar. 28, 2013;153(1):71-85. doi: 10.1016/j.cell.2013.02.036.
- Kahmann et al., G inversion in bacteriophage Mu DNA is stimulated by a site within the invertase gene and a host factor. *Cell.* Jul. 1985;41(3):771-80. doi: 10.1016/s0092-8674(85)80058-1.
- Kaiser et al., Gene therapy. Putting the fingers on gene repair. *Science.* Dec. 23, 2005;310(5756):1894-6.
- Kakiyama et al., A peptide release system using a photo-cleavable linker in a cell array format for cell-toxicity analysis. *Polymer J.* Feb. 27, 2013;45:535-9.
- Kalyaanamoorthy et al., ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* Jun. 2017;14(6):587-589. doi: 10.1038/nmeth.4285. Epub May 8, 2017.
- Kandavelou et al., Targeted manipulation of mammalian genomes using designed zinc finger nucleases. *Biochem Biophys Res Commun.* Oct. 9, 2009;388(1):56-61. doi: 10.1016/j.bbrc.2009.07.112. Epub Jul. 25, 2009.
- Kang et al., Structural Insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. *Mol Cell.* Mar. 27, 2009;33(6):784-90. doi: 10.1016/j.molcel.2009.02.019. Epub Mar. 12, 2009.
- Kang et al., Precision genome engineering through adenine base editing in plants. *Nat Plants.* Jul. 2018;4(7):427-431. doi: 10.1038/s41477-018-0178-x. Epub Jun. 4, 2018. Erratum in: *Nat Plants.* Sep. 2018;4(9):730.
- Kao et al., Cleavage specificity of *Saccharomyces cerevisiae* flap endonuclease 1 suggests a double-flap structure as the cellular substrate. *J Biol Chem.* Apr. 26, 2002;277(17):14379-89. doi: 10.1074/jbc.M110662200. Epub Feb. 1, 2002.
- Kappel et al., Regulating gene expression in transgenic animals. *Curr Opin Biotechnol.* Oct. 1992;3(5):548-53.
- Karimova et al., Discovery of Nigri/nox and Panto/pox site-specific recombinase systems facilitates advanced genome engineering. *Sci Rep.* Jul. 22, 2016;6:30130. doi: 10.1038/srep30130.
- Karimova et al., Vika/vox, a novel efficient and specific Cre/loxP-like site-specific recombination system. *Nucleic Acids Res.* Jan. 2013;41(2):e37. doi: 10.1093/nar/gks1037. Epub Nov. 9, 2012.
- Karpenshif et al., From yeast to mammals: recent advances in genetic control of homologous recombination. *DNA Repair (Amst).* Oct. 1, 2012;11(10):781-8. doi: 10.1016/j.dnarep.2012.07.001. Epub Aug. 11, 2012. Review.

(56)

References Cited**OTHER PUBLICATIONS**

- Karpinsky et al., Directed evolution of a recombinase that excises the provirus of most HIV-1 primary isolates with high specificity. *Nat Biotechnol.* Apr. 2016;34(4):401-9. doi: 10.1038/nbt.3467. Epub Feb. 22, 2016.
- Katafuchi et al., DNA polymerases involved in the incorporation of oxidized nucleotides into DNA: their efficiency and template base preference. *Mutat Res.* Nov. 28, 2010;703(1):24-31. doi: 10.1016/j.mrgentox.2010.06.004. Epub Jun. 11, 2010.
- Kato et al., Improved purification and enzymatic properties of three forms of reverse transcriptase from avian myeloblastosis virus. *J Virol Methods.* Dec. 1984;9(4):325-39. doi: 10.1016/0166-0934(84)90058-2.
- Katoh et al., MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* Apr. 2013;30(4):772-80. doi: 10.1093/molbev/mst010. Epub Jan. 16, 2013.
- Kaufman et al., Translational efficiency of polycistronic mRNAs and their utilization to express heterologous genes in mammalian cells. *EMBO J.* Jan. 1987;6(1):187-93.
- Kavli et al., Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO J.* Jul. 1, 1996;15(13):3442-7.
- Kawasaki et al., Enhanced crossover SCRATCHY: construction and high-throughput screening of a combinatorial library containing multiple non-homologous crossovers. *Nucleic Acids Res.* Nov. 1, 2003;31(21):e126.
- Kay et al., Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat Med.* Jan. 2001;7(1):33-40.
- Kaya et al., A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl. Acad. Sci. USA* Apr. 2016;113(15):4057-62.
- Keijzers et al., Human exonuclease 1 (EXO1) activity characterization and its function on flap structures. *Biosci Rep.* Apr. 25, 2015;35(3):e00206. doi: 10.1042/BSR20150058.
- Kellendonk et al., Regulation of Cre recombinase activity by the synthetic steroid RU 486. *Nucleic Acids Res.* Apr. 15, 1996;24(8):1404-11.
- Kelman, PCNA: structure, functions and interactions. *Oncogene.* Feb. 13, 1997;14(6):629-40. doi: 10.1038/sj.onc.1200886.
- Keravala et al., A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. *Mol Genet Genomics.* Aug. 2006;276(2):135-46. doi: 10.1007/s00438-006-0129-5. Epub May 13, 2006.
- Kessel et al., Murine developmental control genes. *Science.* Jul. 27, 1990;249(4967):374-9. doi: 10.1126/science.1974085.
- Kessler et al., Gene delivery to skeletal muscle results in sustained expression and systemic delivery of a therapeutic protein. *Proc Natl Acad Sci U S A.* Nov. 26, 1996;93(24):14082-7. doi: 10.1073/pnas.93.24.14082.
- Ketha et al., Application of bioinformatics-coupled experimental analysis reveals a new transport-competent nuclear localization signal in the nucleoprotein of Influenza A virus strain. *BMC Cell Biol.* Apr. 28, 2008; 9:22. <https://doi.org/10.1186/1471-2121-9-22>.
- Kiga et al., An engineered *Escherichia coli* tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc Natl Acad Sci U S A.* Jul. 23, 2002;99(15):9715-20. Epub Jul. 3, 2002.
- Kilbride et al., Determinants of product topology in a hybrid Cre-Tn3 resolvase site-specific recombination system. *J Mol Biol.* Jan. 13, 2006;355(2):185-95. Epub Nov. 9, 2005.
- Kilcher et al., Brochothrix thermosphacta bacteriophages feature heterogeneous and highly mosaic genomes and utilize unique prophage insertion sites. *J Bacteriol.* Oct. 2010;192(20):5441-53. doi: 10.1128/JB.00709-10. Epub Aug. 13, 2010.
- Kim et al., DJ-1, a novel regulator of the tumor suppressor PTEN. *Cancer Cell.* 2005;7(3):263-273.
- Kim et al., Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat Biotechnol.* Apr. 2019;37(4):430-435. doi: 10.1038/s41587-019-0050-1. Epub Mar. 4, 2019.
- Kim et al., A library of TAL effector nucleases spanning the human genome. *Nat Biotechnol.* Mar. 2013;31(3):251-8. doi: 10.1038/nbt.2517. Epub Feb. 17, 2013.
- Kim et al., An anionic human protein mediates cationic liposome delivery of genome editing proteins into mammalian cells. *Nat Commun.* Jul. 2, 2019;10(1):2905. doi: 10.1038/s41467-019-10828-3.
- Kim et al., Evaluating and Enhancing Target Specificity of Gene-Editing Nucleases and Deaminases. *Annu Rev Biochem.* Jun. 20, 2019;88:191-220. doi: 10.1146/annurev-biochem-013118-111730. Epub Mar. 18, 2019.
- Kim et al., Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat Biotechnol.* May 2017;35(5):475-480. doi: 10.1038/nbt.3852. Epub Apr. 10, 2017.
- Kim et al., High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One.* 2011;6(4):e18556. doi: 10.1371/journal.pone.0018556. Epub Apr. 29, 2011.
- Kim et al., High-throughput analysis of the activities of xCas9, SpCas9-NG and SpCas9 at matched and mismatched target sequences in human cells. *Nat Biomed Eng.* Jan. 2020;4(1):111-124. doi: 10.1038/s41551-019-0505-1. Epub Jan. 14, 2020.
- Kim et al., Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol.* May 2017;35(5):435-437. doi: 10.1038/nbt.3816. Epub Feb. 27, 2017.
- Kim et al., Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* Jun. 2014;24(6):1012-9. doi: 10.1101/gr.171322.113. Epub Apr. 2, 2014.
- Kim et al., In vivo genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat Commun.* Feb. 21, 2017;8:14500. doi: 10.1038/ncomms14500. PMID: 28220790; PMCID: PMC5473640.
- Kim et al., In vivo high-throughput profiling of CRISPR-Cpf1 activity. *Nat Methods.* Feb. 2017;14(2):153-159. doi: 10.1038/nmeth.4104. Epub Dec. 19, 2016.
- Kim et al., Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol.* Apr. 2017;35(4):371-376. doi: 10.1038/nbt.3803. Epub Feb. 13, 2017.
- Kim et al., Mycobacteriophage Bxb1 integrates into the *Mycobacterium smegmatis* groEL1 gene. *Mol Microbiol.* Oct. 2003;50(2):463-73. doi: 10.1046/j.1365-2958.2003.03723.x.
- Kim et al., Rescue of high-specificity Cas9 variants using sgRNAs with matched 5' nucleotides. *Genome Biol.* Nov. 15, 2017;18(1):218. doi: 10.1186/s13059-017-1355-3.
- Kim et al., Structural and kinetic characterization of *Escherichia coli* TadA, the wobble-specific tRNA deaminase. *Biochemistry.* May 23, 2006;45(20):6407-16. doi: 10.1021/bi0522394. PMID: 16700551.
- Kim et al., TALENs and ZFNs are associated with different mutationsignatures. *Nat Methods.* Mar. 2013;10(3):185. doi: 10.1038/nmeth.2364. Epub Feb. 10, 2013.
- Kim et al., Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* Jul. 2009;19(7):1279-88. doi: 10.1101/gr.089417.108. Epub May 21, 2009.
- Kim et al., The role of apolipoprotein E in Alzheimer's disease. *Neuron.* Aug. 13, 2009;63(3):287-303. doi: 10.1016/j.neuron.2009.06.026.
- Kim et al., Transcriptional repression by zinc finger peptides. Exploring the potential for applications in gene therapy. *J Biol Chem.* Nov. 21, 1997;272(47):29795-800.
- King et al., No gain, no pain: NaV1.7 as an analgesic target. *ACS Chem Neurosci.* Sep. 17, 2014;5(9):749-51. doi: 10.1021/cn500171p. Epub Aug. 11, 2014.
- Kitamura et al., Uracil DNA glycosylase counteracts APOBEC3G-induced hypermutation of hepatitis B viral genomes: excision repair of covalently closed circular DNA. *PLoS Pathog.* 2013;9(5):e1003361. doi: 10.1371/journal.ppat.1003361. Epub May 16, 2013.

(56)

References Cited**OTHER PUBLICATIONS**

- Klapacz et al., Frameshift mutagenesis and microsatellite instability induced by human alkyladenine DNA glycosylase. *Mol Cell.* Mar. 26, 2010;37(6):843-53. doi: 10.1016/j.molcel.2010.01.038.
- Klauser et al., An engineered small RNA-mediated genetic switch based on a ribozyme expression platform. *Nucleic Acids Res.* May 1, 2013;41(10):5542-52. doi: 10.1093/nar/gkt253. Epub Apr. 12, 2013.
- Klein et al., Cocrystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat Struct Mol Biol.* Mar. 2009;16(3):343-4. doi: 10.1038/nsmb.1563. Epub Feb. 22, 2009.
- Kleiner et al., In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J Am Chem Soc.* Aug. 25, 2010;132(33):11779-91. doi: 10.1021/ja104903x.
- Kleinstiver et al., Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol.* Dec. 2015;33(12):1293-1298. doi: 10.1038/nbt.3404. Epub Nov. 2, 2015.
- Kleinstiver et al., Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature.* Jul. 23, 2015;523(7561):481-5 and Supplementary Materials. doi: 10.1038/nature14592. Epub Jun. 22, 2015. 27 pages.
- Kleinstiver et al., Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature.* Jul. 23, 2015;523(7561):481-5. doi: 10.1038/nature14592. Epub Jun. 22, 2015.
- Kleinstiver et al., High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature.* Jan. 28, 2016;529(7587):490-5. doi: 10.1038/nature16526. Epub Jan. 6, 2016.
- Kleinstiver et al., Monomeric site-specific nucleases for genome editing. *Proc Natl Acad Sci U S A.* May 22, 2012;109(21):8061-6. doi: 10.1073/pnas.1117984109. Epub May 7, 2012.
- Klement et al., Discrimination between bacteriophage T3 and T7 promoters by the T3 and T7 RNA polymerases depends primarily upon a three base-pair region located 10 to 12 base-pairs upstream from the start site. *J Mol Biol.* Sep. 5, 1990;215(1):21-9.
- Klippel et al., Isolation and characterization of unusual gin mutants. *EMBO J.* Dec. 1, 1988;7(12):3983-9.
- Klippel et al., The DNA invertase Gin of phage Mu: formation of a covalent complex with DNA via a phosphoserine at amino acid position 9. *EMBO J.* Apr. 1988;7(4):1229-37.
- Klompe et al., Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature.* Jul. 2019;571(7764):219-225. doi: 10.1038/s41586-019-1323-z. Epub Jun. 12, 2019.
- Knott et al., Guide-bound structures of an RNA-targeting A-cleaving CRISPR-Cas13a enzyme. *Nat Struct Mol Biol.* Oct. 2017;24(10):825-833. doi: 10.1038/nsmb.3466. Epub Sep. 11, 2017.
- Koblan et al., Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol.* Oct. 2018;36(9):843-846. doi: 10.1038/nbt.4172. Epub May 29, 2018.
- Kobori et al., Deep Sequencing Analysis of Aptazyme Variants Based on a Pistol Ribozyme. *ACS Synth Biol.* Jul. 21, 2017;6(7):1283-1288. doi: 10.1021/acssynbio.7b00057. Epub Apr. 14, 2017.
- Kohli et al., A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J Biol Chem.* Aug. 21, 2009;284(34):22898-904. doi: 10.1074/jbc.M109.025536. Epub Jun. 26, 2009.
- Kohli et al., Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification. *J Biol Chem.* Dec. 24, 2010;285(52):40956-64. doi: 10.1074/jbc.M110.177402. Epub Oct. 6, 2010.
- Koike-Yusa et al., Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol.* Mar. 2014;32(3):267-73. doi: 10.1038/nbt.2800. Epub Dec. 23, 2013.
- Kolot et al., Site promiscuity of coliphage HK022 integrase as a tool for gene therapy. *Gene Ther.* Jul. 2015;22(7):521-7. doi: 10.1038/gt.2015.9. Epub Mar. 12, 2015.
- Kolot et al., Site-specific recombination in mammalian cells expressing the Int recombinase of bacteriophage HK022. *Mol Biol Rep.* Aug. 1999;26(3):207-13. doi: 10.1023/a:1007096701720.
- Komor et al., CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell.* Jan. 12, 2017;168(1-2):20-36. doi: 10.1016/j.cell.2016.10.044.
- Komor et al., Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv.* Aug. 30, 2017;3(8):eaao4774. doi: 10.1126/sciadv.aao4774. eCollection Aug. 2017.
- Komor et al., Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature.* Apr. 20, 2016;533(7603):420-4. doi: 10.1038/nature17946.
- Komor, Editing the Genome Without Double-Stranded DNA Breaks. *ACS Chem Biol.* Feb. 16, 2018;13(2):383-388. doi: 10.1021/acscchembio.7b00710. Epub Oct. 9, 2017.
- Konermann et al., Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature.* Jan. 29, 2015;517(7536):583-8. doi: 10.1038/nature14136. Epub Dec. 10, 2014.
- Koonin et al., Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol.* 2017;37:67-78. doi: 10.1016/j.mib.2017.05.008.
- Kosicki et al., Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol.* Sep. 2018;36(8):765-771. doi: 10.1038/nbt.4192. Epub Jul. 16, 2018.
- Kotewicz et al., Cloning and overexpression of Moloney murine leukemia virus reverse transcriptase in *Escherichia coli*. *Gene.* 1985;35(3):249-58. doi: 10.1016/0378-1119(85)90003-4.
- Kotewicz et al., Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.* Jan. 11, 1988;16(1):265-77. doi: 10.1093/nar/16.1.265.
- Kotin, Prospects for the use of adeno-associated virus as a vector for human gene therapy. *Hum Gene Ther.* Jul. 1994;5(7):793-801. doi: 10.1089/hum.1994.5.7-793.
- Kouzminova et al., Patterns of chromosomal fragmentation due to uracil-DNA incorporation reveal a novel mechanism of replication-dependent double-stranded breaks. *Mol Microbiol.* Apr. 2008;68(1):202-15. doi: 10.1111/j.1365-2958.2008.06149.x.
- Kowal et al., Exploiting unassigned codons in *Micrococcus luteus* for tRNA-based amino acid mutagenesis. *Nucleic Acids Res.* Nov. 15, 1997;25(22):4685-9.
- Kowalski et al., Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery. *Mol Ther.* Apr. 10, 2019;27(4):710-728. doi: 10.1016/j.ymthe.2019.02.012. Epub Feb. 19, 2019.
- Kozak, An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* Oct. 26, 1987;15(20):8125-48. doi: 10.1093/nar/15.20.8125.
- Kraft et al., Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. *Cell Rep.* Feb. 10, 2015;10(5):833-839. doi: 10.1016/j.celrep.2015.01.016. Epub Feb. 7, 2015.
- Kremer et al., Adenovirus and adeno-associated virus mediated gene transfer. *Br Med Bull.* Jan. 1995;51(1):31-44. doi: 10.1093/oxfordjournals.bmb.a072951.
- Krokan et al., Uracil in DNA—occurrence, consequences and repair. *Oncogene.* Dec. 16, 2002;21(58):8935-48. doi: 10.1038/sj.onc.1205996.
- Krokan et al., Base excision repair. *Cold Spring Harb Perspect Biol.* Apr. 1, 2013;5(4):a012583. doi: 10.1101/cshperspect.a012583.
- Krzywkowski et al., Limited reverse transcriptase activity of phi29 DNA polymerase. *Nucleic Acids Res.* Apr. 20, 2018;46(7):3625-3632. doi: 10.1093/nar/gky190.
- Kumar et al., Gene therapy for chronic neuropathic pain: how does it work and where do we stand today? *Pain Med.* May 2011;12(5):808-22. doi: 10.1111/j.1526-4637.2011.01120.x.
- Kumar et al., Structural and functional consequences of the mutation of a conserved arginine residue in alphaA and alphaB crystallins. *J Biol Chem.* Aug. 20, 1999;274(34):24137-41.

(56)

References Cited**OTHER PUBLICATIONS**

- Kundu et al., Leucine to proline substitution by SNP at position 197 in Caspase-9 gene expression leads to neuroblastoma: a bioinformatics analysis. *3 Biotech.* 2013; 3:225-34.
- Kunkel et al., Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annu Rev Genet.* 2015;49:291-313. doi: 10.1146/annurev-genet-112414-054722.
- Kunz et al., DNA Repair in mammalian cells: Mismatched repair: variations on a theme. *Cell Mol Life Sci.* Mar. 2009;66(6):1021-38. doi: 10.1007/s00018-009-8739-9.
- Kurjan et al., Structure of a yeast pheromone gene (MF alpha): a putative alpha-factor precursor contains four tandem copies of mature alpha-factor. *Cell.* Oct. 1982;30(3):933-43. doi: 10.1016/0092-8674(82)90298-7.
- Kury et al., De Novo Disruption of the Proteasome Regulatory Subunit PSMD12 Causes a Syndromic Neurodevelopmental Disorder. *Am J Hum Genet.* Feb. 2, 2017;100(2):352-363. doi: 10.1016/j.ajhg.2017.01.003. Epub Jan. 26, 2017.
- Kuscu et al., CRISPR-Cas9-AID base editor is a powerful gain-of-function screening tool. *Nat Methods.* Nov. 29, 2016;13(12):983-984. doi: 10.1038/nmeth.4076.
- Kuscu et al., CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat Methods.* Jul. 2017;14(7):710-712. doi: 10.1038/nmeth.4327. Epub Jun. 5, 2017.
- Kuscu et al., Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol.* Jul. 2014;32(7):677-83. doi: 10.1038/nbt.2916. Epub May 18, 2014.
- Kwart et al., Precise and efficient scarless genome editing in stem cells using CORRECT. *Nat Protoc.* Feb. 2017;12(2):329-354. doi: 10.1038/nprot.2016.171. Epub Jan. 19, 2017.
- Kweon et al., Fusion guide RNAs for orthogonal gene manipulation with Cas9 and Cpf1. *Nat Commun.* Nov. 23, 2017;8(1):1723. doi: 10.1038/s41467-017-01650-w. Erratum in: *Nat Commun.* Jan. 16, 2018;9(1):303.
- Kwon et al., Chemical basis of glycine riboswitch cooperativity. *RNA.* Jan. 2008;14(1):25-34. Epub Nov. 27, 2007.
- Köhler et al., A possible approach to site-specific insertion of two different unnatural amino acids into proteins in mammalian cells via nonsense suppression. *Chem Biol.* Nov. 2003;10(11):1095-102.
- Köhler et al., Complete set of orthogonal 21st aminoacyl-tRNA synthetase-amber, ochre and opal suppressor tRNA pairs: concomitant suppression of three different termination codons in an mRNA in mammalian cells. *Nucleic Acids Res.* Dec. 1, 2004;32(21):6200-11. Print 2004.
- Kügler et al., Human synapsin 1 gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. *Gene Ther.* Feb. 2003;10(4):337-47. doi: 10.1038/sj.gt.3301905.
- Lada et al., Mutator effects and mutation signatures of editing deaminases produced in bacteria and yeast. *Biochemistry (Mosc).* Jan. 2011;76(1):131-46.
- Lakich et al., Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nat Genet.* Nov. 1993;5(3):236-41. doi: 10.1038/ng1193-236.
- Lancaster et al., Limited trafficking of a neurotropic virus through inefficient retrograde axonal transport and the type I interferon response. *PLoS Pathog.* Mar. 5, 2010;6(3):e1000791. doi: 10.1371/journal.ppat.1000791.
- Landrum et al., ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* Jan. 4, 2016;44(D1):D862-8. doi: 10.1093/nar/gkv1222. Epub Nov. 17, 2015.
- Landrum et al., ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* Jan. 2014;42(Database issue):D980-5. doi: 10.1093/nar/gkt1113. Epub Nov. 14, 2013.
- Langer et al., Chemical and Physical Structure of Polymers as Carriers for Controlled Release of Bioactive Agents: A Review. *Journal of Macromolecular Science.* 2006;23(1):61-126. DOI: 10.1080/07366578308079439.
- Langer et al., New methods of drug delivery. *Science.* Sep. 28, 1990;249(4976):1527-33.
- Larson et al., CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc.* Nov. 2013;8(11):2180-96. doi: 10.1038/nprot.2013.132. Epub Oct. 17, 2013.
- Lau et al., Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proc Natl Acad Sci U S A.* Dec. 5, 2000;97(25):13573-8.
- Lauer et al., Construction, characterization, and use of two *Listeria monocytogenes* site-specific phage integration vectors. *J Bacteriol.* Aug. 2002;184(15):4177-86. doi: 10.1128/jb.184.15.4177-4186.2002.
- La Vergne et al., Defects in type IIA von Willebrand disease: a cysteine 509 to arginine substitution in the mature von Willebrand factor disrupts a disulphide loop involved in the interaction with platelet glycoprotein Ib-IX. *Br J Haematol.* Sep. 1992;82(1):66-72.
- Lawrence et al., Supercharging proteins can impart unusual resilience. *J Am Chem Soc.* Aug. 22, 2007;129(33):10110-2. Epub Aug. 1, 2007.
- Lawyer et al., High-level expression, purification, and enzymatic characterization of full-length *Thermus aquaticus* DNA polymerase and a truncated form deficient in 5' to 3' exonuclease activity. *PCR Methods Appl.* May 1993;2(4):275-87. doi: 10.1101/gr.2.4.275.
- Lazar et al., Transforming growth factor alpha: mutation of aspartic acid 47 and leucine 48 results in different biological activities. *Mol Cell Biol.* Mar. 1988;8(3):1247-52.
- Lazarevic et al., Nucleotide sequence of the *Bacillus subtilis* temperate bacteriophage SPbeta2c. *Microbiology (Reading).* May 1999;145 (Pt 5):1055-1067. doi: 10.1099/13500872-145-5-1055.
- Le Grice et al., Purification and characterization of recombinant equine infectious anemia virus reverse transcriptase. *J Virol.* Dec. 1991;65(12):7004-7. doi: 10.1128/JVI.65.12.7004-7007.1991.
- Leaver-Fay et al., ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* 2011;487:545-74. doi: 10.1016/B978-0-12-381270-4.00019-6.
- Leconte et al., A population-based experimental model for protein evolution: effects of mutation rate and selection stringency on evolutionary outcomes. *Biochemistry.* Feb. 26, 2013;52(8): 1490-9. doi: 10.1021/bi3016185. Epub Feb. 14, 2013.
- Ledford, Gene-editing hack yields pinpoint precision. *Nature.* Apr. 20, 2016. <http://www.nature.com/news/gene-editing-hack-yields-pinpoint-precision-1.19773>.
- Lee et al., A chimeric thyroid hormone receptor constitutively bound to DNA requires retinoid X receptor for hormone-dependent transcriptional activation in yeast. *Mol Endocrinol.* Sep. 1994;8(9):1245-52.
- Lee et al., A monoclonal antibody that targets a NaV1.7 channel voltage sensor for pain and itch relief. *Cell.* Jun. 5, 2014;157(6):1393-1404. doi: 10.1016/j.cell.2014.03.064. Epub May 22, 2014. Retraction in: *Cell.* Jun. 25, 2020;181(7):1695.
- Lee et al., An allosteric self-splicing ribozyme triggered by a bacterial second messenger. *Science.* Aug. 13, 2010;329(5993):845-8. doi: 10.1126/science.1190713.
- Lee et al., Failure to detect DNA-guided genome editing using *Natronobacterium gregoryi* Argonaute. *Nat Biotechnol.* Nov. 28, 2016;35(1):17-18. doi: 10.1038/nbt.3753.
- Lee et al., Group I Intron-Based Therapeutics Through Trans-Splicing Reaction. *Prog Mol Biol Transl Sci.* 2018;159:79-100. doi: 10.1016/bs.pmbts.2018.07.001. Epub Aug. 9, 2018.
- Lee et al., PIK3CA gene is frequently mutated in breast carcinomas and hepatocellular carcinomas. *Oncogene.* Feb. 17, 2005;24(8):1477-80.
- Lee et al., Recognition of liposomes by cells: in vitro binding and endocytosis mediated by specific lipid headgroups and surface charge density. *Biochim Biophys Acta.* Jan. 31, 1992;1103(2):185-97.
- Lee et al., Ribozyme Mediated gRNA Generation for In Vitro and In Vivo CRISPR/Cas9 Mutagenesis. *PLoS One.* Nov. 10, 2016;11(11):e0166020. doi: 10.1371/journal.pone.0166020. eCollection 2016.
- Lee et al., Simultaneous targeting of linked loci in mouse embryos using base editing. *Sci Rep.* Feb. 7, 2019;9(1):1662. doi: 10.1038/s41598-018-33533-5.

(56)

References Cited**OTHER PUBLICATIONS**

- Lee et al., Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette-Guérin. Proc Natl Acad Sci U S A. Apr. 15, 1991;88(8):3111-5. doi: 10.1073/pnas.88.8.3111.
- Lee et al., Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. Elife. May 2, 2017;6:e25312. doi: 10.7554/elife.25312.
- Lee et al., Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome Res. Jan. 20, 2010: 81-89; Published in Advance Dec. 1, 2009, doi:10.1101/gr.099747.109.
- Lee et al., Targeting fidelity of adenine and cytosine base editors in mouse embryos. Nat Commun. Nov. 15, 2018;9(1):4804. doi: 10.1038/s41467-018-07322-7.
- Lee et al., Transcriptional regulation and its misregulation in disease. Cell. Mar. 14, 2013;152(6):1237-51. doi: 10.1016/j.cell.2013.02.014.
- Lei et al., Efficient targeted gene disruption in *Xenopus* embryos using engineered transcription activator-like effector nucleases (TALENs). Proc Natl Acad Sci U S A. Oct. 23, 2012;109(43):17484-9. doi: 10.1073/pnas.1215421109. Epub Oct. 8, 2012.
- Lei et al., Site-specificity of serine integrase demonstrated by the attB sequence preference of ?BT1 integrase. FEBS Lett. Apr. 2018;592(8):1389-1399. doi: 10.1002/1873-3468.13023. Epub Mar. 25, 2018.
- Leipold et al., A de novo gain-of-function mutation in SCN11A causes loss of pain perception. Nat Genet. Nov. 2013;45(11):1399-404. doi: 10.1038/ng.2767. Epub Sep. 15, 2013.
- Lemos et al., CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. Proc Natl Acad Sci U S A. Feb. 27, 2018;115(9):E2040-E2047. doi: 10.1073/pnas.1716855115. Epub Feb. 13, 2018.
- Lenk et al., Pathogenic mechanism of the FIG4 mutation responsible for Charcot-Marie-Tooth disease CMT4J. PLoS Genet. Jun. 2011;7(6):e1002104. doi: 10.1371/journal.pgen.1002104. Epub Jun. 2, 2011.
- Levy et al., Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. Nat Biomed Eng. 2020;4(1):97-110. doi:10.1038/s41551-019-0501-5.
- Levy et al., Inhibition of calcification of bioprosthetic heart valves by local controlled-release diphosphonate. Science. Apr. 12, 1985;228(4696):190-2.
- Levy et al., Membrane-associated guanylate kinase dynamics reveal regional and developmental specificity of synapse stability. J Physiol. Mar. 1, 2017;595(5):1699-1709. doi: 10.1113/JP273147. Epub Jan. 18, 2017.
- Lew et al., Protein splicing in vitro with a semisynthetic two-component minimal intein. J Biol Chem. Jun. 26, 1998;273(26):15887-90. doi: 10.1074/jbc.273.26.15887.
- Lewis et al., A serum-resistant cytofectin for cellular delivery of antisense oligodeoxynucleotides and plasmid DNA. Proc Natl Acad Sci U S A. Apr. 16, 1996;93(8):3176-81.
- Lewis et al., Building the Class 2 CRISPR-Cas Arsenal. Mol Cell. 2017;65(3):377-379.
- Lewis et al., Codon 129 polymorphism of the human prion protein influences the kinetics of amyloid formation. J Gen Virol. Aug. 2006;87(Pt 8):2443-9.
- Lewis et al., Cytosine deamination and the precipitous decline of spontaneous mutation during Earth's history. Proc Natl Acad Sci U S A. Jul. 19, 2016;113(29):8194-9. doi: 10.1073/pnas.1607580113. Epub Jul. 5, 2016.
- Lewis et al., RNA modifications and structures cooperate to guide RNA-protein interactions. Nat Rev Mol Cell Biol. Mar. 2017;18(3):202-210. doi: 10.1038/nrm.2016.163. Epub Feb. 1, 2017.
- Li et al., A Radioactivity-Based Assay for Screening Human m6A-RNA Methyltransferase, METTL3-METTL14 Complex, and Demethylase ALKBH5. J Biomol Screen. Mar. 2016;21(3):290-7. doi: 10.1177/1087057115623264. Epub Dec. 23, 2015.
- Li et al., Base editing with a Cpf1-cytidine deaminase fusion. Nat Biotechnol. Apr. 2018;36(4):324-327. doi: 10.1038/nbt.4102. Epub Mar. 19, 2018.
- Li et al., Current approaches for engineering proteins with diverse biological properties. Adv Exp Med Biol. 2007;620:18-33.
- Li et al., Disruption of splicing-regulatory elements using CRISPR/Cas9 to rescue spinal muscular atrophy in human iPSCs and mice. National Science Review. Jan. 1, 2020:92-101. DOI: 10.1093/nsr/nwz131. Retrieved from the Internet via <https://academic.oup.com/nsr/article-pdf/7/1/92/33321439/nwz131.pdf>. Last accessed Apr. 28, 2021.
- Li et al., Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. Jul. 15, 2009;25(14):1754-60. doi: 10.1093/bioinformatics/btp324. Epub May 18, 2009.
- Li et al., Generation of Targeted Point Mutations in Rice by a Modified CRISPR/Cas9 System. Mol Plant. Mar. 6, 2017;10(3):526-529. doi: 10.1016/j.molp.2016.12.001. Epub Dec. 8, 2016.
- Li et al., Highly efficient and precise base editing in discarded human tripromuclear embryos. Protein Cell. Aug. 19, 2017. doi: 10.1007/s13238-017-0458-7. [Epub ahead of print].
- Li et al., Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. J Biol Chem. Sep. 22, 1995;270(38):22109-12. doi: 10.1074/jbc.270.38.22109.
- Li et al., Loss of post-translational modification sites in disease. Pac Symp Biocomput. 2010:337-47. doi: 10.1142/9789814295291_0036.
- Li et al., Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res. Aug. 2011;39(14):6315-25. doi: 10.1093/nar/gkr188. Epub Mar. 31, 2011.
- Li et al., Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. Nat Biotechnol. Aug. 2013;31(8):688-91. doi: 10.1038/nbt.2654.
- Li et al., Programmable Single and Multiplex Base-Editing in *Bombyx mori* Using RNA-Guided Cytidine Deaminases. G3 (Bethesda). May 4, 2018;8(5):1701-1709. doi: 10.1534/g3.118.200134.
- Li et al., Protein trans-splicing as a means for viral vector-mediated in vivo gene therapy. Hum Gene Ther. Sep. 2008;19(9):958-64. doi: 10.1089/hum.2008.0009.
- Li et al., RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. Aug. 4, 2011;12:323. doi: 10.1186/1471-2105-12-323.
- Li et al., TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res. Jan. 2011;39(1):359-72. doi: 10.1093/nar/gkq704. Epub Aug. 10, 2010.
- Li, Mechanisms and functions of DNA mismatch repair. Cell Res. Jan. 2008;18(1):85-98. doi: 10.1038/cr.2007.115.
- Liang et al., Correction of ?thalassemia mutant by base editor in human embryos. Protein Cell. Nov. 2017;8(11):811-822. doi: 10.1007/s13238-017-0475-6. Epub Sep. 23, 2017.
- Liang et al., Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc Natl Acad Sci U S A. Apr. 28, 1998;95(9):5172-7. doi: 10.1073/pnas.95.9.5172.
- Liang et al., Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. Send to; J Biotechnol. Aug. 20, 2015;208:44-53. doi: 10.1016/j.jbiotec.2015.04.024.
- Liao et al., One-step assembly of large CRISPR arrays enables multi-functional targeting and reveals constraints on array design. bioRxiv. May 2, 2018. doi: 10.1101/312421. 45 pages.
- Lieber et al., Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol. Sep. 2003;4(9):712-20.
- Liefke et al., The oxidative demethylase ALKBH3 marks hyperactive gene promoters in human cancer cells. Genome Med. Jun. 30, 2015;7(1):66. doi: 10.1186/s13073-015-0180-0.
- Lienert et al., Two- and three-input TALE-based AND logic computation in embryonic stem cells. Nucleic Acids Res. Nov. 2013;41(21):9967-75. doi: 10.1093/nar/gkt758. Epub Aug. 27, 2013.
- Lilley, D.M. The Varkud Satellite Ribozyme. RNA. Feb. 2004;10(2):151-8. doi: 10.1261/rna.5217104.

(56)

References Cited**OTHER PUBLICATIONS**

- Lim et al., Crystal structure of the moloney murine leukemia virus RNase H domain. *J Virol.* Sep. 2006;80(17):8379-89. doi: 10.1128/JVI.00750-06.
- Lim et al., Viral vectors for neurotrophic factor delivery: a gene therapy approach for neurodegenerative diseases of the CNS. *Pharmacol Res.* Jan. 2010;61(1):14-26. doi: 10.1016/j.phrs.2009.10.002. Epub Oct. 17, 2009.
- Lin et al., Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife.* Dec. 15, 2014;3:e04766. doi: 10.7554/elife.04766.
- Lin et al., The human REV1 gene codes for a DNA template-dependent dCMP transferase. *Nucleic Acids Res.* Nov. 15, 1999;27(22):4468-75. doi: 10.1093/nar/27.22.4468.
- Link et al., Engineering ligand-responsive gene-control elements: lessons learned from natural riboswitches. *Gene Ther.* Oct. 2009;16(10):1189-201. doi: 10.1038/gt.2009.81. Epub Jul. 9, 2009. Review.
- Liu et al., C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism. *Molecular Cell* Jan. 2017;65(2):310-22.
- Liu et al., Split dnaE genes encoding multiple novel inteins in Trichodesmium erythraeum. *J Biol Chem.* Jul. 18, 2003;278(29):26315-8. doi: 10.1074/jbc.C30022200. Epub May 24, 2003.
- Liu et al., A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol.* Feb. 2014;10(2):93-5. doi: 10.1038/nchembio.1432. Epub Dec. 6, 2013.
- Liu et al., Adding new chemistries to the genetic code. *Annu Rev Biochem.* 2010;79:413-44. doi: 10.1146/annurev.biochem.052308.105824.
- Liu et al., Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat Rev Neurol.* Feb. 2013;9(2):106-18. doi: 10.1038/nrneurol.2012.263. Epub Jan. 8, 2013.
- Liu et al., Balancing AID and DNA repair during somatic hypermutation. *Trends Immunol.* Apr. 2009;30(4):173-81. doi: 10.1016/j.it.2009.01.007.
- Liu et al., Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell.* Aug. 23, 1991;66(4):807-15. doi: 10.1016/0092-8674(91)90124-h.
- Liu et al., CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature.* Feb. 2019;566(7743):218-223. doi: 10.1038/s41586-019-0908-x. Epub Feb. 4, 2019. Author manuscript entitled CRISPR-CasX is an RNA-dominated enzyme active for human genome editing.
- Liu et al., Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. *PLoS One.* Jan. 20, 2014;9(1):e85755. doi: 10.1371/journal.pone.0085755. eCollection 2014.
- Liu et al., Computational approaches for effective CRISPR guide RNA design and evaluation. *Comput Struct Biotechnol J.* Nov. 29, 2019;18:35-44. doi: 10.1016/j.csbj.2019.11.006.
- Liu et al., Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc Natl Acad Sci U S A.* May 27, 1997;94(11):5525-30.
- Liu et al., Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell.* Apr. 5, 2018;173(2):430-442.e17. doi: 10.1016/j.cell.2018.03.016. Epub Mar. 29, 2018.
- Liu et al., Distance determination by GIY-YIG intron endonucleases: discrimination between repression and cleavage functions. *Nucleic Acids Res.* Mar. 31, 2006;34(6):1755-64. Print 2006.
- Liu et al., Editing DNA Methylation in the Mammalian Genome. *Cell.* Sep. 22, 2016;167(1):233-247.e17. doi: 10.1016/j.cell.2016.08.056.
- Liu et al., Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo. *Proc Natl Acad Sci U S A.* Sep. 16, 1997;94(19):10092-7.
- Liu et al., Fast Colorimetric Sensing of Adenosine and Cocaine Based on a General Sensor Design Involving Aptamers and Nanoparticles. *Angew Chem.* Dec. 16, 2006;45(1):90-4. DOI: 10.1002/anie.200502589.
- Liu et al., Fast Colorimetric Sensing of Adenosine and Cocaine Based on a General Sensor Design Involving Aptamers and Nanoparticles. *Angew Chem.* 2006;118(1):96-100.
- Liu et al., Flap endonuclease 1: a central component of DNA metabolism. *Annu Rev Biochem.* 2004;73:589-615. doi:10.1146/annurev.biochem.73.012803.092453.
- Liu et al., Functional Nucleic Acid Sensors. *Chem Rev.* May 2009;109(5):1948-98. doi: 10.1021/cr030183i.
- Liu et al., Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods.* Mar. 2007;4(3):239-44. Epub Feb. 25, 2007.
- Liu et al., Highly efficient RNA-guided base editing in rabbit. *Nat Commun.* Jul. 13, 2018;9(1):2717. doi: 10.1038/s41467-018-05232-2.
- Liu et al., (6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature.* Feb. 26, 2015;518(7540):560-4. doi: 10.1038/nature14234.
- Liu et al., Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA.* Dec. 2013;19(12):1848-56. doi: 10.1261/ma.041178.113. Epub Oct. 18, 2013.
- Liu et al., Reverse transcriptase of foamy virus. Purification of the enzymes and immunological identification. *Arch Virol.* 1977;55(3):187-200. doi: 10.1007/BF01319905.
- Liu et al., Reverse transcriptase-mediated tropism switching in *Bordetella* bacteriophage. *Science.* Mar. 15, 2002;295(5562):2091-4. doi: 10.1126/science.1067467.
- Liu et al., *Saccharomyces cerevisiae* flap endonuclease 1 uses flap equilibration to maintain triplet repeat stability. *Mol Cell Biol.* May 2004;24(9):4049-64. doi: 10.1128/MCB.24.9.4049-4064.2004.
- Liu et al., The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. *Cell.* Aug. 10, 2017;170(4):714-726.e10. doi: 10.1016/j.cell.2017.06.050. Epub Jul. 27, 2017.
- Loessner et al., Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. *Mol Microbiol.* Jan. 2000;35(2):324-40. doi: 10.1046/j.1365-2958.2000.01720.x.
- Lombardo et al., Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol.* Nov. 2007;25(11):1298-306. Epub Oct. 28, 2007.
- Long et al., Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science.* Jan. 22, 2016;351(6271):400-3. doi: 10.1126/science.aad5725. Epub Dec. 31, 2015.
- Lopez-Girona et al., Cereblon is a direct protein target for immunomodulatory and antiproliferative activities of lenalidomide and pomalidomide. *Leukemia.* Nov. 2012;26(11):2326-35. doi: 10.1038/leu.2012.119. Epub May 3, 2012.
- Lorenz et al., ViennaRNA Package 2.0. *Algorithms Mol Biol.* Nov. 24, 2011;6:26. doi: 10.1186/1748-7188-6-26.
- Losey et al., Crystal structure of *Staphylococcus aureus* tRNA adenosine deaminase tadA in complex with RNA. *Nature Struct. Mol. Biol.* Feb. 2006;13(2):153-9.
- Lu et al., Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. *Mol Plant.* Mar. 6, 2017;10(3):523-525. doi: 10.1016/j.molp.2016.11.013. Epub Dec. 6, 2016.
- Luan et al., Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell.* Feb. 26, 1993;72(4):595-605. doi: 10.1016/0092-8674(93)90078-5.
- Luckow et al., High level expression of nonfused foreign genes with *Autographa californica* nuclear polyhedrosis virus expression vectors. *Virology.* May 1989;170(1):31-9. doi: 10.1016/0042-6822(89)90348-6.
- Lukacsovich et al., Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. *Nucleic Acids Res.* Dec. 25, 1994;22(25):5649-57. doi: 10.1093/nar/22.25.5649.
- Lundberg et al., Delivery of short interfering RNA using endosomolytic cell-penetrating peptides. *FASEB J.* Sep. 2007;21(11):2664-71. Epub Apr. 26, 2007.
- Lundquist et al., Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein. Role of specific carbox-

(56)

References Cited**OTHER PUBLICATIONS**

- ylc amino acids in complex formation with *Escherichia coli* uracil-DNA glycosylase. *J Biol Chem.* Aug. 22, 1997;272(34):21408-19.
- Lynch, Evolution of the mutation rate. *Trends Genet.* Aug. 2010;26(8):345-52. doi: 10.1016/j.tig.2010.05.003. Epub Jun. 30, 2010.
- Lyons et al., Efficient Recognition of an Unpaired Lesion by a DNA Repair Glycosylase. *J. Am. Chem. Soc.*, 2009;131(49):17742-3. DOI: 10.1021/ja908378y.
- Lüke et al., Partial purification and characterization of the reverse transcriptase of the simian immunodeficiency virus TYO-7 isolated from an African green monkey. *Biochemistry*. Feb. 20, 1990;29(7):1764-9. doi: 10.1021/bi00459a015.
- Ma et al., Identification of pseudo attP sites for phage phiC31 integrase in bovine genome. *Biochem Biophys Res Commun.* Jul. 7, 2006;345(3):984-8. doi: 10.1016/j.bbrc.2006.04.145. Epub May 3, 2006.
- Ma et al., In vitro protein engineering using synthetic tRNA(Ala) with different anticodons. *Biochemistry*. Aug. 10, 1993;32(31):7939-45.
- Ma et al., PhiC31 integrase induces efficient site-specific recombination in the *Capra hircus* genome. *DNA Cell Biol.* Aug. 2014;33(8):484-91. doi: 10.1089/dna.2013.2124. Epub Apr. 22, 2014.
- Ma et al., Single-Stranded DNA Cleavage by Divergent CRISPR-Cas9 Enzymes. *Mol Cell.* Nov. 5, 2015;60(3):398-407. doi: 10.1016/j.molcel.2015.10.030.
- Ma et al., Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nature Methods*. Oct. 2016;13:1029-35. doi: 10.1038/nmeth.4027.
- Maas et al., Identification and characterization of a human tRNA-specific adenosine deaminase related to the ADAR family of pre-mRNA editing enzymes. *Proc Natl Acad Sci U S A.* Aug. 3, 1999;96(16):8895-900. doi: 10.1073/pnas.96.16.8895.
- Macbeth et al., Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science*. Sep. 2, 2005;309(5740):1534-9. doi: 10.1126/science.1113150.
- Macrae et al., Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr Opin Struct Biol.* Feb. 2007;17(1):138-45. doi: 10.1016/j.sbi.2006.12.002. Epub Dec. 27, 2006.
- Madura et al., Structural basis for ineffective T-cell responses to MHC anchor residue-improved "heteroclitic" peptides. *Eur J Immunol.* Feb. 2015;45(2):584-91. doi: 10.1002/eji.201445114. Epub Dec. 28, 2014.
- Maeder et al., CRISPR RNA-guided activation of endogenous human genes. *Nat Methods*. Oct. 2013;10(10):977-9. doi: 10.1038/nmeth.2598. Epub Jul. 25, 2013.
- Maeder et al., Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell.* Jul. 25, 2008;31(2):294-301. doi: 10.1016/j.molcel.2008.06.016.
- Maeder et al., Robust, synergistic regulation of human gene expression using TALE activators. *Nat Methods*. Mar. 2013;10(3):243-5. doi: 10.1038/nmeth.2366. Epub Feb. 10, 2013.
- Magin et al., Corf, the Rev/Rex homologue of HTDV/HERV-K, encodes an arginine-rich nuclear localization signal that exerts a trans-dominant phenotype when mutated. *Virology*. Aug. 15, 2000;274(1):11-6. doi: 10.1006/viro.2000.0438.
- Mahfouz et al., De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc Natl Acad Sci U S A.* Feb. 8, 2011;108(6):2623-8. doi: 10.1073/pnas.1019533108. Epub Jan. 24, 2011.
- Maizels et al., Initiation of homologous recombination at DNA nicks. *Nucleic Acids Res.* Aug. 21, 2018;46(14):6962-6973. doi: 10.1093/nar/gky588.
- Maji et al., A High-Throughput Platform to Identify Small-Molecule Inhibitors of CRISPR-Cas9. *Cell*. May 2, 2019;177(4):1067-1079.e19. doi: 10.1016/j.cell.2019.04.009.
- Makarova et al., An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol.* Nov. 2015;13(11):722-36. doi: 10.1038/nrmicro3569. Epub Sep. 28, 2015.
- Makarova et al., Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? *CRISPR J.* Oct. 2018;1(5):325-336. doi: 10.1089/crispr.2018.0033.
- Makarova et al., Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol.* Jun. 2011;9(6):467-77. doi: 10.1038/nrmicro2577. Epub May 9, 2011.
- Makarova et al., Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biology Direct* 2009;4:29. doi: 10.1186/1745-6150-4-29.
- Makeyev et al., Evolutionary potential of an RNA virus. *J Virol.* Feb. 2004;78(4):2114-20.
- Malashkevich et al., Crystal structure of tRNA adenosine deaminase TadA from *Escherichia coli*. Deposited: Mar. 10, 2005 Released: Feb. 21, 2006 doi:10.2210/pdb1z3a/pdb (2006).
- Mali et al., Cas9 as a versatile tool for engineeringbiology. *Nat Methods*. Oct. 2013;10(10):957-63. doi: 10.1038/nmeth.2649.
- Mali et al., CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol.* Sep. 2013;31(9):833-8. doi: 10.1038/nbt.2675. Epub Aug. 1, 2013.
- Mali et al., RNA-guided human genome engineering via Cas9. *Science*. Feb. 15, 2013;339(6121):823-6. doi: 10.1126/science.1232033. Epub Jan. 3, 2013.
- Malito et al., Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. *Proc Natl Acad Sci U S A.* Apr. 3, 2012;109(14):5229-34. doi: 10.1073/pnas.1201964109. Epub Mar. 19, 2012.
- Mandal et al., Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell*. Nov. 6, 2014;15(5):643-52. doi: 10.1016/j.stem.2014.10.004. Epub Nov. 6, 2014.
- Mandal et al., Riboswitches Control Fundamental Biochemical Pathways in *Bacillus Subtilis* and Other Bacteria. *Cell*. May 30, 2003;113(5):577-86. doi: 10.1016/s0092-8674(03)00391-x.
- Mani et al., Design, engineering, and characterization of zinc finger nucleases. *Biochem Biophys Res Commun.* Sep. 23, 2005;335(2):447-57.
- Marceau, Functions of single-strand DNA-binding proteins in DNA replication, recombination, and repair. *Methods Mol Biol.* 2012;922:1-21. doi: 10.1007/978-1-62703-032-8_1.
- Maresca et al., Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Res.* Mar. 2013;23(3):539-46. doi: 10.1101/gr.145441.112. Epub Nov. 14, 2012.
- Marioni et al., DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol.* Jan. 30, 2015;16:25. doi: 10.1186/s13059-015-0584-6.
- Marraffini et al., CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*. Dec. 19, 2008;322(5909):1843-5. doi: 10.1126/science.1165771.
- Martinez et al., Hypermutation of RNA using human immunodeficiency virus type 1 reverse transcriptase and biased dNTP concentrations. *Proc Natl Acad Sci U S A.* Dec. 6, 1994;91(25):11787-91. doi: 10.1073/pnas.91.25.11787.
- Martsolf et al., Complete trisomy 17p a relatively new syndrome. *Ann Genet.* 1988;31(3):172-4.
- Martz, L., Nav-i-gating antibodies for pain. *Science-Business eXchange*. Jun. 12, 2014;7(662):1-2. doi: 10.1038/scibx.2014.662.
- Maruyama et al., Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol.* May 2015;33(5):538-42. doi: 10.1038/nbt.3190. Epub Mar. 23, 2015.
- Marzec et al., Prime Editing: A New Way for Genome Editing. *Trends Cell Biol.* Apr. 2020;30(4):257-259. doi: 10.1016/j.tcb.2020.01.004. Epub Jan. 27, 2020.
- Mascola et al., HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunol Rev.* Jul. 2013;254(1):225-44. doi: 10.1111/imr.12075.

(56)

References Cited**OTHER PUBLICATIONS**

- Mathys et al., Characterization of a self-splicing mini-intein and its conversion into autocatalytic—and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene*. Apr. 2, 1999;231(1-2):1-13. doi: 10.1016/s0378-1119(99)00103-1.
- Matsuura et al., A gene essential for the site-specific excision of actinophage r4 prophage genome from the chromosome of a lysogen. *J Gen Appl Microbiol*. 1995;41(1):53-61.
- Matthews, Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity. *Nat Struct Mol Biol*. May 2016;23(5):426-33. doi: 10.1038/nsmb.3203. Epub Apr. 11, 2016.
- May et al., Emergent lineages of mumps virus suggest the need for a polyvalent vaccine. *Int J Infect Dis*. Jan. 2018;66:1-4. doi: 10.1016/j.ijid.2017.09.024. Epub Oct. 4, 2017.
- McCarroll et al., Copy-number variation and association studies of human disease. *Nat Genet*. Jul. 2007;39(7 Suppl):S37-42. doi: 10.1038/ng2080.
- McDonald et al., Characterization of mutations at the mouse phenylalanine hydroxylase locus. *Genomics*. Feb. 1, 1997;39(3):402-5. doi: 10.1006/geno.1996.4508.
- McInerney et al., Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase. *Mol Biol Int*. 2014;2014:287430. doi: 10.1155/2014/287430. Epub Aug. 17, 2014.
- McKenna et al., Recording development with single cell dynamic lineage tracing. *Development*. Jun. 27, 2019;146(12):dev169730. doi: 10.1242/dev.169730.
- McKenna et al., Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science*. Jul. 29, 2016;353(6298):aaf7907. doi: 10.1126/science.aaf7907. Epub May 26, 2016.
- McNaughton et al., Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. *Proc Natl Acad Sci U S A*. Apr. 14, 2009;106(15):6111-6. doi: 10.1073/pnas.0807883106. Epub Mar. 23, 2009.
- McVey et al., MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet*. Nov. 2008;24(11):529-38. doi: 10.1016/j.tig.2008.08.007. Epub Sep. 21, 2008.
- Mead et al., A novel protective prion protein variant that colocalizes with kuru exposure. *Engl J Med*. Nov. 19, 2009;361(21):2056-65. doi: 10.1056/NEJMoa0809716.
- Mei et al., Recent Progress in CRISPR/Cas9 Technology. *J Genet Genomics*. Feb. 20, 2016;43(2):63-75. doi: 10.1016/j.jgg.2016.01.001. Epub Jan. 18, 2016.
- Meinke et al., Cre Recombinase and Other Tyrosine Recombinases. *Chem Rev*. Oct. 26, 2016;116(20):12785-12820. doi: 10.1021/acs.chemrev.6b00077. Epub May 10, 2016.
- Meng et al., Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol*. Jun. 2008;26(6):695-701. doi: 10.1038/nbt1398. Epub May 25, 2008.
- Menéndez-Arias, Mutation rates and intrinsic fidelity of retroviral reverse transcriptases. *Viruses*. Dec. 2009;1(3):1137-65. doi: 10.3390/v1031137. Epub Dec. 4, 2009.
- Mercer et al., Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Res*. Nov. 2012;40(21):11163-72. doi: 10.1093/nar/gks875. Epub Sep. 26, 2012.
- Mertens et al., Site-specific recombination in bacteriophage Mu: characterization of binding sites for the DNA invertase Gin. *EMBO J*. Apr. 1988;7(4):1219-27.
- Meyer et al., Breathing life into polycations: functionalization with pH-responsive endosomolytic peptides and polyethylene glycol enables siRNA delivery. *J Am Chem Soc*. Mar. 19, 2008;130(11):3272-3. doi: 10.1021/ja710344v. Epub Feb. 21, 2008.
- Meyer et al., Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. Jun. 22, 2012;149(7):1635-46. doi: 10.1016/j.cell.2012.05.003. Epub May 17, 2012.
- Meyer et al., Confirmation of a second natural preQ1 aptamer class in Streptococcaceae bacteria. *RNA*. Apr. 2008;14(4):685-95. doi: 10.1261/rna.937308. Epub Feb. 27, 2008.
- Meyer et al., Library generation by gene shuffling. *Curr Protoc Mol Biol*. Jan. 6, 2014;105: Unit 15.12 . . . doi: 10.1002/0471142727.mbl1512s105.
- Meyer et al., Ribosome biogenesis factor Tsrf3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. *Nucleic Acids Res*. May 19, 2016;44(9):4304-16. doi: 10.1093/nar/gkw244. Epub Apr. 15, 2016.
- Meyer et al., The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat Rev Mol Cell Biol*. May 2014;15(5):313-26. doi: 10.1038/nrm3785. Epub Apr. 9, 2014.
- Michel et al., Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. *Nature*. Aug. 15-21, 1985;316(6029):641-3. doi: 10.1038/316641a0.
- Midoux et al., Chemical vectors for gene delivery: a current review on polymers, peptides and lipids containing histidine or imidazole as nucleic acids carriers. *Br J Pharmacol*. May 2009;157(2):166-78. doi: 10.1111/j.1476-5381.2009.00288.x.
- Mihai et al., PTEN inhibition improves wound healing in lung epithelia through changes in cellular mechanics that enhance migration. *Am J Physiol Lung Cell Mol Physiol*. 2012;302(3):L287-L299.
- Mijakovic et al., Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res*. Mar. 20, 2006;34(5):1588-96. doi: 10.1093/nar/gkj514.
- Miller et al., A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*. Feb. 2011;29(2):143-8. doi: 10.1038/nbt.1755. Epub Dec. 22, 2010.
- Miller et al., An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*. Jul. 2007;25(7):778-85. Epub Jul. 1, 2007.
- Miller et al., Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. *J Virol*. May 1991;65(5):2220-4. doi: 10.1128/JVI.65.5.2220-2224.1991.
- Miller et al., Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat Biotechnol*. Apr. 2020;38(4):471-481. doi: 10.1038/s41587-020-0412-8. Epub Feb. 10, 2020.
- Miller, Human gene therapy comes of age. *Nature*. Jun. 11, 1992;357(6378):455-60. doi: 10.1038/357455a0.
- Mills et al., Protein splicing in trans by purified—and C-terminal fragments of the *Mycobacterium tuberculosis* RecA intein. *Proc Natl Acad Sci U S A*. Mar. 31, 1998;95(7):3543-8. doi: 10.1073/pnas.95.7.3543.
- Minoche et al., Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol*. Nov. 8, 2011;12(11):R112. doi: 10.1186/GB-2011-12-11-r112.
- Minoretti et al., A W148R mutation in the human FOXD4 gene segregating with dilated cardiomyopathy, obsessive-compulsive disorder, and suicidality. *Int J Mol Med*. Mar. 2007;19(3):369-72.
- Mir et al., Two Active Site Divalent Ions in the Crystal Structure of the Hammerhead Ribozyme Bound to a Transition State Analogue. *Biochemistry* . . . Feb. 2, 2016;55(4):633-6. doi: 10.1021/acs.biochem.5b01139. Epub Jan. 19, 2016.
- Mir et al., Type II-C CRISPR-Cas9 Biology, Mechanism, and Application. *ACS Chem Biol*. Feb. 16, 2018;13(2):357-365. doi: 10.1021/acscchembio.7b00855. Epub Dec. 20, 2017.
- Mishina et al., Conditional gene targeting on the pure C57BL/6 genetic background. *Neurosci Res*. Jun. 2007;58(2):105-12. doi: 10.1016/j.neures.2007.01.004. Epub Jan. 18, 2007.
- Mitani et al., Delivering therapeutic genes—matching approach and application. *Trends Biotechnol*. May 1993;11(5):162-6. doi: 10.1016/0167-7799(93)90108-L.
- Mitton-Fry et al., Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science*. Nov. 26, 2010;330(6008):1244-7. doi: 10.1126/science.1195858.
- Miyaoka et al., Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci Rep*. Mar. 31, 2016;6:23549. doi: 10.1038/srep23549.
- Moede et al., Identification of a nuclear localization signal, RRMKWKK, in the homeodomain transcription factor PDX-1. *FEBS Lett*. Nov. 19, 1999;461(3):229-34. doi: 10.1016/s0014-5793(99)01446-5.

(56)

References Cited**OTHER PUBLICATIONS**

- Mohr et al., A Reverse Transcriptase-Cas1 Fusion Protein Contains a Cas6 Domain Required for Both CRISPR RNA Biogenesis and RNA Spacer Acquisition. *Mol Cell.* Nov. 15, 2018;72(4):700-714. e8. doi: 10.1016/j.molcel.2018.09.013. Epub Oct. 18, 2018. Including Supplemental Information.
- Mohr et al., Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA.* Jul. 2013;19(7):958-70. doi: 10.1261/rna.039743.113. Epub May 22, 2013.
- Mojica et al., Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol.* Feb. 2005;60(2):174-82.
- Mol et al., Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell.* Mar. 24, 1995;80(6):869-78. doi: 10.1016/0092-8674(95)90290-2.
- Mol et al., Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. *Cell.* Sep. 8, 1995;82(5):701-8.
- Molla et al., CRISPR/Cas-Mediated Base Editing: Technical Considerations and Practical Applications. *Trends Biotechnol.* Oct. 2019;37(10):1121-1142. doi: 10.1016/j.tibtech.2019.03.008. Epub Apr. 14, 2019.
- Monahan et al., Site-specific incorporation of unnatural amino acids into receptors expressed in Mammalian cells. *Chem Biol.* Jun. 2003;10(6):573-80.
- Monot et al., The specificity and flexibility of 11 reverse transcription priming at imperfect T-tracts. *PLoS Genet.* May 2013;9(5):e1003499. doi: 10.1371/journal.pgen.1003499. Epub May 9, 2013.
- Montange et al., Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature.* Jun. 29, 2006;441(7097):1172-5.
- Moore et al., Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). *PloS One.* 2012;7(5):e37877. doi: 10.1371/journal.pone.0037877. Epub May 24, 2012.
- Mootz et al., Conditional protein splicing: a new tool to control protein structure and function in vitro and in vivo. *J Am Chem Soc.* Sep. 3, 2003;125(35):10561-9.
- Mootz et al., Protein splicing triggered by a small molecule. *J Am Chem Soc.* Aug. 7, 2002;124(31):9044-5 and Supporting Information. doi: 10.1021/ja0267690. 4 pages.
- Mootz et al., Protein splicing triggered by a small molecule. *J Am Chem Soc.* Aug. 7, 2002;124(31):9044-5.
- Morbitzer et al., Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res.* Jul. 2011;39(13):5790-9. doi: 10.1093/nar/gkr151. Epub Mar. 18, 2011.
- Moreno-Mateos et al., CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods.* Oct. 2015;12(10):982-8. doi: 10.1038/nmeth.3543. Epub Aug. 31, 2015.
- Morita et al., The site-specific recombination system of actinophage TG1. *FEMS Microbiol Lett.* Aug. 2009;297(2):234-40. doi: 10.1111/j.1574-6968.2009.01683.x.
- Morris et al., A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat Biotechnol.* Dec. 2001;19(12):1173-6.
- Moscou et al., A simple cipher governs DNA recognition by TAL effectors. *Science.* Dec. 11, 2009;326(5959):1501. doi: 10.1126/science.1178817.
- Mougiakos et al., Characterizing a thermostable Cas9 for bacterial genome editing and silencing. *Nat Commun.* Nov. 21, 2017;8(1):1647. doi: 10.1038/s41467-017-01591-4.
- Muir et al., Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A.* Jun. 9, 1998;95(12):6705-10. doi: 10.1073/pnas.95.12.6705.
- Muller et al., Nucleotide exchange and excision technology (NExT) DNA shuffling: a robust method for DNA fragmentation and directed evolution. *Nucleic Acids Res.* Aug. 1, 2005;33(13):e117. doi: 10.1093/nar/gni116. PMID: 16061932; PMCID: PMC1182171.
- Muller, U.F., Design and Experimental Evolution of trans-Splicing Group I Intron Ribozymes. *Molecules.* Jan. 2, 2017;22(1):75. doi: 10.3390/molecules22010075.
- Mullins et al., Transgenesis in nonmurine species. *Hypertension.*

Mumtsidu et al., Structural features of the single-stranded DNA-binding protein of Epstein-Barr virus. *J Struct Biol.* Feb. 2008;161(2):172-87. doi: 10.1016/j.jsb.2007.10.014. Epub Nov. 1, 2007.

Murphy, Phage recombinases and their applications. *Adv Virus Res.* 2012;83:367-414. doi: 10.1016/B978-0-12-394438-2.00008-6. Review.

Mussolini et al., A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* Nov. 2011;39(21):9283-93. doi: 10.1093/nar/gkr597. Epub Aug. 3, 2011.

Mussolini et al., TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol.* Oct. 2012;23(5):644-50. doi: 10.1016/j.copbio.2012.01.013. Epub Feb. 17, 2012.

Muzyczka et al., Adeno-associated virus (AAV) vectors: will they work? *J Clin Invest.* Oct. 1994;94(4):1351. doi: 10.1172/JCI117468.

Myerowitz et al., The major defect in Ashkenazi Jews with Tay-Sachs disease is an insertion in the gene for the alpha-chain of beta-hexosaminidase. *J Biol Chem.* Dec. 15, 1988;263(35):18587-9.

Myers et al., Insulin signal transduction and the IRS proteins. *Annu Rev Pharmacol Toxicol.* 1996;36:615-58. doi: 10.1146/annurev.pa.36.040196.003151.

Nabel et al., Direct gene transfer for immunotherapy and immunization. *Trends Biotechnol.* May 1993;11(5):211-5. doi: 10.1016/0167-7799(93)90117-R.

Nahar et al., A G-quadruplex motif at the 3' end of sgRNAs improves CRISPR-Cas9 based genome editing efficiency. *Chem Commun (Camb).* Mar. 7, 2018;54(19):2377-2380. doi: 10.1039/c7cc08893k. Epub Feb. 16, 2018.

Nahvi et al., Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res.* Jan. 2, 2004;32(1):143-50.

Nakade et al., Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat Commun.* Nov. 20, 2014;5:5560. doi: 10.1038/ncomms6560.

Nakamura et al., Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* Jan. 1, 2000;28(1):292. doi: 10.1093/nar/28.1.292.

Naorem et al., DGR mutagenic transposition occurs via hypermutagenic reverse transcription primed by nicked template RNA. *Proc Natl Acad Sci U S A.* Nov. 21, 2017;114(47):E10187-E10195. doi: 10.1073/pnas.1715952114. Epub Nov. 6, 2017.

Narayanan et al., Clamping down on weak terminal base pairs: oligonucleotides with molecular caps as fidelity-enhancing elements at the 5'- and 3'-terminal residues. *Nucleic Acids Res.* May 20, 2004;32(9):2901-11. Print 2004.

Navaratnam et al., An overview of cytidine deaminases. *Int J Hematol.* Apr. 2006;83(3):195-200.

NCBI Reference Sequence: NM_002427.3. Wu et al., May 3, 2014. 5 pages.

Neel et al., Riboswitches: Classification, function and in silico approach. *International Journal of Pharma Sciences and Research.* 2010;1(9):409-420.

Nelson et al., Filamentous phage DNA cloning vectors: a noninfective mutant with a nonpolar deletion in gene III. *Virology.* 1981;108(2): 338-50.

Nelson et al., The unstable repeats—three evolving faces of neurological disease. *Neuron.* Mar. 6, 2013;77(5):825-43. doi: 10.1016/j.neuron.2013.02.022.

Nern et al., Multiple new site-specific recombinases for use in manipulating animal genomes. *Proc Natl Acad Sci U S A.* Aug. 23, 2011;108(34):14198-203. doi: 10.1073/pnas.1111704108. Epub Aug. 9, 2011.

(56)

References Cited**OTHER PUBLICATIONS**

- Nguyen et al., Evolutionary drivers of thermoadaptation in enzyme catalysis. *Science*. Jan. 20, 2017;355(6322):289-294. doi: 10.1126/science.aah3717. Epub Dec. 22, 2016.
- Nguyen et al., IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. Jan. 2015;32(1):268-74. doi: 10.1093/molbev/msu300. Epub Nov. 3, 2014.
- Ni et al., A PCSK9-binding antibody that structurally mimics the EGF(A) domain of LDL-receptor reduces LDL cholesterol in vivo. *J Lipid Res*. 2011;52:76-86.
- Ni et al., Nucleic acid aptamers: clinical applications and promising new horizons. *Curr Med Chem*. 2011;18(27):4206-14. Review.
- Nishida et al., Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science*. Sep. 16, 2016;353(6305):1248. pii: aaf8729. doi: 10.1126/science.aaf8729. Epub Aug. 4, 2016.
- Nishikura, Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem*. 2010;79:321-349. doi:10.1146/annurev-biochem-060208-105251.
- Nishimasu et al., Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. Feb. 27, 2014;156(5):935-49. doi: 10.1016/j.cell.2014.02.001. Epub Feb. 13, 2014.
- Nishimasu et al., Crystal Structure of *Staphylococcus aureus* Cas9. *Cell*. Aug. 27, 2015;162(5):1113-26. doi: 10.1016/j.cell.2015.08.007.
- Nishimasu et al., Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science*. Sep. 21, 2018;361(6408):1259-1262. doi: 10.1126/science.aas9129. Epub Aug. 30, 2018.
- Noack et al., Epitranscriptomics: A New Regulatory Mechanism of Brain Development and Function. *Front Neurosci*. Feb. 20, 2018;12:85. doi: 10.3389/fnins.2018.00085. 9 pages.
- Nomura et al., Controlling Mammalian Gene Expression by Allosteric Hepatitis Delta Virus Ribozymes. *ACS Synth Biol*. Dec. 20, 2013;2(12):684-9. doi: 10.1021/sb400037a. Epub May 22, 2013.
- Nomura et al., Synthetic mammalian riboswitches based on guanine aptazyme. *Chem Commun (Camb)*. Jul. 21, 2012;48(57):7215-7. doi: 10.1039/c2cc33140c. Epub Jun. 13, 2012.
- Noris et al., A phenylalanine-55 to serine amino-acid substitution in the human glycoprotein IX leucine-rich repeat is associated with Bernard-Soulier syndrome. *Br J Haematol*. May 1997;97(2):312-20.
- Nottingham et al., RNA-seq of human reference RNA samples using a thermostable group II intron reverse transcriptase. *RNA*. Apr. 2016;22(4):597-613. doi: 10.1261/rna.055558.115. Epub Jan. 29, 2016.
- Nowak et al., Characterization of single-stranded DNA-binding proteins from the psychrophilic bacteria *Desulfotalea psychrophila*, *Flavobacterium psychrophilum*, *Psychrobacter arcticus*, *Psychrobacter cryohalolentis*, *Psychromonas ingrahamii*, *Psychroflexus torquis*, and *Photobacterium profundum*. *BMC Microbiol*. Apr. 14, 2014;14:91. doi: 10.1186/1471-2180-14-91.
- Nowak et al., Guide RNA Engineering for Versatile Cas9 Functionality. *Nucleic Acids Res*. Nov. 16, 2016;44(20):9555-9564. doi: 10.1093/nar/gkw908. Epub Oct. 12, 2016.
- Nowak et al., Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid. *Nucleic Acids Res*. Apr. 1, 2013;41(6):3874-87. doi: 10.1093/nar/gkt053. Epub Feb. 4, 2013.
- Numrych et al., A comparison of the effects of single-base and triple-base changes in the integrase arm-type binding sites on the site-specific recombination of bacteriophage lambda. *Nucleic Acids Res*. Jul. 11, 1990;18(13):3953-9. doi: 10.1093/nar/18.13.3953.
- Nyerges et al., A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci U S A*. Mar. 1, 2016;113(9):2502-7. doi: 10.1073/pnas.1520040113. Epub Feb. 16, 2016.
- O'Connell et al., Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*. Dec. 11, 2014;516(7530):263-6. doi: 10.1038/nature13769. Epub Sep. 28, 2014.
- O'Maille et al., Structure-based combinatorial protein engineering (Scope). *J Mol Biol*. Aug. 23, 2002;321(4):677-91.
- Oakes et al., CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification. *Cell*. Jan. 10, 2019;176(1-2):254-267.e16. doi: 10.1016/j.cell.2018.11.052.
- Oakes et al., Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat Biotechnol*. Jun. 2016;34(6):646-51. doi: 10.1038/nbt.3528. Epub May 2, 2016.
- Oakes et al., Protein engineering of Cas9 for enhanced function. *Methods Enzymol*. 2014;546:491-511.
- Odsbu et al., Specific-terminal interactions of the *Escherichia coli* SeqA protein are required to form multimers that restrain negative supercoils and form foci. *Genes Cells*. Nov. 2005;10(11):1039-49.
- Oeemig et al., Solution structure of DnaE intein from *Nostoc punctiforme*: structural basis for the design of a new split intein suitable for site-specific chemical modification. *FEBS Lett*. May 6, 2009;583(9):1451-6.
- Offord, Advances in Genome Editing. *The Scientist*. Apr. 20, 2016. <http://www.the-scientist.com/?articles.view/articleNo/45903/title/Advances-in-Genome-Editing/>.
- Oh et al., Positional cloning of a gene for Hermansky-Pudlak syndrome, a disorder of cytoplasmic organelles. *Nat Genet*. Nov. 1996;14(3):300-6. doi: 10.1038/ng1196-300.
- Ohe et al., Purification and properties of xanthine dehydrogenase from *Streptomyces cyanogenus*. *J Biochem*. Jul. 1979;86(1):45-53.
- Olivares et al., Site-specific genomic integration produces therapeutic Factor IX levels in mice. *Nat Biotechnol*. Nov. 2002;20(11):1124-8. doi: 10.1038/nbt753. Epub Oct. 15, 2002.
- Olorunniyi et al., Purification and In Vitro Characterization of Zinc Finger Recombinases. *Methods Mol Biol*. 2017;1642:229-245. doi: 10.1007/978-1-4939-7169-5_15.
- Olorunniyi et al., Site-specific recombinases: molecular machines for the Genetic Revolution. *Biochem J*. Mar. 15, 2016;473(6):673-84. doi: 10.1042/BJ20151112.
- Olorunniyi et al., Synapsis and catalysis by activated Tn3 resolvase mutants. *Nucleic Acids Res*. Dec. 2008;36(22):7181-91. doi: 10.1093/nar/gkn885. Epub Nov. 10, 2008.
- Orlando et al., Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res*. Aug. 2010;38(15):e152. doi: 10.1093/nar/gkq512. Epub Jun. 8, 2010.
- Orthwein et al., A mechanism for the suppression of homologous recombination in G1 cells. *Nature*. Dec. 17, 2015;528(7582):422-6. doi: 10.1038/nature16142. Epub Dec. 9, 2015.
- Ortiz-Urda et al., Stable nonviral genetic correction of inherited human skin disease. *Nat Med*. Oct. 2002;8(10):1166-70. doi: 10.1038/nm766. Epub Sep. 16, 2002. Erratum in: *Nat Med*. Feb. 2003;9(2):237.
- Osborn et al., Base Editor Correction of COL7A1 in Recessive Dystrophic Epidermolysis Bullosa Patient-Derived Fibroblasts and iPSCs. *J Invest Dermatol*. Feb. 2020;140(2):338-347.e5. doi: 10.1016/j.jid.2019.07.701. Epub Aug. 19, 2019.
- Osborn et al., TALEN-based gene correction for epidermolysis bullosa. *Mol Ther*. Jun. 2013;21(6):1151-9. doi: 10.1038/mt.2013.56. Epub Apr. 2, 2013.
- Ostermeier et al., A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat Biotechnol*. Dec. 1999;17(12):1205-9.
- Ostertag et al., Biology of mammalian L1 retrotransposons. *Annu Rev Genet*. 2001;35:501-38. doi: 10.1146/annurev.genet.35.102401.091032.
- Otomo et al., Improved segmental isotope labeling of proteins and application to a larger protein. *J Biomol NMR*. Jun. 1999;14(2):105-14. doi: 10.1023/a:1008308128050.
- Otomo et al., NMR observation of selected segments in a larger protein: central-segment isotope labeling through intein-mediated ligation. *Biochemistry*. Dec. 7, 1999;38(49):16040-4. doi: 10.1021/bi991902j.
- Otto et al., The probability of fixation in populations of changing size. *Genetics*. Jun. 1997;146(2):723-33.
- Packer et al., Methods for the directed evolution of proteins. *Nat Rev Genet*. Jul. 2015;16(7):379-94. doi: 10.1038/nrg3927. Epub Jun. 9, 2015.

(56)

References Cited**OTHER PUBLICATIONS**

- Packer et al., Phage-assisted continuous evolution of proteases with altered substrate specificity. *Nat Commun.* Oct. 16, 2017;8(1):956. doi: 10.1038/s41467-017-01055-9.
- Paige et al., RNA mimics of green fluorescent protein. *Science.* Jul. 29, 2011;333(6042):642-6. doi:10.1126/science.1207339.
- Paiva et al., Targeted protein degradation: elements of PROTAC design. *Curr Opin Chem Biol.* Jun. 2019;50:111-119. doi: 10.1016/j.cbpa.2019.02.022. Epub Apr. 17, 2019.
- Pan et al., Biological and biomedical applications of engineered nucleases. *Mol Biotechnol.* Sep. 2013;55(1):54-62. doi: 10.1007/s12033-012-9613-9.
- Paquet et al., Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature.* May 5, 2016;533(7601):125-9. doi: 10.1038/nature17664. Epub Apr. 27, 2016.
- Park et al., Digenome-seq web tool for profiling CRISPR specificity. *Nat Methods.* May 30, 2017;14(6):548-549. doi: 10.1038/nmeth.4262.
- Park et al., Highly efficient editing of the β -globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. *Nucleic Acids Res.* Sep. 5, 2019;47(15):7955-7972. doi: 10.1093/nar/gkz475.
- Park et al., Sendai virus, an RNA virus with no risk of genomic integration, delivers CRISPR/Cas9 for efficient gene editing. *Mol Ther Methods Clin Dev.* Aug. 24, 2016;3:16057. doi: 10.1038/mtm.2016.57.
- Parker et al., Admixture mapping identifies a quantitative trait locus associated with FEV1/FVC in the COPDGene Study. *Genet Epidemiol.* Nov. 2014;38(7):652-9. doi: 10.1002/gepi.21847. Epub Aug. 11, 2014.
- Patel et al., Flap endonucleases pass 5'-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5'-ends. *Nucleic Acids Res.* May 2012;40(10):4507-19. doi: 10.1093/nar/gks051. Epub Feb. 8, 2012.
- Pattanayak et al., Determining the specificities of TALENs, Cas9, and other genome-editing enzymes. *Methods Enzymol.* 2014;546:47-78. doi: 10.1016/B978-0-12-801185-0.00003-9.
- Pattanayak et al., High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol.* Sep. 2013;31(9):839-43. doi: 10.1038/nbt.2673. Epub Aug. 11, 2013.
- Pattanayak et al., Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods.* Aug. 7, 2011;8(9):765-70. doi: 10.1038/nmeth.1670.
- Pavletich et al., Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science.* May 10, 1991;252(5007):809-17.
- Payson et al., Protein phosphorylation in signaling—50 years and counting. *Trends Biochem Sci.* Jun. 2005;30(6):286-90. doi: 10.1016/j.tibs.2005.04.013.
- Pearl, Structure and function in the uracil-DNA glycosylase superfamily. *Mutat Res.* Aug. 30, 2000;460(3-4):165-81.
- Peck et al., Directed evolution of a small-molecule-triggered intein with improved splicing properties in mammalian cells. *Chem Biol.* May 27, 2011;18(5):619-30. doi: 10.1016/j.chembiol.2011.02.014.
- Pellenz et al., New human chromosomal safe harbor sites for genome engineering with CRISPR/Cas9, TAL effector and homing endonucleases. Aug. 20, 2018. bioRxiv doi: <https://doi.org/10.1101/396390>.
- Pelletier, CRISPR-Cas systems for the study of the immune function. Nov. 15, 2016. <https://doi.org/10.1002/9780470015902.a0026896>.
- Pennisi et al., The CRISPR craze. *Science.* Aug. 23, 2013;341(6148):833-6. doi: 10.1126/science.341.6148.833.
- Pennisi et al., The tale of the TALEs. *Science.* Dec. 14, 2012;338(6113):1408-11. doi: 10.1126/science.338.6113.1408.
- Perach et al., Catalytic features of the recombinant reverse transcriptase of bovine leukemia virus expressed in bacteria. *Virology.* Jun. 20, 1999;259(1):176-89. doi: 10.1006/viro.1999.9761.
- Perez et al., Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol.* Jul. 2008;26(7):808-16. doi: 10.1038/nbt1410. Epub Jun. 29, 2008.
- Perez-Pinera et al., Advances in targeted genome editing. *Curr Opin Chem Biol.* Aug. 2012;16(3-4):268-77. doi: 10.1016/j.cbpa.2012.06.007. Epub Jul. 20, 2012.
- Perez-Pinera et al., RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods.* Oct. 2013;10(10):973-6. doi: 10.1038/nmeth.2600. Epub Jul. 25, 2013.
- Perler et al., Protein splicing and autoproteolysis mechanisms. *Curr Opin Chem Biol.* Oct. 1997;1(3):292-9. doi: 10.1016/s1367-5931(97)80065-8.
- Perler et al., Protein splicing elements: inteins and exteins—a definition of terms and recommended nomenclature. *Nucleic Acids Res.* Apr. 11, 1994;22(7):1125-7. doi: 10.1093/nar/22.7.1125.
- Perler, InBase, the New England Biolabs Intein Database. *Nucleic Acids Res.* Jan. 1, 1999;27(1):346-7. doi: 10.1093/nar/27.1.346.
- Perler, Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. *Cell.* Jan. 9, 1998;92(1):1-4. doi: 10.1016/s0092-8674(00)80892-2.
- Perreault et al., Mixed deoxyribo- and ribo-oligonucleotides with catalytic activity. *Nature.* Apr. 5, 1990;344(6266):565-7. doi: 10.1038/344565a0.
- Petek et al., Frequent endonuclease cleavage at off-target locations in vivo. *Mol Ther.* May 2010;18(5):983-6. doi: 10.1038/mt.2010.35. Epub Mar. 9, 2010.
- Petersen-Mahrt et al., AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature.* Jul. 4, 2002;418(6893):99-103.
- Petolino et al., Editing Plant Genomes: a new era of crop improvement. *Plant Biotechnol J.* Feb. 2016;14(2):435-6. doi: 10.1111/pbi.12542.
- Peyrottes et al., Oligodeoxynucleoside phosphoramidates (P-NH2): synthesis and thermal stability of duplexes with DNA and RNA targets. *Nucleic Acids Res.* May 15, 1996;24(10):1841-8.
- Pfeiffer et al., Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. *Mutagenesis.* Jul. 2000;15(4):289-302. doi: 10.1093/mutage/15.4.289.
- Phillips, The challenge of gene therapy and DNA delivery. *J Pharm Pharmacol.* Sep. 2001;53(9):1169-74.
- Pickart et al., Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta.* Nov. 29, 2004;1695(1-3):55-72. doi: 10.1016/j.bbamer.2004.09.019.
- Pieken et al., Kinetic characterization of ribonuclease-resistant 2'-modified hammerhead ribozymes. *Science.* Jul. 19, 1991;253(5017):314-7. doi: 10.1126/science.1857967.
- Pinkert et al., An albumin enhancer located 10 kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice. *Genes Dev.* May 1987;1(3):268-76. doi: 10.1101/gad.1.3.268.
- Pirakitkul et al., PCRless library mutagenesis via oligonucleotide recombination in yeast. *Protein Sci.* Dec. 2010;19(12):2336-46. doi: 10.1002/pro.513.
- Plasterk et al., DNA inversions in the chromosome of *Escherichia coli* and in bacteriophage Mu: relationship to other site-specific recombination systems. *Proc Natl Acad Sci U S A.* Sep. 1983;80(17):5355-8.
- Plosky et al., CRISPR-Mediated Base Editing without DNA Double-Strand Breaks. *Mol Cell.* May 19, 2016;62(4):477-8. doi: 10.1016/j.molcel.2016.05.006.
- Plciennik et al., PCNA function in the activation and strand direction of MutLox endonuclease in mismatch repair. *Proc Natl Acad Sci U S A.* Sep. 14, 2010;107(37):16066-71. doi: 10.1073/pnas.1010662107. Epub Aug. 16, 2010.
- Poller et al., A leucine-to-proline substitution causes a defective alpha 1-antichymotrypsin allele associated with familial obstructive lung disease. *Genomics.* Sep. 1993;17(3):740-3.
- Popp et al., Sortagging: a versatile method for protein labeling. *Nat Chem Biol.* Nov. 2007;3(11):707-8. doi: 10.1038/nchembio.2007.31. Epub Sep. 23, 2007.
- Porteus, Design and testing of zinc finger nucleases for use in mammalian cells. *Methods Mol Biol.* 2008;435:47-61. doi: 10.1007/978-1-59745-232-8_4.

(56)

References Cited**OTHER PUBLICATIONS**

- Posnick et al., Imbalanced base excision repair increases spontaneous mutation and alkylation sensitivity in *Escherichia coli*. *J Bacteriol*. Nov. 1999;181(21):6763-71.
- Pospíšilová et al., Hydrolytic cleavage of N6-substituted adenine derivatives by eukaryotic adenine and adenosine deaminases. *Biosci Rep*. 2008;28(6):335-347. doi: 10.1042/BSR20080081.
- Pourcel et al., CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*. Mar. 2005;151(Pt 3):653-63.
- Prasad et al., Rev1 is a base excision repair enzyme with 5'-deoxyribose phosphate lyase activity. *Nucleic Acids Res*. Dec. 15, 2016;44(22):10824-10833. doi: 10.1093/nar/gkw869. Epub Sep. 28, 2016.
- Prashant et al., CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology* 2013;31(9):833-8.
- Prorocic et al., Zinc-finger recombinase activities in vitro. *Nucleic Acids Res*. Nov. 2011;39(21):9316-28. doi: 10.1093/nar/gkr652. Epub Aug. 17, 2011.
- Proudfoot et al., Zinc finger recombinases with adaptable DNA sequence specificity. *PLoS One*. Apr. 29, 2011;6(4):e19537. doi: 10.1371/journal.pone.0019537.
- Pruscha et al., Mechanistic studies of a signaling pathway activated by the organic dimerizer FK1012. *Chem Biol*. Nov. 1994;1(3):163-72. doi: 10.1016/1074-5521(94)90006-x.
- Prykhozhij et al., CRISPR multitargeter: a web tool to find common and unique CRISPR single guide RNA targets in a set of similar sequences. *PLoS One*. Mar. 5, 2015;10(3):e0119372. doi: 10.1371/journal.pone.0119372. eCollection 2015.
- Pu et al., Evolution of a split RNA polymerase as a versatile biosensor platform. *Nat Chem Biol*. Apr. 2017;13(4):432-438. doi: 10.1038/nchembio.2299. Epub Feb. 13, 2017.
- Putnam et al., Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with *Escherichia coli* uracil-DNA glycosylase. *J Mol Biol*. Mar. 26, 1999;287(2):331-46.
- Qi et al., Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals. *Nucleic Acids Res*. Jul. 2012;40(12):5775-86. doi: 10.1093/nar/gks168. Epub Mar. 1, 2012.
- Qi et al., Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. Feb. 28, 2013;152(5):1173-83. doi: 10.1016/j.cell.2013.02.022.
- Qu et al., Global mapping of binding sites for phic31 integrase in transgenic maden-darby bovine kidney cells using ChIP-seq. *Hereditas*. Jan. 14, 2019;156:3. doi: 10.1186/s41065-018-0079-z.
- Queen et al., Immunoglobulin gene transcription is activated by downstream sequence elements. *Cell*. Jul. 1983;33(3):741-8. doi: 10.1016/0092-8674(83)90016-8.
- Radany et al., Increased spontaneous mutation frequency in human cells expressing the phage PBS2-encoded inhibitor of uracil-DNA glycosylase. *Mutat Res*. Sep. 15, 2000;461(1):41-58. doi: 10.1016/s0921-8777(00)00040-9.
- Raghavan et al., Abstract 27: Therapeutic Targeting of Human Lipid Genes with in vivo CRISPR-Cas9 Genome Editing. Oral Abstract Presentations: Lipoprotein Metabolism and Therapeutic Targets. *Arterioscler Thromb Vasc Biol*. 2015;35(Suppl. 1):Abstract 27. 5 pages.
- Raillard et al., Targeting sites within HIV-1 cDNA with a DNA-cleaving ribozyme. *Biochemistry*. Sep. 10, 1996;35(36):11693-701. doi: 10.1021/bi960845g.
- Raina et al., PROTAC-induced BET protein degradation as a therapy for castration-resistant prostate cancer. *Proc Natl Acad Sci U S A*. Jun. 28, 2016;113(26):7124-9. doi: 10.1073/pnas.1521738113. Epub Jun. 6, 2016.
- Rakonjac et al., Roles of PIII in filamentous phage assembly. *J Mol Biol*. 1998; 282(1):25-41.
- Ramakrishna et al., Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res*. Jun. 2014;24(6):1020-7. doi: 10.1101/gr.171264.113. Epub Apr. 2, 2014.
- Ramamurthy et al., Identification of immunogenic B-cell epitope peptides of rubella virus El glycoprotein towards development of highly specific immunoassays and/or vaccine. *Conference Abstract*. 2019.
- Ramirez et al., Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res*. Jul. 2012;40(12):5560-8. doi: 10.1093/nar/gks179. Epub Feb. 28, 2012.
- Ramirez et al., Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods*. May 2008;5(5):374-5. doi: 10.1038/nmeth0508-374.
- Ran et al., Double Nicking by RNA-guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell*. Sep. 12, 2013;154(6):1380-9. doi: 10.1016/j.cell.2013.08.021. Epub Aug. 29, 2013.
- Ran et al., Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. Nov. 2013;8(11):2281-308. doi: 10.1038/nprot.2013.143. Epub Oct. 24, 2013.
- Ran et al., In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*. Apr. 9, 2015;520(7546):186-91. doi: 10.1038/nature14299. Epub Apr. 1, 2015.
- Ranzau et al., Genome, Epigenome, and Transcriptome Editing via Chemical Modification of Nucleobases in Living Cells. *Biochemistry*. Feb. 5, 2019;58(5):330-335. doi: 10.1021/acs.biochem.8b00958. Epub Dec. 12, 2018.
- Rashel et al., A novel site-specific recombination system derived from bacteriophage phiMR11. *Biochem Biophys Res Commun*. Apr. 4, 2008;368(2):192-8. doi: 10.1016/j.bbrc.2008.01.045. Epub Jan. 22, 2008.
- Rasila et al., Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, *Escherichia coli* mutator strain, and hydroxylamine treatment. *Anal Biochem*. May 1, 2009;388(1):71-80. doi: 10.1016/j.ab.2009.02.008. Epub Feb. 10, 2009.
- Raskin et al., Substitution of a single bacteriophage T3 residue in bacteriophage T7 RNA polymerase at position 748 results in a switch in promoter specificity. *J Mol Biol*. Nov. 20, 1992;228(2):506-15.
- Raskin et al., T7 RNA polymerase mutants with altered promoter specificities. *Proc Natl Acad Sci U S A*. Apr. 15, 1993;90(8):3147-51.
- Rath et al., Fidelity of end joining in mammalian episomes and the impact of Metnase on joint processing. *BMC Mol Biol*. Mar. 22, 2014;15:6. doi: 10.1186/1471-2199-15-6.
- Rauch et al., Programmable RNA Binding Proteins for Imaging and Therapeutics. *Biochemistry*. Jan. 30, 2018;57(4):363-364. doi: 10.1021/acs.biochem.7b01101. Epub Nov. 17, 2017.
- Ravishankar et al., X-ray analysis of a complex of *Escherichia coli* uracil DNA glycosylase (EcUDG) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG. *Nucleic Acids Res*. 26 (21): 4880-4887 (1998).
- Ray et al., A compendium of RNA-binding motifs for decoding gene regulation. *Nature*. Jul. 11, 2013;499(7457):172-7. doi: 10.1038/nature12311.
- Ray et al., Homologous recombination: ends as the means. *Trends Plant Sci*. Oct. 2002;7(10):435-40.
- Rebar et al., Phage display methods for selecting zinc finger proteins with novel DNA-binding specificities. *Methods Enzymol*. 1996;267:129-49.
- Rebuzzini et al., New mammalian cellular systems to study mutations introduced at the break site by non-homologous end-joining. *DNA Repair (Amst)*. May 2, 2005;4(5):546-55.
- Rees et al., Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci Adv*. May 8, 2019;5(5):eaax5717. doi: 10.1126/sciadv.aax5717.
- Rees et al., Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet*. Dec. 2018;19(12):770-788. doi: 10.1038/s41576-018-0059-1.

(56)

References Cited**OTHER PUBLICATIONS**

- Rees et al., Development of hRad51-Cas9 nickase fusions that mediate HDR without double-stranded breaks. *Nat Commun.* May 17, 2019;10(1):2212. doi: 10.1038/s41467-019-09983-4.
- Rees et al., Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat Commun.* Jun. 6, 2017;8:15790. doi: 10.1038/ncomms15790.
- Relph et al., Recent developments and current status of gene therapy using viral vectors in the United Kingdom. *BMJ.* 2004;329(7470):839-842. doi: 10.1136/bmj.329.7470.839.
- Remy et al., Gene transfer with a series of lipophilic DNA-binding molecules. *Bioconjug Chem.* Nov.-Dec. 1994;5(6):647-54. doi: 10.1021/bc00030a021.
- Ren et al., In-line Alignment and Mg²⁺ Coordination at the Cleavage Site of the env22 Twister Ribozyme. *Nat Commun.* Nov. 20, 2014;5:5534. doi: 10.1038/ncomms6534.
- Ren et al., Pistol Ribozyme Adopts a Pseudoknot Fold Facilitating Site-Specific In-Line Cleavage. *Nat Chem Biol.* Sep. 2016;12(9):702-8. doi: 10.1038/nchembio.2125. Epub Jul. 11, 2016.
- Reynaud et al., What role for AID: mutator, or assembler of the immunoglobulin mutasome? *Nat Immunol.* Jul. 2003;4(7):631-8.
- Reyon et al., FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol.* May 2012;30(5):460-5. doi: 10.1038/nbt.2170.
- Ribeiro et al., Protein Engineering Strategies to Expand CRISPR-Cas9 Applications. *Int J Genomics.* Aug. 2, 2018;2018:1652567. doi: 10.1155/2018/1652567.
- Richardson et al., Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol.* Mar. 2016;34(3):339-44. doi: 10.1038/nbt.3481. Epub Jan. 20, 2016.
- Richter et al., Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. *Viruses.* Oct. 19, 2012;4(10):2291-311. doi: 10.3390/v4102291.
- Riechmann et al., The C-terminal domain of TolA is the coreceptor for filamentous phage infection of *E. coli*. *Cell.* 1997; 90(2):351-60. PMID:9244308.
- Ringrose et al., The Kw recombinase, an integrase from *Kluyveromyces waltii*. *Eur J Biochem.* Sep. 15, 1997;248(3):903-12. doi: 10.1111/j.1432-1033.1997.00903.x.
- Risso et al., Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β -lactamases. *J Am Chem Soc.* Feb. 27, 2013;135(8):2899-902. doi: 10.1021/ja311630a. Epub Feb. 14, 2013.
- Ritchie et al., limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* Apr. 20, 2015;43(7):e47. doi: 10.1093/nar/gkv007. Epub Jan. 20, 2015.
- Robertson et al., DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci.* Mar. 2009;66(6):981-93. doi: 10.1007/s00018-009-8736-z.
- Robertson et al., Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature.* Mar. 29, 1990;344(6265):467-8. doi: 10.1038/344467a0.
- Robinson et al., The protein tyrosine kinase family of the human genome. *Oncogene.* Nov. 20, 2000;19(49):5548-57. doi: 10.1038/sj.onc.1203957.
- Rogozin et al., Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat Immunol.* Jun. 2007;8(6):647-56. doi: 10.1038/ni1463. Epub Apr. 29, 2007.
- Rong et al., Homologous recombination in human embryonic stem cells using CRISPR/Cas9 nickase and a long DNA donor template. *Protein Cell.* Apr. 2014;5(4):258-60. doi: 10.1007/s13238-014-0032-5.
- Rongrong et al., Effect of deletion mutation on the recombination activity of Cre recombinase. *Acta Biochim Pol.* 2005;52(2):541-4. Epub May 15, 2005.
- Roth et al., A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nat Chem Biol.* Jan. 2014;10(1):56-60. doi: 10.1038/nchembio.1386. Epub Nov. 17, 2013.
- Roth et al., Purification and characterization of murine retroviral reverse transcriptase expressed in *Escherichia coli*. *J Biol Chem.* Aug. 5, 1985;260(16):9326-35.
- Rouet et al., Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A.* Jun. 21, 1994;91(13):6064-8. doi: 10.1073/pnas.91.13.6064.
- Rouet et al., Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol.* Dec. 1994;14(12):8096-106. doi: 10.1128/mcb.14.12.8096.
- Rouet et al., Receptor-Mediated Delivery of CRISPR-Cas9 Endonuclease for Cell-Type-Specific Gene Editing. *J Am Chem Soc.* May 30, 2018;140(21):6596-6603. doi: 10.1021/jacs.8b01551. Epub May 18, 2018.
- Roundtree et al., YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *Elife.* Oct. 6, 2017;6:e31311. doi: 10.7554/elife.31311.
- Rowland et al., Regulatory mutations in Sin recombinase support a structure-based model of the synapsosome. *Mol Microbiol.* Oct. 2009;74(2):282-98. doi: 10.1111/j.1365-2958.2009.06756.x. Epub Jun. 8, 2009.
- Rowland et al., Sin recombinase from *Staphylococcus aureus*: synaptic complex architecture and transposon targeting. *Mol Microbiol.* May 2002;44(3):607-19. doi: 10.1046/j.1365-2958.2002.02897.x.
- Rowley, Chromosome translocations: dangerous liaisons revisited. *Nat Rev Cancer.* Dec. 2001;1(3):245-50. doi: 10.1038/35106108.
- Rubio et al., An adenosine-to-inosine tRNA-editing enzyme that can perform C-to-U deamination of DNA. *Proc Natl Acad Sci U S A.* May 8, 2007;104(19):7821-6. doi: 10.1073/pnas.0702394104. Epub May 1, 2007. PMID: 17483465; PMCID: PMC1876531.
- Rubio et al., Transfer RNA travels from the cytoplasm to organelles. *Wiley Interdiscip Rev RNA.* Nov.-Dec. 2011;2(6):802-17. doi: 10.1002/wrna.93. Epub Jul. 11, 2011.
- Rudolph et al., Synthetic riboswitches for the conditional control of gene expression in *Streptomyces coelicolor*. *Microbiology.* Jul. 2013;159(Pt 7):1416-22. doi: 10.1099/mic.0.067322-0. Epub May 15, 2013.
- Rutherford et al., Attachment site recognition and regulation of directionality by the serine integrases. *Nucleic Acids Res.* Sep. 2013;41(17):8341-56. doi: 10.1093/nar/gkt580. Epub Jul. 2, 2013.
- Ryu et al., Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. *Nat Biotechnol.* Jul. 2018;36(6):536-539. doi: 10.1038/nbt.4148. Epub Apr. 27, 2018.
- Rüfer et al., Non-contact positions impose site selectivity on Cre recombinase. *Nucleic Acids Res.* Jul. 1, 2002;30(13):2764-71. doi: 10.1093/nar/gkf399.
- Sadelain et al., Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer.* Dec. 1, 2011;12(1):51-8. doi: 10.1038/nrc3179.
- Sadowski, The Flp recombinase of the 2-microns plasmid of *Saccharomyces cerevisiae*. *Prog Nucleic Acid Res Mol Biol.* 1995;51:53-91.
- Safari et al., CRISPR Cpf1 proteins: structure, function and implications for genome editing. *Cell Biosci.* May 9, 2019;9:36. doi: 10.1186/s13578-019-0298-7.
- Sage et al., Proliferation of functional hair cells in vivo in the absence of the retinoblastoma protein. *Science.* Feb. 18, 2005;307(5712):1114-8. Epub Jan. 13, 2005.
- Sakuma et al., MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCH systems. *Nat Protoc.* Jan. 2016;11(1):118-33. doi: 10.1038/nprot.2015.140. Epub Dec. 17, 2015.
- Sale et al., Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Biol.* Feb. 23, 2012;13(3):141-52. doi: 10.1038/nrm3289.
- Saleh-Gohari et al., Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res.* Jul. 13, 2004;32(12):3683-8. Print 2004.

(56)

References Cited**OTHER PUBLICATIONS**

- Samal et al., Cationic polymers and their therapeutic potential. *Chem Soc Rev.* Nov. 7, 2012;41(21):7147-94. doi: 10.1039/c2cs35094g. Epub Aug. 10, 2012.
- Samanta et al., A reverse transcriptase ribozyme. *Elife.* Sep. 26, 2017;6:e31153. doi: 10.7554/eLife.31153.
- Samulski et al., Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *J Virol.* Sep. 1989;63(9):3822-8. doi: 10.1128/JVI.63.9.3822-3828.1989.
- Sander et al., CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol.* Apr. 2014;32(4):347-55. doi: 10.1038/nbt.2842. Epub Mar. 2, 2014.
- Sander et al., In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res.* Oct. 2013;41(19):e181. doi: 10.1093/nar/gkt716. Epub Aug. 14, 2013.
- Sander et al., Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol.* Aug. 5, 2011;29(8):697-8. doi: 10.1038/nbt.1934.
- Sang et al., A unique uracil-DNA binding protein of the uracil DNA glycosylase superfamily. *Nucleic Acids Res.* Sep. 30, 2015;43(17):8452-63. doi: 10.1093/nar/gkv854. Epub Aug. 24, 2015.
- Sang, Prospects for transgenesis in the chick. *Mech Dev.* Sep. 2004;121(9):1179-86.
- Sanjana et al., A transcription activator-like effector toolbox for genome engineering. *Nat Protoc.* Jan. 5, 2012;7(1):171-92. doi: 10.1038/nprot.2011.431.
- Santiago et al., Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proc Natl Acad Sci U S A.* Apr. 15, 2008;105(15):5809-14. doi: 10.1073/pnas.0800940105. Epub Mar. 21, 2008.
- Santoro et al., Directed evolution of the site specificity of Cre recombinase. *Proc Natl Acad Sci U S A.* Apr. 2, 2002;99(7):4185-90. Epub Mar. 19, 2002.
- Saparbaev et al., Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases. *Proc Natl Acad Sci U S A.* Jun. 21, 1994;91(13):5873-7. doi: 10.1073/pnas.91.13.5873.
- Saprauskas et al., The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* Nov. 2011;39(21):9275-82. doi: 10.1093/nar/gkr606. Epub Aug. 3, 2011.
- Sapunar et al., Dorsal root ganglion—a potential new therapeutic target for neuropathic pain. *J Pain Res.* 2012;5:31-8. doi: 10.2147/JPR.S26603. Epub Feb. 16, 2012.
- Saraconi et al., The RNA editing enzyme APOBEC1 induces somatic mutations and a compatible mutational signature is present in esophageal adenocarcinomas. *Genome Biol.* Jul. 31, 2014;15(7):417. doi: 10.1186/s13059-014-0417-z.
- Sarkar et al., HIV-1 proviral DNA excision using an evolved recombinase. *Science.* Jun. 29, 2007;316(5833):1912-5. doi: 10.1126/science.1141453.
- Sashital et al., Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol Cell.* Jun. 8, 2012;46(5):606-15. doi: 10.1016/j.molcel.2012.03.020. Epub Apr. 19, 2012.
- Sasidharan et al., The selection of acceptable protein mutations. *PNAS*; Jun. 12, 2007;104(24):10080-5. www.pnas.org/cgi/doi/10.1073.pnas.0703737104.
- Satomura et al., Precise genome-wide base editing by the CRISPR Nickase system in yeast. *Sci Rep.* May 18, 2017;7(1):2095. doi: 10.1038/s41598-017-02013-7.
- Saudek et al., A preliminary trial of the programmable implantable medication system for insulin delivery. *Engl J Med.* Aug. 31, 1989;321(9):574-9.
- Sauer et al., DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-cl regions of P1-related phages. *Nucleic Acids Res.* Nov. 18, 2004;32(20):6086-95. doi: 10.1093/nar/gkh941.
- Savic et al., Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *Elife.* May 29, 2018;7:e33761. doi: 10.7554/eLife.33761.
- Saville et al., A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell.* May 18, 1990;61(4):685-96. doi: 10.1016/0092-8674(90)90480-3.
- Savva et al., The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature.* Feb. 9, 1995;373(6514):487-93. doi: 10.1038/373487a0.
- Schaaper et al., Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J Biol Chem.* Nov. 15, 1993;268(32):23762-5.
- Schaaper et al., Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: the nature of in vivo DNA replication errors. *Proc Natl Acad Sci U S A.* Sep. 1987;84(17):6220-4.
- Schaefer et al., Understanding RNA modifications: the promises and technological bottlenecks of the ‘epitranscriptome’. *Open Biol.* May 2017;7(5):170077. doi: 10.1098/rsob.170077.
- Schechner et al., Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat Methods.* Jul. 2015;12(7):664-70. doi: 10.1038/nmeth.3433. Epub Jun. 1, 2015. Author manuscript entitled CRISPR Display: A modular method for locus-specific targeting of long noncoding RNAs and synthetic RNA devices in vivo.
- Schek et al., Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. *Mol Cell Biol.* Dec. 1992;12(12):5386-93. doi: 10.1128/mcb.12.12.5386.
- Schenk et al., MPDU1 mutations underlie a novel human congenital disorder of glycosylation, designated type If. *J Clin Invest.* Dec. 2001;108(11):1687-95. doi: 10.1172/JCI13419.
- Schmitz et al., Behavioral abnormalities in prion protein knockout mice and the potential relevance of PrP(C) for the cytoskeleton. *Prion.* 2014;8(6):381-6. doi: 10.4161/19336896.2014.983746.
- Schriefer et al., Low pressure DNA shearing: a method for random DNA sequence analysis. *Nucleic Acids Res.* Dec. 25, 1990;18(24):7455-6.
- Schultz et al., Expression and secretion in yeast of a 400-kDa envelope glycoprotein derived from Epstein-Barr virus. *Gene.* 1987;54(1):113-23. doi: 10.1016/0378-1119(87)90353-2.
- Schultz et al., Oligo-2'-fluoro-2'-deoxyribonucleotide N3'-->P5' phosphoramidates: synthesis and properties. *Nucleic Acids Res.* Aug. 1, 1996;24(15):2966-73.
- Schwank et al., Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell.* Dec. 5, 2013;13(6):653-8. doi: 10.1016/j.stem.2013.11.002.
- Schwartz et al., Post-translational enzyme activation in an animal via optimized conditional protein splicing. *Nat Chem Biol.* Jan. 2007;3(1):50-4. Epub Nov. 26, 2006.
- Schwarze et al., In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science.* Sep. 3, 1999;285(5433):1569-72.
- Schöller et al., Interactions, localization, and phosphorylation of the m6A generating METTL3-METTL14-WTAP complex. *RNA.* Apr. 2018;24(4):499-512. doi: 10.1261/rna.064063.117. Epub Jan. 18, 2018.
- Sclimenti et al., Directed evolution of a recombinase for improved genomic integration at a native human sequence. *Nucleic Acids Res.* Dec. 15, 2001;29(24):5044-51.
- SCORE Results for Luetticken et al., Complete genome sequence of a *Streptococcus dysgalactiae* subsp. RT equisimilis strain possessing Lancefield's group A antigen. RL Submitted to the EMBL/GenBank/DDBJ databases. May 2012. 3 pages.
- SCORE Results for Okumura et al., Evolutionary paths of streptococcal and staphylococcal superantigens. RL *BMC Genomics.* 2012;13:404-404. 3 pages.
- SCORE Results for Shimomura et al., Complete Genome Sequencing and Analysis of a Lancefield Group G RT *Streptococcus dysgalactiae* Subsp. Equisimilis Strain Causing Streptococcal RT Toxic Shock Syndrome (STSS). RL *BMC Genomics.* 2011;12:17-17. 3 pages.

(56)

References Cited**OTHER PUBLICATIONS**

- Scott et al., Production of cyclic peptides and proteins in vivo. *Proc Natl Acad Sci U S A.* Nov. 23, 1999;96(24):13638-43. doi: 10.1073/pnas.96.24.13638.
- Sebastián-Martín et al., Transcriptional inaccuracy threshold attenuates differences in RNA-dependent DNA synthesis fidelity between retroviral reverse transcriptases. *Sci Rep.* Jan. 12, 2018;8(1):627. doi: 10.1038/s41598-017-18974-8.
- Seed, An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature.* Oct. 29-Nov. 4, 1987;329(6142):840-2. doi: 10.1038/329840a0.
- Sefton et al., Implantable pumps. *Crit Rev Biomed Eng.* 1987;14(3):201-40.
- Segal et al., Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci U S A.* Mar. 16, 1999;96(6):2758-63.
- Sells et al., Delivery of protein into cells using polycationic liposomes. *Biotechniques.* Jul. 1995;19(1):72-6, 78.
- Semenova et al., Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci U S A.* Jun. 21, 2011;108(25):10098-103. doi: 10.1073/pnas.1104144108. Epub Jun. 6, 2011.
- Semple et al., Rational design of cationic lipids for siRNA delivery. *Nat Biotechnol.* Feb. 2010;28(2):172-6. doi: 10.1038/nbt.1602. Epub Jan. 17, 2010.
- Serganov et al., Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature.* Mar. 12, 2009;458(7235):233-7. doi: 10.1038/nature07642. Epub Jan. 25, 2009.
- Serganov et al., Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem Biol.* Dec. 2004;11(12):1729-41.
- Serganov et al., Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature.* Jun. 29, 2006;441(7097):1167-71. Epub May 21, 2006.
- Seripa et al., The missing ApoE allele. *Ann Hum Genet.* Jul. 2007;71(Pt 4):496-500. Epub Jan. 22, 2007.
- Serrano-Heras et al., Protein p56 from the *Bacillus subtilis* phage phi29 inhibits DNA-binding ability of uracil-DNA glycosylase. *Nucleic Acids Res.* 2007;35(16):5393-401. Epub Aug. 13, 2007.
- Setten et al., The current state and future directions of RNAi-based therapeutics. *Nat Rev Drug Discov.* Jun. 2019;18(6):421-446. doi: 10.1038/s41573-019-0017-4.
- Severinov et al., Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *J Biol Chem.* Jun. 26, 1998;273(26):16205-9. doi: 10.1074/jbc.273.26.16205.
- Sha et al., Monobodies and other synthetic binding proteins for expanding protein science. *Protein Sci.* May 2017;26(5):910-924. doi: 10.1002/pro.3148. Epub Mar. 24, 2017.
- Shah et al., Inteins: nature's gift to protein chemists. *Chem Sci.* 2014;5(1):446-461.
- Shah et al., Kinetic control of one-pot trans-splicing reactions by using a wild-type and designed split intein. *Angew Chem Int Ed Engl.* Jul. 11, 2011;50(29):6511-5. doi: 10.1002/anie.201102909. Epub Jun. 8, 2011.
- Shah et al., Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol.* May 2013;10(5):891-9. doi: 10.4161/rna.23764. Epub Feb. 12, 2013.
- Shah et al., Target-specific variants of Flp recombinase mediate genome engineering reactions in mammalian cells. *FEBS J.* Sep. 2015;282(17):3323-33. doi: 10.1111/febs.13345. Epub Jul. 1, 2015.
- Shaikh et al., Chimeras of the Flp and Cre recombinases: tests of the mode of cleavage by Flp and Cre. *J Mol Biol.* Sep. 8, 2000;302(1):27-48.
- Shalem et al., Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science.* Jan. 3, 2014;343(6166):84-7. doi: 10.1126/science.1247005. Epub Dec. 12, 2013.
- Shalem et al., High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet.* May 2015;16(5):299-311. doi: 10.1038/nrg3899. Epub Apr. 9, 2015.
- Sharbeen et al., Ectopic restriction of DNA repair reveals that UNG2 excises AID-induced uracils predominantly or exclusively during G1 phase. *J Exp Med.* May 7, 2012;209(5):965-74. doi: 10.1084/jem.20112379. Epub Apr. 23, 2012.
- Sharer et al., The ARF-like 2 (ARL2)-binding protein, BART. Purification, cloning, and initial characterization. *J Biol Chem.* Sep. 24, 1999;274(39):27553-61. doi: 10.1074/jbc.274.39.27553.
- Sharma et al., Efficient introduction of aryl bromide functionality into proteins in vivo. *FEBS Lett.* Feb. 4, 2000;467(1):37-40.
- Sharma et al., Identification of novel methyltransferases, Bmt5 and Bmt6, responsible for the m3U methylations of 25S rRNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* Mar. 2014;42(5):3246-60. doi: 10.1093/nar/gkt1281. Epub Dec. 11, 2013.
- Sharon et al., Functional Genetic Variants Revealed by Massively Parallel Precise Genome Editing. *Cell.* Oct. 4, 2018;175(2):544-557.e16. doi: 10.1016/j.cell.2018.08.057. Epub Sep. 20, 2018.
- Shaw et al., Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet.* Apr. 1, 2004;13 Spec No. 1:R57-64. doi: 10.1093/hmg/ddh073. Epub Feb. 5, 2004.
- Shecherbакова et al., Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat Methods.* Aug. 2013;10(8):751-4. doi: 10.1038/nmeth.2521. Epub Jun. 16, 2013.
- Shechner et al., Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat Methods.* Jul. 2015;12(7):664-70. doi: 10.1038/nmeth.3433. Epub Jun. 1, 2015.
- Shee et al., Engineered proteins detect spontaneous DNA breakage in human and bacterial cells. *Elife.* Oct. 29, 2013;2:e01222. doi: 10.7554/elife.01222.
- Shen et al., Herpes simplex virus 1 (HSV-1) for cancer treatment. *Cancer Gene Ther.* Nov. 2006;13(11):975-92. doi: 10.1038/sj.cgt.7700946. Epub Apr. 7, 2006.
- Shen et al., Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature.* Nov. 2018;563(7733):646-651. doi: 10.1038/s41586-018-0686-x. Epub Nov. 7, 2018.
- Shen, Data processing, Modeling and Analysis scripts for CRISPR-inDelphi. GitHub—maxwshen/inDELPHI-dataProcessingAnalysis at 6b68e3cec73c9358fef6e5f178a935f3c2a4118f. Apr. 10, 2018. Retrieved online via <https://github.com/maxwshen/inDELPHI-dataProcessingAnalysis/tree/6b68e3cec73c9358fef6e5f178a935f3c2a4118f>. Last retrieved on Jul. 26, 2021. 2 pages.
- Sheridan, First CRISPR-Cas patent opens race to stake out intellectual property. *Nat Biotechnol.* 2014;32(7):599-601.
- Sheridan, Gene therapy finds its niche. *Nat Biotechnol.* Feb. 2011;29(2):121-8. doi: 10.1038/nbt.1769.
- Sherwood et al., Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat Biotechnol.* Feb. 2014;32(2):171-178. doi: 10.1038/nbt.2798. Epub Jan. 19, 2014.
- Shi et al., Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat Struct Mol Biol.* Feb. 2017;24(2):131-139. doi: 10.1038/nsmb.3344. Epub Dec. 19, 2016.
- Shi et al., YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res.* Mar. 2017;27(3):315-328. doi: 10.1038/cr.2017.15. Epub Jan. 20, 2017.
- Shimantani et al., Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat Biotechnol.* May 2017;35(5):441-443. doi: 10.1038/nbt.3833. Epub Mar. 27, 2017.
- Shimojima et al., Spinocerebellar ataxias type 27 derived from a disruption of the fibroblast growth factor 14 gene with mimicking phenotype of paroxysmal non-kinesigenic dyskinesia. *Brain Dev.* Mar. 2012;34(3):230-3. doi: 10.1016/j.braindev.2011.04.014. Epub May 19, 2011.
- Shin et al., CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nat Commun.* May 31, 2017;8:15464. doi: 10.1038/ncomms15464.
- Shindo et al., A Comparison of Two Single-Stranded DNA Binding Models by Mutational Analysis of APOBEC3G. *Biology (Basel).* Aug. 2, 2012;1(2):260-76. doi: 10.3390/biology1020260.
- Shingledecker et al., Molecular dissection of the Mycobacterium tuberculosis RecA intein: design of a minimal intein and of a

(56)

References Cited**OTHER PUBLICATIONS**

- trans-splicing system involving two intein fragments. *Gene*. Jan. 30, 1998;207(2):187-95. doi: 10.1016/s0378-1119(97)00624-0.
- Shmakov et al., Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems. *Molecular Cell*. Nov. 2015;60(3):385-97.
- Shmakov et al., Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Microbiol*. Mar. 2017;15(3):169-182. doi: 10.1038/nrmicro.2016.184. Epub Jan. 23, 2017.
- Shultz et al., A genome-wide analysis of FRT-like sequences in the human genome. *PLoS One*. Mar. 23, 2011;6(3):e18077. doi: 10.1371/journal.pone.0018077.
- Siebert et al., An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res*. Mar. 25, 1995;23(6):1087-8.
- Silas et al., Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science*. Feb. 26, 2016;351(6276):aad4234. doi: 10.1126/science.aad4234.
- Silva et al., Selective disruption of the DNA polymerase III α - β complex by the umuD gene products. *Nucleic Acids Res*. Jul. 2012;40(12):5511-22. doi: 10.1093/nar/gks229. Epub Mar. 9, 2012.
- Simonelli et al., Base excision repair intermediates are mutagenic in mammalian cells. *Nucleic Acids Res*. Aug. 2, 2005;33(14):4404-11. Print 2005.
- Singh et al., Cross-talk between diverse serine recombinases. *J Mol Biol*. Jan. 23, 2014;426(2):318-31. doi: 10.1016/j.jmb.2013.10.013. Epub Oct. 22, 2013.
- Singh et al., Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat Commun*. Sep. 14, 2016;7:12778. doi: 10.1038/ncomms12778.
- Singh et al., Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). *Proc Natl Acad Sci U S A*. May 22, 2018;115(21):5444-5449. doi: 10.1073/pnas.1718686115. Epub May 7, 2018.
- Sirk et al., Expanding the zinc-finger recombinase repertoire: directed evolution and mutational analysis of serine recombinase specificity determinants. *Nucleic Acids Res*. Apr. 2014;42(7):4755-66. doi: 10.1093/nar/gkt1389. Epub Jan. 21, 2014.
- Siu et al., Riboregulated toehold-gated gRNA for programmable CRISPR-Cas9 function. *Nat Chem Biol*. Mar. 2019;15(3):217-220. doi: 10.1038/s41589-018-0186-1. Epub Dec. 10, 2018.
- Sivalingam et al., Biosafety assessment of site-directed transgene integration in human umbilical cord-lining cells. *Mol Ther*. Jul. 2010;18(7):1346-56. doi: 10.1038/mt.2010.61. Epub Apr. 27, 2010.
- Sjöblom et al., The consensus coding sequences of human breast and colorectal cancers. *Science*. Oct. 13, 2006;314(5797):268-74. Epub Sep. 7, 2006.
- Skretas et al., Regulation of protein activity with small-molecule-controlled inteins. *Protein Sci*. Feb. 2005;14(2):523-32. Epub Jan. 4, 2005.
- Slaymaker et al., Rationally engineered Cas9 nucleases with improved specificity. *Science*. Jan. 1, 2016;351(6268):84-8. doi: 10.1126/science.aad5227. Epub Dec. 1, 2015.
- Sledz et al., Structural insights into the molecular mechanism of the m(6)A writer complex. *Elife*. Sep. 14, 2016;5:e18434. doi: 10.7554/elife.18434.
- Slupphaug et al., A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*. Nov. 7, 1996;384(6604):87-92. doi: 10.1038/384087a0.
- Smargon et al., Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol Cell*. Feb. 16, 2017;65(4):618-630.e7. doi: 10.1016/j.molcel.2016.12.023. Epub Jan. 5, 2017.
- Smith et al., Diversity in the serine recombinases. *Mol Microbiol*. Apr. 2002;44(2):299-307. Review.
- Smith et al., Expression of a dominant negative retinoic acid receptor γ in Xenopus embryos leads to partial resistance to retinoic acid. *Roux Arch Dev Biol*. Mar. 1994;203(5):254-265. doi: 10.1007/BF00360521.
- Smith et al., Herpesvirus transport to the nervous system and back again. *Annu Rev Microbiol*. 2012;66:153-76. doi: 10.1146/annurev-micro-092611-150051. Epub Jun. 15, 2012.
- Smith et al., Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Mol Cell Biol*. Dec. 1983;3(12):2156-65. doi: 10.1128/mcb.3.12.2156.
- Smith et al., Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*. Jul. 15, 1988;67(1):31-40. doi: 10.1016/0378-1119(88)90005-4.
- Smith, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*. Jun. 14, 1985;238(4705):1315-7.
- Smith, Phage-encoded Serine Integrases and Other Large Serine Recombinases. *Microbiol Spectr*. Aug. 2015;3(4). doi: 10.1128/microbiolspec.MDNA3-0059-2014.
- Somanathan et al., AAV vectors expressing LDLR gain-of-function variants demonstrate increased efficacy in mouse models of familial hypercholesterolemia. *Circ Res*. Aug. 29, 2014;115(6):591-9. doi: 10.1161/CIRCRESAHA.115.304008. Epub Jul. 14, 2014.
- Sommerfelt et al., Receptor interference groups of 20 retroviruses plating on human cells. *Virology*. May 1990;176(1):58-69. doi: 10.1016/0042-6822(90)90230-o.
- Song et al., Adenine base editing in an adult mouse model of tyrosinaemia. *Nat Biomed Eng*. Jan. 2020;4(1):125-130. doi: 10.1038/s41551-019-0357-8. Epub Feb. 25, 2019.
- Southworth et al., Control of protein splicing by intein fragment reassembly. *EMBO J*. Feb. 16, 1998;17(4):918-26. doi: 10.1093/emboj/17.4.918.
- Southworth et al., Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques*. Jul. 1999;27(1):110-4, 116, 118-20. doi: 10.2144/99271st04.
- Spencer et al., A general strategy for producing conditional alleles of Src-like tyrosine kinases. *Proc Natl Acad Sci U S A*. Oct. 10, 1995;92(21):9805-9. doi: 10.1073/pnas.92.21.9805.
- Spencer et al., Controlling signal transduction with synthetic ligands. *Science*. Nov. 12, 1993;262(5136):1019-24. doi: 10.1126/science.7694365.
- Spencer et al., Functional analysis of Fas signaling in vivo using synthetic inducers of dimerization. *Curr Biol*. Jul. 1, 1996;6(7):839-47. doi: 10.1016/s0960-9822(02)00607-3.
- Srivastava et al., An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell*. Dec. 21, 2012;151(7):1474-87. doi: 10.1016/j.cell.2012.11.054.
- Stadtman, Selenocysteine. *Annu Rev Biochem*. 1996;65:83-100.
- Stamos et al., Structure of a Thermostable Group II Intron Reverse Transcriptase with Template-Primer and Its Functional and Evolutionary Implications. *Mol Cell*. Dec. 7, 2017;68(5):926-939.e4. doi: 10.1016/j.molcel.2017.10.024. Epub Nov. 16, 2017.
- Steele et al., The prion protein knockout mouse: a phenotype under challenge. *Prion*. Apr.-Jun. 2007;1(2):83-93. doi: 10.4161/pri.1.2.4346. Epub Apr. 25, 2007.
- Steiner et al., The neurotropic herpes viruses: herpes simplex and varicella-zoster. *Lancet Neurol*. Nov. 2007;6(11):1015-28. doi: 10.1016/S1474-4422(07)70267-3.
- Stella et al., Structure of the Cpf1 endonuclease R-loop complex after target DNA cleavage. *Nature*. Jun. 22, 2017;546(7659):559-563. doi: 10.1038/nature22398. Epub May 31, 2017.
- Stenglein et al., APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nat Struct Mol Biol*. Feb. 2010;17(2):222-9. doi: 10.1038/nsmb.1744. Epub Jan. 10, 2010.
- Stenson et al., The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet*. Jun. 2017;136(6):665-677. doi: 10.1007/s00439-017-1779-6. Epub Mar. 27, 2017.
- Stephens et al., The landscape of cancer genes and mutational processes in breast cancer. *Nature*. Jun. 2012;486:400-404. doi: 10.1038/nature11017.
- Sternberg et al., Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*. Nov. 5, 2015;527(7576):110-3. doi: 10.1038/nature15544. Epub Oct. 28, 2015.

(56)

References Cited**OTHER PUBLICATIONS**

- Sternberg et al., DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. Mar. 6, 2014;507(7490):62-7. doi: 10.1038/nature13011. Epub Jan. 29, 2014.
- Sterne-Weiler et al., Exon identity crisis: disease-causing mutations that disrupt the splicing code. *Genome Biol.* Jan. 23, 2014;15(1):201. doi: 10.1186/gb4150.
- Stevens et al., A promiscuous split intein with expanded protein engineering applications. *Proc Natl Acad Sci U S A.* Aug. 8, 2017;114(32):8538-8543. doi: 10.1073/pnas.1701083114. Epub Jul. 24, 2017.
- Stockwell et al., Probing the role of homomeric and heteromeric receptor interactions in TGF-beta signaling using small molecule dimerizers. *Curr Biol.* Jun. 18, 1998;8(13):761-70. doi: 10.1016/s0960-9822(98)70299-4.
- Strecker et al., Engineering of CRISPR-Cas12b for human genome editing. *Nat Commun.* Jan. 22, 2019;10(1):212. doi: 10.1038/s41467-018-08224-4.
- Strecker et al., RNA-guided DNA insertion with CRISPR-associated transposases. *Science*. Jul. 5, 2019;365(6448):48-53. doi: 10.1126/science.aax9181. Epub Jun. 6, 2019.
- Strutt et al., RNA-dependent RNA targeting by CRISPR-Cas9. *eLife*. Jan. 5, 2018;7:e32724. doi: 10.7554/eLife.32724.
- Su et al., Human DNA polymerase ? has reverse transcriptase activity in cellular environments. *J Biol Chem.* Apr. 12, 2019;294(15):6073-6081. doi: 10.1074/jbc.RA119.007925. Epub Mar. 6, 2019.
- Sudarsan et al., An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev.* Nov. 1, 2003;17(21):2688-97.
- Sudarsan et al., Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science*. Jul. 18, 2008;321(5887):411-3. doi: 10.1126/science.1159519.
- Suess et al., A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo. *Nucleic Acids Res.* Mar. 5, 2004;32(4):1610-4.
- Sullenger et al., Ribozyme-mediated repair of defective mRNA by targeted, trans-splicing. *Nature*. Oct. 13, 1994;371(6498):619-22. doi: 10.1038/371619a0.
- Sun et al., Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Mol Biosyst.* Apr. 2012;8(4):1255-63. doi: 10.1039/c2mb05461b. Epub Feb. 3, 2012.
- Sun et al., The CRISPR/Cas9 system for gene editing and its potential application in pain research. *Transl Periop & Pain Med.* Aug. 3, 2016;1(3):22-33.
- Surun et al., High Efficiency Gene Correction in Hematopoietic Cells by Donor-Template-Free CRISPR/Cas9 Genome Editing. *Mol Ther Nucleic Acids*. Mar. 2, 2018;10:1-8. doi: 10.1016/j.omtn.2017.11.001. Epub Nov. 10, 2017.
- Suzuki et al., Crystal structures reveal an elusive functional domain of pyrrolyls-tRNA synthetase. *Nat Chem Biol.* Dec. 2017;13(12):1261-1266. doi: 10.1038/nchembio.2497. Epub Oct. 16, 2017.
- Suzuki et al., In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature*. Dec. 1, 2016;540(7631):144-149. doi: 10.1038/nature20565. Epub Nov. 16, 2016.
- Suzuki et al., VCre/VloxP and SCre/SloxP: new site-specific recombination systems for genome engineering. *Nucleic Acids Res.* Apr. 2011;39(8):e49. doi: 10.1093/nar/gkq1280. Epub Feb. 1, 2011.
- Swarts et al., Argonaute of the archaeon Pyrococcus furiosus is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.* May 26, 2015;43(10):5120-9. doi: 10.1093/nar/gkv415. Epub Apr. 29, 2015.
- Swarts et al., DNA-guided DNA interference by a prokaryotic Argonaute. *Nature*. Mar. 13, 2014;507(7491):258-61. doi: 10.1038/nature12971. Epub Feb. 16, 2014.
- Swarts et al., The evolutionary journey of Argonaute proteins. *Nat Struct Mol Biol.* Sep. 2014;21(9):743-53. doi: 10.1038/nsmb.2879. Szczepek et al., Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol.* Jul. 2007;25(7):786-93. Epub Jul. 1, 2007.
- Tabebardar et al., In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science*. Jan. 22, 2016;351(6271):407-411. doi: 10.1126/science.aad5177. Epub Dec. 31, 2015.
- Tagalakis et al., Lack of RNA-DNA oligonucleotide (chimeroplast) mutagenic activity in mouse embryos. *Mol Reprod Dev.* Jun. 2005;71(2):140-4.
- Tahara et al., Potent and Selective Inhibitors of 8-Oxoguanine DNA Glycosylase. *J Am Chem Soc.* Feb. 14, 2018;140(6):2105-2114. doi: 10.1021/jacs.7b09316. Epub Feb. 5, 2018.
- Tajiri et al., Functional cooperation of MutT, MutM and MutY proteins in preventing mutations caused by spontaneous oxidation of guanine nucleotide in *Escherichia coli*. *Mutat Res.* May 1995;336(3):257-67. doi: 10.1016/0921-8777(94)00062-b.
- Takimoto et al., Stereochemical basis for engineered pyrrolyls-tRNA synthetase and the efficient in vivo incorporation of structurally divergent non-native amino acids. *ACS Chem Biol.* Jul. 15, 2011;6(7):733-43. doi: 10.1021/cb200057a. Epub May 5, 2011.
- Tambunan et al., Vaccine Design for H5N1 Based on B- and T-cell Epitope Predictions. *Bioinform Biol Insights*. Apr. 28, 2016;10:27-35. doi: 10.4137/BBI.S38378.
- Tanenbaum et al., A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. Oct. 23, 2014;159(3):635-46. doi: 10.1016/j.cell.2014.09.039. Epub Oct. 9, 2014.
- Tanese et al., Expression of enzymatically active reverse transcriptase in *Escherichia coli*. *Proc Natl Acad Sci U S A.* Aug. 1985;82(15):4944-8. doi: 10.1073/pnas.82.15.4944.
- Tang et al., Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat Commun.* Jun. 28, 2017;8:15939. doi: 10.1038/ncomms15939.
- Tang et al., Evaluation of Bioinformatic Programmes for the Analysis of Variants within Splice Site Consensus Regions. *Adv Bioinformatics*. 2016;2016:5614058. doi: 10.1155/2016/5614058. Epub May 24, 2016.
- Tang et al., Rewritable multi-event analog recording in bacterial and mammalian cells. *Science*. Apr. 13, 2018;360(6385):eaap8992. doi: 10.1126/science.aap8992. Epub Feb. 15, 2018.
- Tassabehji, Williams-Beuren syndrome: a challenge for genotype-phenotype correlations. *Hum Mol Genet.* Oct. 15, 2003;12 Spec No. 2:R229-37. doi: 10.1093/hmg/ddg299. Epub Sep. 2, 2003.
- Taube et al., Reverse transcriptase of mouse mammary tumour virus: expression in bacteria, purification and biochemical characterization. *Biochem J.* Feb. 1, 1998;329 (Pt 3)(Pt 3):579-87. doi: 10.1042/bj3290579. Erratum in: *Biochem J* Jun. 15, 1998;332(Pt 3):808.
- Tebas et al., Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *Engl J Med.* Mar. 6, 2014;370(10):901-10. doi: 10.1056/NEJMoa1300662.
- Tee et al., Polishing the craft of genetic diversity creation in directed evolution. *Biotechnol Adv.* Dec. 2013;31(8):1707-21. doi: 10.1016/j.biotechadv.2013.08.021. Epub Sep. 6, 2013.
- Telenti et al., The *Mycobacterium xenopi* Gyra protein splicing element: characterization of a minimal intein. *J Bacteriol.* Oct. 1997;179(20):6378-82. doi: 10.1128/jb.179.20.6378-6382.1997.
- Telesnitsky et al., RNase H domain mutations affect the interaction between Moloney murine leukemia virus reverse transcriptase and its primer-template. *Proc Natl Acad Sci U S A.* Feb. 15, 1993;90(4):1276-80. doi: 10.1073/pnas.90.4.1276.
- Teng et al., Mutational analysis of apolipoprotein B mRNA editing enzyme (APOBEC1), structure-function relationships of RNA editing and dimerization. *J Lipid Res.* Apr. 1999;40(4):623-35.
- Tessarollo et al., Targeted mutation in the neurotrophin-3 gene results in loss of muscle sensory neurons. *Proc Natl Acad Sci U S A.* Dec. 6, 1994;91(25):11844-8.
- Tesson et al., Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol.* Aug. 5, 2011;29(8):695-6. doi: 10.1038/nbt.1940.
- Thompson et al., Cellular uptake mechanisms and endosomal trafficking of supercharged proteins. *Chem Biol.* Jul. 27, 2012;19(7):831-43. doi: 10.1016/j.chembiol.2012.06.014.

(56)

References Cited**OTHER PUBLICATIONS**

- Thompson et al., Engineering and identifying supercharged proteins for macromolecule delivery into mammalian cells. *Methods Enzymol.* 2012;503:293-319. doi: 10.1016/B978-0-12396962-0.00012-4.
- Thompson et al., The Future of Multiplexed Eukaryotic Genome Engineering. *ACS Chem Biol.* Feb. 16, 2018;13(2):313-325. doi: 10.1021/acscchembio.7b00842. Epub Dec. 28, 2017.
- Thomson et al., Mutational analysis of loxP sites for efficient Cre-mediated insertion into genomic DNA. *Genesis.* Jul. 2003;36(3):162-7. doi: 10.1002/gene.10211.
- Thorpe et al., Functional correction of episomal mutations with short DNA fragments and RNA-DNA oligonucleotides. *J Gene Med.* Mar.-Apr. 2002;4(2):195-204.
- Thuronyi et al., Continuous evolution of base editors with expanded target compatibility and improved activity. *Nat Biotechnol.* Sep. 2019;37(9):1070-1079. doi: 10.1038/s41587-019-0193-0. Epub Jul. 22, 2019.
- Thyagarajan et al., Creation of engineered human embryonic stem cell lines using phiC31 integrase. *Stem Cells.* Jan. 2008;26(1):119-26. doi: 10.1634/stemcells.2007-0283. Epub Oct. 25, 2007.
- Thyagarajan et al., Mammalian genomes contain active recombinase recognition sites. *Gene.* Feb. 22, 2000;244(1-2):47-54.
- Thyagarajan et al., Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Mol Cell Biol.* Jun. 2001;21(12):3926-34.
- Tinland et al., The T-DNA-linked VirD2 protein contains two distinct functional nuclear localization signals. *Proc Natl Acad Sci U S A.* Aug. 15, 1992;89(16):7442-6. doi: 10.1073/pnas.89.16.7442.
- Tirumalai et al., Recognition of core-type DNA sites by lambda integrase. *J Mol Biol.* Jun. 12, 1998;279(3):513-27.
- Tom et al., Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *J Biol Chem.* Apr. 7, 2000;275(14):10498-505. doi: 10.1074/jbc.275.14.10498.
- Tone et al., Single-stranded DNA binding protein Gp5 of *Bacillus subtilis* phage ?29 is required for viral DNA replication in growth-temperature dependent fashion. *Biosci Biotechnol Biochem.* 2012;76(12):2351-3. doi: 10.1271/bbb.120587. Epub Dec. 7, 2012.
- Toor et al., Crystal structure of a self-spliced group II intron. *Science.* Apr. 4, 2008;320(5872):77-82. doi: 10.1126/science.1153803.
- Toro et al., On the Origin and Evolutionary Relationships of the Reverse Transcriptases Associated With Type III CRISPR-Cas Systems. *Front Microbiol.* Jun. 15, 2018;9:1317. doi: 10.3389/fmicb.2018.01317.
- Toro et al., The Reverse Transcriptases Associated with CRISPR-Cas Systems. *Sci Rep.* Aug. 2, 2017;7(1):7089. doi: 10.1038/s41598-017-0782-y.
- Torres et al., Non-integrative lentivirus drives high-frequency cre-mediated cassette exchange in human cells. *PLoS One.* 2011;6(5):e19794. doi: 10.1371/journal.pone.0019794. Epub May 23, 2011.
- Tourdot et al., A general strategy to enhance immunogenicity of low-affinity HLA-A2. 1-associated peptides: implication in the identification of cryptic tumor epitopes. *Eur J Immunol.* Dec. 2000;30(12):3411-21.
- Townsend et al., Role of HFE in iron metabolism, hereditary haemochromatosis, anaemia of chronic disease, and secondary iron overload. *Lancet.* Mar. 2, 2002;359(9308):786-90. doi: 10.1016/S0140-6736(02)07885-6.
- Tracewell et al., Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr Opin Chem Biol.* Feb. 2009;13(1):3-9. doi: 10.1016/j.cbpa.2009.01.017. Epub Feb. 25, 2009.
- Tratschin et al., A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the prokaryotic gene for chloramphenicol acetyltransferase. *Mol Cell Biol.* Oct. 1984;4(10):2072-81. doi: 10.1128/mcb.4.10.2072.
- Tratschin et al., Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. *Mol Cell Biol.* Nov. 1985;5(11):3251-60. doi: 10.1128/mcb.5.11.3251.
- Trausch et al., The structure of a tetrahydrofolate-sensing riboswitch reveals two ligand binding sites in a single aptamer. *Structure.* Oct. 12, 2011;19(10):1413-23. doi: 10.1016/j.str.2011.06.019. Epub Sep. 8, 2011.
- Traxler et al., A genome-editing strategy to treat ?-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat Med.* Sep. 2016;22(9):987-90. doi: 10.1038/nm.4170. Epub Aug. 15, 2016.
- Trojan et al., Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology.* Jan. 2002;122(1):211-9. doi: 10.1053/gast.2002.30296.
- Trudeau et al., On the Potential Origins of the High Stability of Reconstructed Ancestral Proteins. *Mol Biol Evol.* Oct. 2016;33(10):2633-41. doi: 10.1093/molbev/msw138. Epub Jul. 12, 2016.
- Truong et al., Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res.* Jul. 27, 2015;43(13):6450-8. doi: 10.1093/nar/gkv601. Epub Jun. 16, 2015. With Supplementary Data.
- Tsai et al., CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods.* Jun. 2017;14(6):607-614. doi: 10.1038/nmeth.4278. Epub May 1, 2017.
- Tsai et al., Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol.* Jun. 2014;32(6):569-76. doi: 10.1038/nbt.2908. Epub Apr. 25, 2014.
- Tsai et al., GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol.* Feb. 2015;33(2):187-97. doi: 10.1038/nbt.3117. Epub Dec. 16, 2014.
- Tsang et al., Specialization of the DNA-cleaving activity of a group I ribozyme through in vitro evolution. *J Mol Biol.* Sep. 13, 1996;262(1):31-42. doi: 10.1006/jmbi.1996.0496.
- Tsutakawa et al., Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell.* Apr. 15, 2011;145(2):198-211. doi: 10.1016/j.cell.2011.03.004.
- Turan et al., Recombinase-mediated cassette exchange (RMCE)—a rapidly-expanding toolbox for targeted genomic modifications. *Gene.* Feb. 15, 2013;515(1):1-27. doi: 10.1016/j.gene.2012.11.016. Epub Nov. 29, 2012.
- Turan et al., Recombinase-mediated cassette exchange (RMCE): traditional concepts and current challenges. *J Mol Biol.* Mar. 25, 2011;407(2):193-221. doi: 10.1016/j.jmb.2011.01.004. Epub Jan. 15, 2011.
- Turan et al., Site-specific recombinases: from tag-and-target- to tag-and-exchange-based genomic modifications. *Faseb J.* Dec. 2011;25(12):4088-107. doi: 10.1096/fj.11-186940. Epub Sep. 2, 2011. Review.
- Tycko et al., Pairwise library screen systematically interrogates *Staphylococcus aureus* Cas9 specificity in human cells. *bioRxiv.* doi: <https://doi.org/10.1101/269399> Posted Feb. 22, 2018.
- UniProt Consortium, UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* Mar. 16, 2018;46(5):2699. doi: 10.1093/nar/gky092.
- UniProt Submission; UniProt, Accession No. P01011. Last modified Jun. 11, 2014, version 2. 15 pages.
- UniProt Submission; UniProt, Accession No. P01011. Last modified Sep. 18, 2013, version 2. 15 pages.
- UniProt Submission; UniProt, Accession No. P04264. Last modified Jun. 11, 2014, version 6. 15 pages.
- UniProt Submission; UniProt, Accession No. P04275. Last modified Jul. 9, 2014, version 107. 29 pages.
- UniProtein A0A1V6. Dec. 11, 2019.
- UniProtKB Submission; Accession No. F0NH53. May 3, 2011. 4 pages.
- UniProtKB Submission; Accession No. F0NN87. May 3, 2011. 4 pages.
- UniProtKB Submission; Accession No. P0DOC6. No Author Listed., Oct. 5, 2016. 5 pages.
- UniProtKB Submission; Accession No. T0D7A2. Oct. 16, 2013. 10 pages.
- Urasaki et al., Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive

(56)

References Cited**OTHER PUBLICATIONS**

- sequence in the subterminal region essential for transposition. *Genetics*. Oct. 2006;174(2):639-49. doi: 10.1534/genetics.106.060244. Epub Sep. 7, 2006.
- Urnov et al., Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*. Sep. 2010;11(9):636-46. doi: 10.1038/nrg2842.
- Urnov et al., Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. Jun. 2, 2005;435(7042):646-51. Epub Apr. 3, 2005.
- Usman et al., Exploiting the chemical synthesis of RNA. *Trends Biochem Sci*. Sep. 1992;17(9):334-9. doi: 10.1016/0968-0004(92)90306-t.
- Vagner et al., Efficiency of homologous DNA recombination varies along the *Bacillus subtilis* chromosome. *J Bacteriol*. Sep. 1988;170(9):3978-82.
- Van Brunt et al., Genetically Encoded Azide Containing Amino Acid in Mammalian Cells Enables Site-Specific Antibody-Drug Conjugates Using Click Cycloaddition Chemistry. *Bioconjug Chem*. Nov. 18, 2015;26(11):2249-60. doi: 10.1021/acs.bioconjchem.5b00359. Epub Sep. 11, 2015.
- Van Brunt et al., Molecular Farming: Transgenic Animals as Bioreactors. *Biotechnology (Y)*. 1988;6(10):1149-1154. doi: 10.1038/nbt1088-1149.
- Van Duyne et al., Teaching Cre to follow directions. *Proc Natl Acad Sci U S A*. Jan. 6, 2009;106(1):4-5. doi: 10.1073/pnas.0811624106. Epub Dec. 31, 2008.
- Van Overbeek et al., DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Mol Cell*. Aug. 18, 2016;63(4):633-646. doi: 10.1016/j.molcel.2016.06.037. Epub Aug. 4, 2016.
- Van Swieten et al., A mutation in the fibroblast growth factor 14 gene is associated with autosomal dominant cerebellar ataxia [corrected]. *Am J Hum Genet*. Jan. 2003;72(1):191-9. Epub Dec. 13, 2002.
- Van Wijk et al., Identification of 51 novel exons of the Usher syndrome type 2A (USH2A) gene that encode multiple conserved functional domains and that are mutated in patients with Usher syndrome type II. *Am J Hum Genet*. Apr. 2004;74(4):738-44. doi: 10.1086/383096. Epub Mar. 10, 2004.
- Vanamee et al., FokI requires two specific DNA sites for cleavage. *J Mol Biol*. May 25, 2001;309(1):69-78.
- Varga et al., Progressive vascular smooth muscle cell defects in a mouse model of Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci U S A*. Feb. 28, 2006;103(9):3250-5. doi: 10.1073/pnas.0600012103. Epub Feb. 21, 2006.
- Vellore et al., A group II intron-type open reading frame from the thermophile *Bacillus* (*Geobacillus*) stearothermophilus encodes a heat-stable reverse transcriptase. *Appl Environ Microbiol*. Dec. 2004;70(12):7140-7. doi: 10.1128/AEM.70.12.7140-7.1407.2004.
- Venken et al., Genome-wide manipulations of *Drosophila melanogaster* with transposons, Flp recombinase, and Fc31 integrase. *Methods Mol Biol*. 2012;859:203-28. doi: 10.1007/978-1-61779-603-6_12.
- Verma, The reverse transcriptase. *Biochim Biophys Acta*. Mar. 21, 1977;473(1):1-38. doi: 10.1016/0304-419X(77)90005-1.
- Vigne et al., Third-generation adenovectors for gene therapy. *Restor Neurol Neurosci*. Jan. 1, 1995;8(1):35-6. doi: 10.3233/RNN-1995-81208.
- Vik et al., Endonuclease V cleaves at inosines in RNA. *Nat Commun*. 2013;4:2271. doi: 10.1038/ncomms3271.
- Vilenchik et al., Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *Proc Natl Acad Sci U S A*. Oct. 28, 2003;100(22):12871-6. doi: 10.1073/pnas.2135498100. Epub Oct. 17, 2003.
- Villiger et al., Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat Med*. Oct. 2018;24(10):1519-1525. doi: 10.1038/s41591-018-0209-1. Epub Oct. 8, 2018.
- Vitreschak et al., Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element. *RNA*. Sep. 2003;9(9):1084-97.
- Voigt et al., Rational evolutionary design: the theory of in vitro protein evolution. *Adv Protein Chem*. 2000;55:79-160.
- Vriend et al., Nick-initiated homologous recombination: Protecting the genome, one strand at a time. *DNA Repair (Amst)*. Feb. 2017;50:1-13. doi: 10.1016/j.dnarep.2016.12.005. Epub Dec. 29, 2016.
- Wacey et al., Disentangling the perturbational effects of amino acid substitutions in the DNA-binding domain of p53. *Hum Genet*. Jan. 1999;104(1):15-22.
- Wadia et al., Modulation of cellular function by TAT mediated transduction of full length proteins. *Curr Protein Pept Sci*. Apr. 2003;4(2):97-104.
- Wadia et al., Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat Med*. Mar. 2004;10(3):310-5. Epub Feb. 8, 2004.
- Wah et al., Structure of FokI has implications for DNA cleavage. *Proc Natl Acad Sci U S A*. Sep. 1, 1998;95(18):10564-9.
- Wals et al., Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Front Chem*. Apr. 1, 2014;2:15. doi: 10.3389/fchem.2014.00015. eCollection 2014.
- Wan et al., Material solutions for delivery of CRISPR/Cas-based genome editing tools: Current status and future outlook. *Materials Today*. Jun. 2019;26:40-66. doi: 10.1016/j.mattod.2018.12.003.
- Wang et al., CRISPR-Cas9 and CRISPR-Assisted Cytidine Deaminase Enable Precise and Efficient Genome Editing in *Klebsiella pneumoniae*. *Appl Environ Microbiol*. 2018;84(23):e01834-18. Published Nov. 15, 2018. doi: 10.1128/AEM.01834-18.
- Wang et al., AID upmutants isolated using a high-throughput screen highlight the immunity/cancer balance limiting DNA deaminase activity. *Nat Struct Mol Biol*. Jul. 2009;16(7):769-76. doi: 10.1038/nsmb.1623. Epub Jun. 21, 2009.
- Wang et al., Continuous directed evolutions of proteins with improved soluble expression. *Nature Chemical Biology*. Nat Publishing Group. Aug. 20, 2018; 14(10):972-980.
- Wang et al., CRISPR-Cas9 Targeting of PCSK9 in Human Hepatocytes In Vivo-Brief Report. *Arterioscler Thromb Vasc Biol*. May 2016;36(5):783-6. doi: 10.1161/ATVBAHA.116.307227. Epub Mar. 3, 2016.
- Wang et al., Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. *Proc Natl Acad Sci U S A*. Feb. 29, 2016. pii: 201520244. [Epub ahead of print].
- Wang et al., Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. *Cell Res*. Oct. 2017;27(1):1289-92. doi: 10.1038/cr.2017.111. Epub Aug. 29, 2017.
- Wang et al., Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc Natl Acad Sci U S A*. Nov. 30, 2004;101(48):16745-9. Epub Nov. 19, 2004.
- Wang et al., Expanding the genetic code. *Annu Rev Biophys Biomol Struct*. 2006;35:225-49. Review.
- Wang et al., Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. Jan. 3, 2014;343(6166):80-4. doi: 10.1126/science.1246981. Epub Dec. 12, 2013.
- Wang et al., Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes. *Biotechniques*. 2015;59,201-2;204:206-8.
- Wang et al., (6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell*. Jun. 4, 2015;161(6):1388-99. doi: 10.1016/j.cell.2015.05.014.
- Wang et al., N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*. Jan. 2, 2014;505(7481):117-20. doi: 10.1038/nature12730. Epub Nov. 27, 2013.
- Wang et al., Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature*. Oct. 8, 2009;461(7265):754-61. doi: 10.1038/nature08434.
- Wang et al., One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. May 9, 2013;153(4):910-8. doi: 10.1016/j.cell.2013.04.025. Epub May 2, 2013.
- Wang et al., Optimized paired-sgRNA/Cas9 cloning and expression cassette triggers high-efficiency multiplex genome editing in kiwi-fruit. *Plant Biotechnol J*. Aug. 2018;16(8):1424-1433. doi: 10.1111/pbi.12884. Epub Feb. 6, 2018.

(56)

References Cited**OTHER PUBLICATIONS**

- Wang et al., Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. Aug. 13, 2009;460(7257):894-8. Epub Jul. 26, 2009.
- Wang et al., Reading RNA methylation codes through methyl-specific binding proteins. *RNA Biol*. 2014;11(6):669-72. doi: 10.4161/rna.28829. Epub Apr. 24, 2014.
- Wang et al., Recombinase technology: applications and possibilities. *Plant Cell Rep*. Mar. 2011;30(3):267-85. doi: 10.1007/s00299-010-0938-1. Epub Oct. 24, 2010.
- Wang et al., Riboswitches that sense S-adenosylhomocysteine and activate genes involved in coenzyme recycling. *Mol Cell*. Mar. 28, 2008;29(6):691-702. doi: 10.1016/j.molcel.2008.01.012.
- Wang et al., *Staphylococcus aureus* protein SAUGI acts as a uracil-DNA glycosylase inhibitor. *Nucleic Acids Res*. Jan. 2014;42(2):1354-64. doi: 10.1093/nar/gkt964. Epub Oct. 22, 2013.
- Wang et al., Structural basis of (6)-adenosine methylation by the METTL3-METTL14 complex. *Nature*. Jun. 23, 2016;534(7608):575-8. doi: 10.1038/nature18298. Epub May 25, 2016.
- Wang et al., Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res*. Jul. 2012;22(7):1316-26. doi: 10.1101/gr.122879.111. Epub Mar. 20, 2012.
- Wang et al., Uracil-DNA glycosylase inhibitor gene of bacteriophage PBS2 encodes a binding protein specific for uracil-DNA glycosylase. *J Biol Chem*. Jan. 15, 1989;264(2):1163-71.
- Warren et al., A chimeric Cre recombinase with regulated directionality. *Proc Natl Acad Sci USA*. Nov. 25, 2008;105(47):18278-83. doi: 10.1073/pnas.0809949105. Epub Nov. 14, 2008.
- Warren et al., Mutations in the amino-terminal domain of lambda-integrase have differential effects on integrative and excisive recombination. *Mol Microbiol*. Feb. 2005;55(4):1104-12.
- Watowich, The erythropoietin receptor: molecular structure and hematopoietic signaling pathways. *J Investig Med*. Oct. 2011;59(7):1067-72. doi: 10.2310/JIM.0b013e31820fb28c.
- Waxman et al., Regulating excitability of peripheral afferents: emerging ion channel targets. *Nat Neurosci*. Feb. 2014;17(2):153-63. doi: 10.1038/nn.3602. Epub Jan. 28, 2014.
- Weber et al., Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One*. 2011;6(5):e19722. doi: 10.1371/journal.pone.0019722. Epub May 19, 2011.
- Weill et al., DNA polymerases in adaptive immunity. *Nat Rev Immunol*. Apr. 2008;8(4):302-12. doi: 10.1038/nri2281. Epub Mar. 14, 2008.
- Weinberg et al., New Classes of Self-Cleaving Ribozymes Revealed by Comparative Genomics Analysis. *Nat Chem Biol*. Aug. 2015;11(8):606-10. doi: 10.1038/nchembio.1846. Epub Jul. 13, 2015.
- Weinberg et al., The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches. *RNA*. May 2008;14(5):822-8. doi: 10.1261/rna.988608. Epub Mar. 27, 2008.
- Weinberger et al., Disease-causing mutations C277R and C277Y modify gating of human CIC-1 chloride channels in myotonia congenita. *J Physiol*. Aug. 1, 2012;590(Pt 15):3449-64. doi: 0.1113/j.physiol.2012.232785. Epub May 28, 2012.
- Weinert et al., Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science*. Apr. 19, 2019;364(6437):286-289. doi: 10.1126/science.aav9023. Epub Apr. 18, 2019.
- Weiss et al., Loss-of-function mutations in sodium channel Nav1.7 cause anosmia. *Nature*. Apr. 14, 2011;472(7342):186-90. doi: 10.1038/nature09975. Epub Mar. 23, 2011.
- Wen et al., Inclusion of a universal tetanus toxoid CD4(+) T cell epitope P2 significantly enhanced the immunogenicity of recombinant rotavirus 7VP8* subunit parenteral vaccines. *Vaccine*. Jul. 31, 2014;32(35):4420-4427. doi: 10.1016/j.vaccine.2014.06.060. Epub Jun. 21, 2014.
- West et al., Gene expression in adeno-associated virus vectors: the effects of chimeric mRNA structure, helper virus, and adenovirus VA1 RNA. *Virology*. Sep. 1987;160(1):38-47. doi: 10.1016/0042-6822(87)90041-9.
- Wharton et al., A new-specificity mutant of 434 repressor that defines an amino acid-base pair contact. *Nature*. Apr. 30-May 6, 1987;326(6116):888-91.
- Wharton et al., Changing the binding specificity of a repressor by redesigning an alpha-helix. *Nature*. Aug. 15-21, 1985;316(6029):601-5.
- Wheeler et al., The thermostability and specificity of ancient proteins. *Curr Opin Struct Biol*. Jun. 2016;38:37-43. doi: 10.1016/j.sbi.2016.05.015. Epub Jun. 9, 2016.
- Wiedenheft et al., RNA-guided genetic silencing systems in bacteria and archaea. *Nature*. Feb. 15, 2012;482(7385):331-8. doi: 10.1038/nature10886. Review.
- Wienert et al., KLF1 drives the expression of fetal hemoglobin in British HPFH. *Blood*. Aug. 10, 2017;130(6):803-807. doi: 10.1182/blood-2017-02-767400. Epub Jun. 28, 2017.
- Wijesinghe et al., Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G. *Nucleic Acids Res*. Oct. 2012;40(18):9206-17. doi: 10.1093/nar/gks685. Epub Jul. 13, 2012.
- Wijnker et al., Managing meiotic recombination in plant breeding. *Trends Plant Sci*. Dec. 2008;13(12):640-6. doi: 10.1016/j.tplants.2008.09.004. Epub Oct. 22, 2008.
- Williams et al., Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Comput Biol*. Jun. 23, 2006;2(6):e69. doi: 10.1371/journal.pcbi.0020069. Epub Jun. 23, 2006.
- Wills et al., Pseudoknot-dependent read-through of retroviral gag termination codons: importance of sequences in the spacer and loop 2. *EMBO J*. Sep. 1, 1994;13(17):4137-44. doi: 10.1002/j.1460-2075.1994.tb06731.x.
- Wilson et al., Assessing annotation transfer for genomics: quantifying the relations between protein sequence, structure and function through traditional and probabilistic scores. *J Mol Biol* 2000;297:233-49.
- Wilson et al., Formation of infectious hybrid virions with gibbon ape leukemia virus and human T-cell leukemia virus retroviral envelope glycoproteins and the gag and pol proteins of Moloney murine leukemia virus. *J Virol*. May 1989;63(5):2374-8. doi: 10.1128/JVI.63.5.2374-2378.1989.
- Wilson et al., In Vitro Selection of Functional Nucleic Acids. *Annu Rev Biochem*. 1999;68:611-47. doi: 10.1146/annurev.biochem.68.1.611.
- Wilson et al., Kinase dynamics. Using ancient protein kinases to unravel a modern cancer drug's mechanism. *Science*. Feb. 20, 2015;347(6224):882-6. doi: 10.1126/science.aaa1823.
- Winkler et al., An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A*. Dec. 10, 2002;99(25):15908-13. Epub Nov. 27, 2002.
- Winkler et al., Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*. Mar. 18, 2004;428(6980):281-6.
- Winkler et al., Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*. Oct. 31, 2002;419(6910):952-6. Epub Oct. 16, 2002.
- Winoto et al., A novel, inducible and T cell-specific enhancer located at the 3' end of the T cell receptor alpha locus. *EMBO J*. Mar. 1989;8(3):729-33.
- Winter et al., Drug Development. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science*. Jun. 19, 2015;348(6241):1376-81. doi: 10.1126/science.aab1433. Epub May 21, 2015.
- Winter et al., Targeted exon skipping with AAV-mediated split adenine base editors. *Cell Discov*. Aug. 20, 2019;5:41. doi: 10.1038/s41421-019-0109-7.
- Wold, Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem*. 1997;66:61-92. doi: 10.1146/annurev.biochem.66.1.61.
- Wolf et al., tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J*. Jul. 15, 2002;21(14):3841-51.
- Wolfe et al., Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. *J Mol Biol*. Feb. 5, 1999;285(5):1917-34.

(56)

References Cited**OTHER PUBLICATIONS**

- Wong et al., A statistical analysis of random mutagenesis methods used for directed protein evolution. *J Mol Biol.* Jan. 27, 2006;355(4):858-71. Epub Nov. 17, 2005.
- Wong et al., The Diversity Challenge in Directed Protein Evolution. *Comb Chem High Throughput Screen.* May 2006;9(4):271-88.
- Wood et al., A genetic system yields self-cleaving inteins for bioseparations. *Nat Biotechnol.* Sep. 1999;17(9):889-92. doi: 10.1038/12879.
- Wood et al., Targeted genome editing across species using ZFNs and TALENs. *Science.* Jul. 15, 2011;333(6040):307. doi: 10.1126/science.1207773. Epub Jun. 23, 2011.
- Woods et al., The phenotype of congenital insensitivity to pain due to the NaV1.9 variant p.L811P. *Eur J Hum Genet.* May 2015;23(5):561-3. doi: 10.1038/ejhg.2014.166. Epub Aug. 13, 2014.
- Wright et al., Continuous in vitro evolution of catalytic function. *Science.* Apr. 25, 1997;276(5312):614-7.
- Wright et al., Rational design of a split-Cas9 enzyme complex. *Proc Natl Acad Sci U S A.* Mar. 10, 2015;112(10):2984-9. doi: 10.1073/pnas.1501698112. Epub Feb. 23, 2015.
- Wu et al., Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell.* Dec. 5, 2013;13(6):659-62. doi: 10.1016/j.stem.2013.10.016.
- Wu et al., Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol.* Jul. 2014;32(7):670-6. doi: 10.1038/nbt.2889. Epub Apr. 20, 2014.
- Wu et al., Human single-stranded DNA binding proteins: guardians of genome stability. *Acta Biochim Biophys Sin (Shanghai).* Jul. 2016;48(7):671-7. doi: 10.1093/abbs/gmw044. Epub May 23, 2016.
- Wu et al., Protein trans-splicing and functional mini-inteins of a cyanobacterial dnaB intein. *Biochim Biophys Acta.* Sep. 8, 1998;1387(1-2):422-32. doi: 10.1016/s0167-4838(98)00157-5.
- Wu et al., Protein trans-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803. *Proc Natl Acad Sci U S A.* Aug. 4, 1998;95(16):9226-31. doi: 10.1073/pnas.95.16.9226.
- Wu et al., Readers, writers and erasers of N6-methylated adenosine modification. *Curr Opin Struct Biol.* Dec. 2017;47:67-76. doi: 10.1016/j.sbi.2017.05.011. Epub Jun. 16, 2017.
- Xiang et al., RNA m6A methylation regulates the ultraviolet-induced DNA damage response. *Nature.* Mar. 23, 2017;543(7646):573-576. doi: 10.1038/nature21671. Epub Mar. 15, 2017.
- Xiao et al., Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew Chem Int Ed Engl.* Dec. 23, 2013;52(52):14080-3. doi: 10.1002/anie.201308137. Epub Nov. 8, 2013.
- Xiao et al., Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell.* Feb. 18, 2016;61(4):507-519. doi: 10.1016/j.molcel.2016.01.012. Epub Feb. 11, 2016.
- Xie et al., Adjusting the attB site in donor plasmid improves the efficiency of ?C31 integrase system. *DNA Cell Biol.* Jul. 2012;31(7):1335-40. doi: 10.1089/dna.2011.1590. Epub Apr. 10, 2012.
- Xiong et al., Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* Oct. 1990;9(10):3353-62.
- Xu et al., Accuracy and efficiency define Bxb1 integrase as the best of fifteen candidate serine recombinases for the integration of DNA into the human genome. *BMC Biotechnol.* Oct. 20, 2013;13:87. doi: 10.1186/1472-6750-13-87.
- Xu et al., Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc Natl Acad Sci U S A.* Jan. 19, 1999;96(2):388-93. doi: 10.1073/pnas.96.2.388.
- Xu et al., Multiplex nucleotide editing by high-fidelity Cas9 variants with improved efficiency in rice. *BMC Plant Biol.* 2019;19(1):511. Published Nov. 2, 2019. doi: 10.1186/s12870-019-2131-1. Includes supplementary data and materials.
- Xu et al., Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. *EMBO J.* Dec. 1, 1994;13(23):5517-22.
- Xu et al., PTMD: A Database of Human Disease-associated Post-translational Modifications. *Genomics Proteomics Bioinformatics.* Aug. 2018;16(4):244-251. doi: 10.1016/j.gpb.2018.06.004. Epub Sep. 21, 2018.
- Xu et al., Sequence determinants of improved CRISPR sgRNA design. *Genome Res.* Aug. 2015;25(8):1147-57. doi: 10.1101/gr.191452.115. Epub Jun. 10, 2015.
- Xu et al., Structures of human ALKBH5 demethylase reveal a unique binding mode for specific single-stranded N6-methyladenosine RNA demethylation. *J Biol Chem.* Jun. 20, 2014;289(25):17299-311. doi: 10.1074/jbc.M114.550350. Epub Apr. 28, 2014.
- Xu et al., The mechanism of protein splicing and its modulation by mutation. *EMBO J.* Oct. 1, 1996;15(19):5146-53.
- Yahata et al., Unified, Efficient, and Scalable Synthesis of Halichondrins: Zirconium/Nickel-Mediated One-Pot Ketone Synthesis as the Final Coupling Reaction. *Angew Chem Int Ed Engl.* Aug. 28, 2017;56(36):10796-10800. doi: 10.1002/anie.201705523. Epub Jul. 28, 2017.
- Yamada et al., Crystal Structure of the Minimal Cas9 from *Campylobacter jejuni* Reveals the Molecular Diversity in the CRISPR-Cas9 Systems. *Mol Cell.* Mar. 16, 2017;65(6):P1109-1121. doi: org/10.1016/j.molcel.2017.02.007.
- Yamamoto et al., The pros and cons of inducible transgenic technology: a review. *Neurobiol Dis.* Dec. 2001;8(6):923-32.
- Yamamoto et al., Virological and immunological bases for HIV-1 vaccine design. *Uirusu* 2007;57(2):133-139. https://doi.org/10.2222/jsv.57.133.
- Yamano et al., Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell.* May 5, 2016;165(4):949-62 and Supplemental Info. doi: 10.1016/j.cell.2016.04.003. Epub Apr. 21, 2016.
- Yamano et al., Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell.* May 5, 2016;165(4):949-62. doi: 10.1016/j.cell.2016.04.003. Epub Apr. 21, 2016.
- Yamazaki et al., Segmental Isotope Labeling for Protein NMR Using Peptide Splicing. *J Am Chem Soc.* May 22, 1998;120(22):5591-2. https://doi.org/10.1021/ja9807760.
- Yan et al., Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein. *Mol Cell.* Apr. 19, 2018;70(2):327-339.e5. doi: 10.1016/j.molcel.2018.02.028. Epub Mar. 15, 2018.
- Yan et al., Functionally diverse type V CRISPR-Cas systems. *Science.* Jan. 4, 2019;363(6422):88-91. doi: 10.1126/science.aav7271. Epub Dec. 6, 2018.
- Yan et al., Highly Efficient A•T to G•C Base Editing by Cas9n-Guided tRNA Adenosine Deaminase in Rice. *Mol Plant.* Apr. 2, 2018;11(4):631-634. doi: 10.1016/j.molp.2018.02.008. Epub Feb. 22, 2018.
- Yang et al., APOBEC: From mutator to editor. *J Genet Genomics.* Sep. 20, 2017;44(9):423-437. doi: 10.1016/j.jgg.2017.04.009. Epub Aug. 7, 2017.
- Yang et al., Construction of an integration-proficient vector based on the site-specific recombination mechanism of enterococcal temperate phage phiFC1. *J Bacteriol.* Apr. 2002;184(7):1859-64. doi: 10.1128/jb.184.7.1859-1864.2002.
- Yang et al., Engineering and optimising deaminase fusions for genome editing. *Nat Commun.* Nov. 2, 2016;7:13330. doi: 10.1038/ncomms13330.
- Yang et al., Genome editing with targeted deaminases. *BioRxiv.* Preprint. First posted online Jul. 28, 2016.
- Yang et al., Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science.* Nov. 27, 2015;350(6264):1101-4. doi: 10.1126/science.aad1191. Epub Oct. 11, 2015.
- Yang et al., Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of TadA deaminase with Cas9 variants. *Protein Cell.* Sep. 2018;9(9):814-819. doi: 10.1007/s13238-018-0568-x.
- Yang et al., Mutations in SCN9A, encoding a sodium channel alpha subunit, in patients with primary erythermalgia. *J Med Genet.* Mar. 2004;41(3):171-4. doi: 10.1136/jmg.2003.012153.
- Yang et al., New CRISPR-Cas systems discovered. *Cell Res.* Mar. 2017;27(3):313-314. doi: 10.1038/cr.2017.21. Epub Feb. 21, 2017.

(56)

References Cited**OTHER PUBLICATIONS**

- Yang et al., One Prime for All Editing. *Cell*. Dec. 12, 2019;179(7):1448-1450. doi: 10.1016/j.cell.2019.11.030.
- Yang et al., One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. Sep. 12, 2013;154(6):1370-9. doi: 10.1016/j.cell.2013.08.022. Epub Aug. 29, 2013.
- Yang et al., PAM-dependent Target DNA Recognition and Cleavage by C2C1 CRISPR-Cas endonuclease. *Cell*. Dec. 2016;167(7):1814-28.
- Yang et al., Permanent genetic memory with >1-byte capacity. *Nat Methods*. Dec. 2014;11(12):1261-6. doi: 10.1038/nmeth.3147. Epub Oct. 26, 2014.
- Yang et al., Preparation of RNA-directed DNA polymerase from spleens of Balb-c mice infected with Rauscher leukemia virus. *Biochem Biophys Res Commun*. Apr. 28, 1972;47(2):505-11. doi: 10.1016/0006-291x(72)90743-7.
- Yang et al., Small-molecule control of insulin and PDGF receptor signaling and the role of membrane attachment. *Curr Biol*. Jan. 1, 1998;8(1):11-8. doi: 10.1016/s0960-9822(98)70015-6.
- Yang, Development of Human Genome Editing Tools for the Study of Genetic Variations and Gene Therapies. Doctoral Dissertation. Harvard University. 2013. Accessible via nrs.harvard.edu/urn-3:HUL.InstRepos:11181072. 277 pages.
- Yang, Nucleases: diversity of structure, function and mechanism. *Q Rev Biophys*. Feb. 2011;44(1):1-93. doi: 10.1017/S0033583510000181. Epub Sep. 21, 2010.
- Yang, PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. Aug. 2007;24(8):1586-91. doi: 10.1093/molbev/msm088. Epub May 4, 2007.
- Yanover et al., Extensive protein and DNA backbone sampling improves structure-based specificity prediction for C2H2 zinc fingers. *Nucleic Acids Res*. Jun. 2011;39(11):4564-76. doi: 10.1093/nar/gkr048. Epub Feb. 22, 2011.
- Yasui et al., MisCoding Properties of 2'-Deoxyinosine, a Nitric Oxide-Derived DNA Adduct, during Translesion Synthesis Catalyzed by Human DNA Polymerases. *J Molec Biol*. Apr. 4, 2008;377(4):1015-23.
- Yasui, Alternative excision repair pathways. *Cold Spring Harb Perspect Biol*. Jun. 1, 2013;5(6):a012617. doi: 10.1101/cshperspect.a012617.
- Yasukawa et al., Characterization of Moloney murine leukaemia virus/avian myeloblastosis virus chimeric reverse transcriptases. *J Biochem*. Mar. 2009;145(3):315-24. doi: 10.1093/jb/mvn166. Epub Dec. 6, 2008.
- Yazaki et al., Hereditary systemic amyloidosis associated with a new apolipoprotein AII stop codon mutation Stop78Arg. *Kidney Int*. Jul. 2003;64(1):11-6.
- Yeh et al., In vivo base editing of post-mitotic sensory cells. *Nat Commun*. Jun. 5, 2018;9(1):2184. doi: 10.1038/s41467-018-04580-3.
- Yin et al., Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol*. Jun. 2014;32(6):551-3. doi: 10.1038/nbt.2884. Epub Mar. 30, 2014.
- Yokoe et al., Spatial dynamics of GFP-tagged proteins investigated by local fluorescence enhancement. *Nat Biotechnol*. Oct. 1996;14(10):1252-6. doi: 10.1038/nbt1096-1252.
- Young et al., Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J Biol Chem*. Apr. 9, 2010;285(15):11039-44. doi: 10.1074/jbc.R109.091306. Epub Feb. 10, 2010.
- Yu et al., Circular permutation: a different way to engineer enzyme structure and function. *Trends Biotechnol*. Jan. 2011;29(1):18-25. doi: 10.1016/j.tibtech.2010.10.004. Epub Nov. 17, 2010.
- Yu et al., Liposome-mediated in vivo ELA gene transfer suppressed dissemination of ovarian cancer cells that overexpress HER-2/neu. *Oncogene*. Oct. 5, 1995;11(7):1383-8.
- Yu et al., Progress towards gene therapy for HIV infection. *Gene Ther*. Jan. 1994;1(1):13-26.
- Yu et al., Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell*. Feb. 5, 2015;16(2):142-7. doi: 10.1016/j.stem.2015.01.003.
- Yu et al., Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res*. Sep. 2010;38(17):5706-17. doi: 10.1093/nar/gkq379. Epub May 11, 2010.
- Yuan et al., Laboratory-directed protein evolution. *Microbiol Mol Biol Rev*. 2005; 69(3):373-92. PMID: 16148303.
- Yuan et al., Tetrameric structure of a serine integrase catalytic domain. *Structure*. Aug. 6, 2008;16(8):1275-86. doi: 10.1016/j.str.2008.04.018.
- Yuen et al., Control of transcription factor activity and osteoblast differentiation in mammalian cells using an evolved small-molecule-dependent intein. *J Am Chem Soc*. Jul. 12, 2006;128(27):8939-46.
- Zakas et al., Enhancing the pharmaceutical properties of protein drugs by ancestral sequence reconstruction. *Nat Biotechnol*. Jan. 2017;35(1):35-37. doi: 10.1038/nbt.3677. Epub Sep. 26, 2016.
- Zalatan et al., Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*. Jan. 15, 2015;160(1-2):339-50. doi: 10.1016/j.cell.2014.11.052. Epub Dec. 18, 2014.
- Zelphati et al., Intracellular delivery of proteins with a new lipid-mediated delivery system. *J Biol Chem*. Sep. 14, 2001;276(37):35103-10. Epub Jul. 10, 2001.
- Zeng et al., Correction of the Marfan Syndrome Pathogenic FBN1 Mutation by Base Editing in Human Cells and Heterozygous Embryos. *Mol Ther*. Nov. 7, 2018;26(11):2631-2637. doi: 10.1016/j.ymthe.2018.08.007. Epub Aug. 14, 2018.
- Zetsche et al., A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat Biotechnol*. Feb. 2015;33(2):139-42. doi: 10.1038/nbt.3149.
- Zetsche et al., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. Oct. 22, 2015;163(3):759-71 and Supplemental Info. doi: 10.1016/j.cell.2015.09.038. Epub Sep. 25, 2015.
- Zetsche et al., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. Oct. 22, 2015;163(3):759-71. doi: 10.1016/j.cell.2015.09.038. Epub Sep. 25, 2015.
- Zettler et al., The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett*. Mar. 4, 2009;583(5):909-14. doi: 10.1016/j.febslet.2009.02.003. Epub Feb. 10, 2009.
- Zhang et al., II-Clamp-mediated cysteine conjugation. *Nat Chem*. Feb. 2016;8(2):120-8. doi: 10.1038/nchem.2413. Epub Dec. 21, 2015.
- Zhang et al., A new strategy for the site-specific modification of proteins in vivo. *Biochemistry*. Jun. 10, 2003;42(22):6735-46.
- Zhang et al., Circular intronic long noncoding RNAs. *Mol Cell*. Sep. 26, 2013;51(6):792-806. doi: 10.1016/j.molcel.2013.08.017. Epub Sep. 12, 2013.
- Zhang et al., Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells. *Sci Rep*. Jun. 2014;4:5405.
- Zhang et al., Conditional gene manipulation: Cre-ating a new biological era. *J Zhejiang Univ Sci B*. Jul. 2012;13(7):511-24. doi: 10.1631/jzus.B1200042. Review.
- Zhang et al., Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet*. 2009;10:451-81. doi: 10.1146/annurev.genom.9.081307.164217.
- Zhang et al., CRISPR/Cas9 for genome editing: progress, implications and challenges. *Hum Mol Genet*. Sep. 15, 2014;23(R1):R40-6. doi: 10.1093/hmg/ddu125. Epub Mar. 20, 2014.
- Zhang et al., Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*. Feb. 2011;29(2):149-53. doi: 10.1038/nbt.1775. Epub Jan. 19, 2011.
- Zhang et al., Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol*. Nov. 2003;50(4):1111-24. doi: 10.1046/j.1365-2958.2003.03734.x.
- Zhang et al., Myoediting: Toward Prevention of Muscular Dystrophy by Therapeutic Genome Editing. *Physiol Rev*. Jul. 1, 2018;98(3):1205-1240. doi: 10.1152/physrev.00046.2017.

(56)

References Cited**OTHER PUBLICATIONS**

- Zhang et al., Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat Commun.* Jul. 25, 2017;8(1):118. doi: 10.1038/s41467-017-00175-6.
- Zhang et al., Reversible RNA Modification N1-methyladenosine (m¹A) in mRNA and tRNA. *Genomics Proteomics Bioinformatics.* Jun. 2018;16(3):155-161. doi: 10.1016/j.gpb.2018.03.003. Epub Jun. 14, 2018.
- Zhang et al., Ribozymes and Riboswitches: Modulation of RNA Function by Small Molecules. *Biochemistry.* Nov. 2, 2010;49(43):9123-31. doi: 10.1021/bi1012645.
- Zhang et al., Stabilized plasmid-lipid particles for regional gene therapy: formulation and transfection properties. *Gene Ther.* Aug. 1999;6(8):1438-47.
- Zhao et al., An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. *RNA.* Feb. 2018;24(2):183-195. doi: 10.1261/rna.063479.117. Epub Nov. 6, 2017.
- Zhao et al., Crystal structures of a group II intron maturase reveal a missing link in spliceosome evolution. *Nat Struct Mol Biol.* Jun. 2016;23(6):558-65. doi: 10.1038/nsmb.3224. Epub May 2, 2016.
- Zhao et al., Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol.* Jan. 2017;18(1):31-42. doi: 10.1038/nrm.2016.132. Epub Nov. 3, 2016.
- Zheng et al., ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell.* Jan. 10, 2013;49(1):18-29. doi: 10.1016/j.molcel.2012.10.015. Epub Nov. 21, 2012.
- Zheng et al., DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. *Nucleic Acids Res.* Apr. 7, 2017;45(6):3369-3377. doi: 10.1093/nar/gkx050.
- Zheng et al., Highly efficient base editing in bacteria using a Cas9-cytidine deaminase fusion. *Commun Biol.* Apr. 19, 2018;1:32. doi: 10.1038/s42003-018-0035-5.
- Zheng et al., Structural basis for the complete resistance of the human prion protein mutant G127V to prion disease. *Sci Rep.* Sep. 4, 2018;8(1):13211. doi: 10.1038/s41598-018-31394-6.
- Zhong et al., Rational Design of Aptazyme Riboswitches for Efficient Control of Gene Expression in Mammalian Cells. *Elife.* Nov. 2, 2016;5:e18858. doi: 10.7554/elife.18858.
- Zhou et al., Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature.* Oct. 22, 2015;526(7574):591-4. doi: 10.1038/nature15377. Epub Oct. 12, 2015.
- Zhou et al., GISSD: Group I Intron Sequence and Structure Database. *Nucleic Acids Res.* Jan. 2008;36(Database issue):D31-7. doi: 10.1093/nar/gkm766. Epub Oct. 16, 2007.
- Zhou et al., Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature.* Jul. 2019;571(7764):275-278. doi: 10.1038/s41586-019-1314-0. Epub Jun. 10, 2019.
- Zhou et al., Protective V127 prion variant prevents prion disease by interrupting the formation of dimer and fibril from molecular dynamics simulations. *Sci Rep.* Feb. 24, 2016;6:21804. doi: 10.1038/srep21804.
- Zhou et al., Seamless Genetic Conversion of SMN2 to SMN1 via CRISPR/Cpf1 and Single-Stranded Oligodeoxynucleotides in Spinal Muscular Atrophy Patient-Specific Induced Pluripotent Stem Cells. *Hum Gene Ther.* Nov. 2018;29(11):1252-1263. doi: 10.1089/hum.2017.255. Epub May 9, 2018.
- Zielenski, Genotype and phenotype in cystic fibrosis. *Respiration.* 2000;67(2):117-33. doi: 10.1159/000029497.
- Zimmerly et al., An Unexplored Diversity of Reverse Transcriptases in Bacteria. *Microbiol Spectr.* Apr. 2015;3(2):MDNA3-0058-2014. doi: 10.1128/microbiolspec.MDNA3-0058-2014.
- Zimmerly et al., Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell.* Aug. 25, 1995;82(4):545-54. doi: 10.1016/0092-8674(95)90027-6.
- Zimmermann et al., Molecular interactions and metal binding in the theophylline-binding core of an RNA aptamer. *RNA.* May 2000;6(5):659-67.
- Zolotukhin et al., Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors. *Methods.* Oct. 2002;28(2):158-67. doi: 10.1016/s1046-2023(02)00220-7.
- Zong et al., Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat Biotechnol.* May 2017;35(5):438-440. doi: 10.1038/nbt.3811. Epub Feb. 27, 2017.
- Zorko et al., Cell-penetrating peptides: mechanism and kinetics of cargo delivery. *Adv Drug Deliv Rev.* Feb. 28, 2005;57(4):529-45. Epub Jan. 22, 2005.
- Zou et al., Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell.* Jul. 2, 2009;5(1):97-110. doi: 10.1016/j.stem.2009.05.023. Epub Jun. 18, 2009.
- Zufferey et al., Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J Virol.* Apr. 1999;73(4):2886-92. doi: 10.1128/JVI.73.4.2886-2892.1999.
- Zuker et al., Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* Jan. 10, 1981;9(1):133-48. doi: 10.1093/nar/9.1.133.
- Zuo et al., Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science.* Apr. 19, 2019;364(6437):289-292. doi: 10.1126/science.aav9973. Epub Feb. 28, 2019.
- Zuris et al., Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol.* 2015;33:73-80.
- [No Author Listed], "Lambda DNA" from Catalog & Technical Reference. New England Biolabs Inc. 2002/2003. pp. 133 and 270-273.
- [No Author Listed], Gag-Pol polyprotein. UniProtKB/Swiss-Prot No. P03355. Sep. 18, 2019. 18 pages.
- [No Author Listed], *Homo sapiens* signal transducer and activator of transcription 3 (STAT3), transcript variant 1, mRNA. NCBI Ref Seq No. NM_139276.2. Retrieved from https://www.ncbi.nlm.nih.gov/nucleotide/nm_139276.2. Feb. 26, 2020. 8 pages.
- Ai et al., C-terminal Loop Mutations Determine Folding and Secretion Properties of PCSK9. *iMedPub J: Biochem Mol Biol J.* Nov. 5, 2016;2(3):17. doi: 10.21767/2471-8084.100026. 12 pages.
- Andersen et al., High frequency of T cells specific for cryptic epitopes in melanoma patients. *Oncimmunology.* Jul. 1, 2013;2(7):e25374. doi: 10.4161/onci.25374. 7 pages.
- Anzalone et al., Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature.* Dec. 2019;576(7785):149-157 and Suppl. Info. doi: 10.1038/s41586-019-1711-4. Epub Oct. 21, 2019. 72 pages.
- Baba et al., Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol.* 2006;2:2006.0008. doi: 10.1038/msb4100050. Epub Feb. 21, 2006.
- Bass, B.L., RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem.* 2002;71:817-46. doi: 10.1146/annurev.biochem.71.110601.135501. Epub Nov. 9, 2001.
- Bibikova et al., Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics.* Jul. 2002;161(3):1169-75. doi: 10.1093/genetics/161.3.1169.
- Blaauw et al., SMN1 gene duplications are associated with sporadic ALS. *Neurology.* Mar. 13, 2012;78(11):776-80. doi: 10.1212/WNL.0b013e318249f697. Epub Feb. 8, 2012.
- Bothmer et al., Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. *Nat Commun.* Jan. 9, 2017;8:13905. doi: 10.1038/ncomms13905.
- Brutlag et al., Improved sensitivity of biological sequence database searches. *Comput Appl Biosci.* Jul. 1990;6(3):237-45. doi: 10.1093/bioinformatics/6.3.237.
- Buonaguro et al., Translating tumor antigens into cancer vaccines. *Clin Vaccine Immunol.* Jan. 2011;18(1):23-34. doi: 10.1128/CVI.00286-10. Epub Nov. 3, 2010.
- Campos-Perez et al., DNA fusion vaccine designs to induce tumor-lytic CD8+ T-cell attack via the immunodominant cysteine-containing epitope of NY-ESO 1. *Int J Cancer.* Sep. 15, 2013;133(6):1400-7. doi: 10.1002/ijc.28156. Epub Apr. 11, 2013.

(56)

References Cited**OTHER PUBLICATIONS**

- Canny et al., Inhibition of 53BP1 Favors Homology-Dependent DNA Repair and Increases CRISPR-Cas9 Genome-Editing Efficiency. *Nat Biotechnol.* Jan. 2018;36(1):95-102. doi: 10.1038/nbt.4021. Epub Nov. 27, 2017.
- Cao et al., Rapamycin reverses cellular phenotypes and enhances mutant protein clearance in Hutchinson-Gilford progeria syndrome cells. *Sci Transl Med.* Jun. 29, 2011;3(89):89ra58. doi: 10.1126/scitranslmed.3002346.
- Carlier et al., Genome Sequence of Burkholderia cenocepacia H111, a Cystic Fibrosis Airway Isolate. *Genome Announc.* Apr. 10, 2014;2(2):e00298-14. doi: 10.1128/genomeA.00298-14.
- Cartegni et al., Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *Am J Hum Genet.* Jan. 2006;78(1):63-77. doi: 10.1086/498853. Epub Nov. 16, 2005.
- Chang et al., Degradation of survival motor neuron (SMN) protein is mediated via the ubiquitin/proteasome pathway. *Neurochem Int.* Dec. 2004;45(7):1107-12. doi: 10.1016/j.neuint.2004.04.005.
- Chatterjee et al., Robust Genome Editing of Single-Base PAM Targets; with Engineered ScCas9 Variants. *bioRxiv.* doi: 10.1101/620351. Posted Apr. 26, 2019.
- Cheng et al., [Cloning, expression and activity identification of human innate immune protein apolipoprotein B mRNA editing enzyme catalytic subunit 3A(APOBEC3A)]. Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi. Chinese Journal of Cellular and Molecular Immunology, Feb. 2017;33(2):179-84. Chinese.
- Cho et al., A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes Dev.* Mar. 1, 2010;24(5):438-42. doi: 10.1101/gad.1884910.
- Chowell et al., TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. *Proc Natl Acad Sci U S A.* Apr. 7, 2015;112(14):E1754-62. doi: 10.1073/pnas.1500973112. Epub Mar. 23, 2015.
- Corcia et al., The importance of the SMN genes in the genetics of sporadic ALS. *Amyotroph Lateral Scler.* Oct.-Dec. 2009;10(5-6):436-40. doi: 10.3109/17482960902759162.
- Corti et al., Genetic correction of human induced pluripotent stem cells from patients with spinal muscular atrophy. *Sci Transl Med.* Dec. 19, 2012;4(165):165ra162. doi: 10.1126/scitranslmed.3004108.
- Cuccharini et al., Enhanced expression of the central survival of motor neuron (SMN) protein during the pathogenesis of osteoarthritis. *J Cell Mol Med.* Jan. 2014;18(1):115-24. doi: 10.1111/jcmm.12170. Epub Nov. 17, 2013.
- Davis et al., Assaying Repair at DNA Nicks. *Methods Enzymol.* 2018;601:71-89. doi: 10.1016/bs.mie.2017.12.001. Epub Feb. 1, 2018.
- Davis et al., Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proc Natl Acad Sci U S A.* Mar. 11, 2014;111(10):E924-32. doi: 10.1073/pnas.1400236111. Epub Feb. 20, 2014.
- Davis et al., Two Distinct Pathways Support Gene Correction by Single-Stranded Donors at DNA Nicks. *Cell Rep.* Nov. 8, 2016;17(7):1872-1881. doi: 10.1016/j.celrep.2016.10.049.
- De Sandre-Giovannoli et al., Lamin a truncation in Hutchinson-Gilford progeria. *Science.* Jun. 27, 2003;300(5628):2055. doi: 10.1126/science.1084125. Epub Apr. 17, 2003.
- Dickinson et al., A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. *Nat Commun.* Oct. 30, 2014;5:5352. doi: 10.1038/ncomms6352.
- Drenth et al., Mutations in sodium-channel gene SCN9A cause a spectrum of human genetic pain disorders. *J Clin Invest.* Dec. 2007;117(12):3603-9. doi: 10.1172/JCI33297.
- Dugar et al., CRISPR RNA-Dependent Binding and Cleavage of Endogenous RNAs by the *Campylobacter jejuni* Cas9. *Mol Cell.* Mar. 1, 2018;69(5):893-905.e7. doi: 10.1016/j.molcel.2018.01.032.
- D'Ydewalle et al., The Antisense Transcript SMN-AS1 Regulates SMN Expression and Is a Novel Therapeutic Target for Spinal Muscular Atrophy. *Neuron.* Jan. 4, 2017;93(1):66-79 and Supplemental Information. doi: 10.1016/j.neuron.2016.11.033. Epub Dec. 22, 2016.
- Eisenberg et al., A-to-I RNA editing—immune protector and transcriptome diversifier. *Nat Rev Genet.* Aug. 2018;19(8):473-490. doi: 10.1038/s41576-018-0006-1.
- Ekstrand et al., Frequent alterations of the PI3K/AKT/mTOR pathways in hereditary nonpolyposis colorectal cancer. *Fam Cancer.* Jun. 2010;9(2):125-9. doi: 10.1007/s10689-009-9293-1.
- Entin-Meer et al., The role of phenylalanine-119 of the reverse transcriptase of mouse mammary tumour virus in DNA synthesis, ribose selection and drug resistance. *Biochem J.* Oct. 15, 2002;367(Pt 2):381-91. doi: 10.1042/BJ20020712.
- Friedman, J. H., Greedy function approximation: A gradient boosting machine. *Ann. Statist.* Oct. 2001;29(5):1189-232. doi: 10.1214/aos/1013203451.
- GenBank Submission; NIH/NCBI, Accession No. NC_000001.11. Gregory et al., Jun. 6, 2016. 3 pages.
- GenBank Submission; NIH/NCBI, Accession No. NG_008692.2. McClintock et al., Aug. 27, 2018. 33 pages.
- GenBank Submission; NIH/NCBI, Accession No. NM_206933.2. Khalaleh et al., Sep. 16, 2018. 12 pages.
- GenBank Submission; NIH/NCBI, Accession No. NP_001075493. 1. Schiaffella et al., Jun. 24, 2018. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. NP_001157741. 1. Zeng et al., Sep. 17, 2018. 3 pages.
- GenBank Submission; NIH/NCBI, Accession No. NP_001157742. 1. Zeng et al., Oct. 21, 2018. 3 pages.
- GenBank Submission; NIH/NCBI, Accession No. NP_033040.2. Liu et al., Jun. 23, 2018. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. NP_996816.2. Fu et al., Sep. 22, 2019. 9 pages.
- GenBank Submission; NIH/NCBI, Accession No. XP_003314669. 1. No Author Listed, Mar. 20, 2018. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. XP_026671085. 1. No Author Listed, Oct. 17, 2018. 1 page.
- Grati et al., Localization of PDZD7 to the stereocilia ankle-link associates this scaffolding protein with the Usher syndrome protein network. *J Neurosci.* Oct. 10, 2012;32(41):14288-93. doi: 10.1523/JNEUROSCI.3071-12.2012.
- Gutschner et al., Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep.* Feb. 16, 2016;14(6):1555-1566. doi: 10.1016/j.celrep.2016.01.019. Epub Feb. 4, 2016.
- Hagen et al., A high rate of polymerization during synthesis of mouse mammary tumor virus DNA alleviates hypermutation by APOBEC3 proteins. *PLoS Pathog.* Feb. 15, 2019;15(2):e1007533. doi: 10.1371/journal.ppat.1007533.
- Harrington et al., Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science.* Nov. 16, 2018;362(6416):839-842. doi: 10.1126/science.aav4294. Epub Oct. 18, 2018.
- Hart et al., High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell.* Dec. 3, 2015;163(6):1515-26. doi: 10.1016/j.cell.2015.11.015. Epub Nov. 25, 2015.
- Hawley-Nelson et al., Transfection of Cultured Eukaryotic Cells Using Cationic Lipid Reagents. *Curr Prot Mol Biol.* Jan. 2008;9,4. 1-9.4.17. doi: 10.1021/0471142727.mb0904s81. 17 pages.
- Hendel et al., Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol.* Sep. 2015;33(9):985-989. doi: 10.1038/nbt.3290. Epub Jun. 29, 2015. Author Manuscript. 14 pages.
- Heyer et al., Regulation of homologous recombination in eukaryotes. *Annu Rev Genet.* 2010;44:113-39. doi: 10.1146/annurev-genet-051710-150955. Author Manuscript. 33 pages.
- Huang et al., Gain-of-function mutations in sodium channel Na(v)1.9 in painful neuropathy. *Brain.* Jun. 2014;137(Pt 6):1627-42. doi: 10.1093/brain/awu079. Epub Apr. 27, 2014.
- Ishizuka et al., Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. *Nature.* Jan. 2019;565(7737):43-48. doi: 10.1038/s41586-018-0768-9. Epub Dec. 17, 2018.

(56)

References Cited**OTHER PUBLICATIONS**

- Iyama et al., DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair (Amst)*. Aug. 2013;12(8):620-36. doi: 10.1016/j.dnarep.2013.04.015. Epub May 16, 2013.
- Jakimo et al., A Cas9 with Complete PAM Recognition for Adenine Dinucleotides. *bioRxiv* preprint. Sep. 27, 2018. doi.org/10.1101/429654. 29 pages.
- Kan et al., Mechanisms of precise genome editing using oligonucleotide donors. *Genome Res*. Jul. 2017;27(7):1099-1111. doi: 10.1101/gr.214775.116. Epub Mar. 29, 2017.
- Kim et al., Adenine base editors catalyze cytosine conversions in human cells. *Nat Biotechnol*. Oct. 2019;37(10):1145-1148. doi: 10.1038/s41587-019-0254-4. Epub Sep. 23, 2019.
- Kim et al., RAD51 mutants cause replication defects and chromosomal instability. *Mol Cell Biol*. Sep. 2012;32(18):3663-80. doi: 10.1128/MCB.00406-12. Epub Jul. 9, 2012.
- Knott et al., CRISPR-Cas guides the future of genetic engineering. *Science*. Aug. 31, 2018;361(6405):866-869. doi: 10.1126/science.aat5011.
- Konishi et al., Amino acid substitutions away from the RNase H catalytic site increase the thermal stability of Moloney murine leukemia virus reverse transcriptase through RNase H inactivation. *Biochem Biophys Res Commun*. Nov. 14, 2014;454(2):269-74. doi: 10.1016/j.bbrc.2014.10.044. Epub Oct. 17, 2014.
- Kreiter et al., Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature*. Apr. 30, 2015;520(7549):692-6. doi: 10.1038/nature14426. Epub Apr. 22, 2015. Erratum in: *Nature*. Jul. 16, 2015;523(7560):370.
- Kuan et al., A systematic evaluation of nucleotide properties for CRISPR sgRNA design. *BMC Bioinformatics*. Jun. 6, 2017;18(1):297. doi: 10.1186/s12859-017-1697-6.
- Kweon et al., A CRISPR-based base-editing screen for the functional assessment of BRCA1 variants. *Oncogene*. Jan. 2020;39(1):30-35. doi: 10.1038/s41388-019-0968-2. Epub Aug. 29, 2019.
- Langer et al., Chemical and Physical Structure of Polymers as Carriers for Controlled Release of Bioactive Agents: A Review. *J Macromol Sci, Part C*, 1983;23(1):61-126. doi: 10.1080/07366578308079439.
- Le et al., SMN1Delta7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. *Hum Mol Genet*. Mar. 15, 2005;14(6):845-57. doi: 10.1093/hmg/ddi078. Epub Feb. 9, 2005.
- Lefebvre et al., Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*. Jan. 13, 1995;80(1):155-65. doi: 10.1016/0092-8674(95)90460-3.
- Lesinski et al., The potential for targeting the STAT3 pathway as a novel therapy for melanoma. *Future Oncol*. Jul. 2013;9(7):925-7. doi: 10.2217/fon.13.83. Author Manuscript. 4 pages.
- Lin et al., [Construction and evaluation of DnaB split intein high expression vector and a six amino acids cyclic peptide library]. Sheng Wu Gong Cheng Xue Bao. Nov. 2008;24(11):1924-30. Chinese.
- Lindahl, T., Instability and decay of the primary structure of DNA. *Nature*. Apr. 22, 1993;362(6422):709-15. doi: 10.1038/362709a0.
- Liu et al., Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat Struct Mol Biol*. Oct. 2010; 17(10):1260-2. doi: 10.1038/nsmb.1904. Epub Aug. 22, 2010.
- Liu et al., Improving Editing Efficiency for the Sequences with NGH PAM Using xCas9-Derived Base Editors. *Mol Ther Nucleic Acids*. Sep. 6, 2019;17:626-635. doi: 10.1016/j.omtn.2019.06.024. Epub Jul. 12, 2019.
- Liu et al., Intrinsic Nucleotide Preference of Diversifying Base Editors Guides Antibody Ex Vivo Affinity Maturation. *Cell Rep*. Oct. 23, 2018;25(4):884-892.e3. doi: 10.1016/j.celrep.2018.09.090.
- Liu et al., Usherin is required for maintenance of retinal photoreceptors and normal development of cochlear hair cells. *Proc Natl Acad Sci U S A*. Mar. 13, 2007;104(11):4413-8. doi: 10.1073/pnas.0610950104. Epub Mar. 5, 2007.
- Lorson et al., A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci U S A*. May 25, 1999;96(11):6307-11. doi: 10.1073/pnas.96.11.6307.
- Lutz et al., Postsymptomatic restoration of SMN rescues the disease phenotype in a mouse model of severe spinal muscular atrophy. *J Clin Invest*. Aug. 2011;121(8):3029-41. doi: 10.1172/JCI57291. Epub Jul. 25, 2011.
- Ma et al., Human RAD52 interactions with replication protein A and the RAD51 presynaptic complex. *J Biol Chem*. Jul. 14, 2017;292(28):11702-11713. doi: 10.1074/jbc.M117.794545. Epub May 27, 2017.
- Maerker et al., A novel Usher protein network at the periciliary reloading point between molecular transport machineries in vertebrate photoreceptor cells. *Hum Mol Genet*. Jan. 1, 2008;17(1):71-86. doi: 10.1093/hmg/ddm285. Epub Sep. 28, 2007.
- Mali et al., CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. Sep. 2013;31(9):Supplemental Info. doi: 10.1038/nbt.2675. Epub Aug. 1, 2013.
- Marcovitz et al., Frustration in protein-DNA binding influences conformational switching and target search kinetics. *Proc Natl Acad Sci U S A*. Nov. 1, 2011;108(44):17957-62. doi: 10.1073/pnas.1109594108. Epub Oct. 14, 2011.
- Marsden et al., The Tumor-Associated Variant RAD51 G151D Induces a Hyper-Recombination Phenotype. *PLOS Genet*. Aug. 11, 2016;12(8):e1006208. doi: 10.1371/journal.pgen.1006208.
- Mason et al., Non-enzymatic roles of human RAD51 at stalled replication forks. *bioRxiv*. Jul. 31, 2019; doi.org/10.1101/359380. 36 pages. *bioRxiv* preprint first posted online Jul. 31, 2019.
- Melero et al., Therapeutic vaccines for cancer: an overview of clinical trials. *Nat Rev Clin Oncol*. Sep. 2014;11(9):509-24. doi: 10.1038/nrclinonc.2014.111. Epub Jul. 8, 2014.
- Mendell et al., Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N Engl J Med*. Nov. 2, 2017;377(18):1713-1722. doi: 10.1056/NEJMoa1706198.
- Min et al., Deep learning in bioinformatics. *Brief Bioinform*. Sep. 1, 2017;18(5):851-869. doi: 10.1093/bib/bbw068.
- Monani et al., A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum Mol Genet*. Jul. 1999;8(7):1177-83. doi: 10.1093/hmg/8.7.1177.
- Murray et al., Selective vulnerability of motor neurons and dissociation of pre- and post-synaptic pathology at the neuromuscular junction in mouse models of spinal muscular atrophy. *Hum Mol Genet*. Apr. 1, 2008;17(7):949-62. doi: 10.1093/hmg/ddm367. Epub Dec. 8, 2007.
- Murugan et al., The Revolution Continues: Newly Discovered Systems Expand the CRISPR-Cas Toolkit. *Mol Cell*. Oct. 5, 2017;68(1):15-25. doi: 10.1016/j.molcel.2017.09.007.
- Nelson et al., In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science*. Jan. 22, 2016;351(6271):403-7. doi: 10.1126/science.aad5143. Epub Dec. 31, 2015.
- Ottesen, ISS-N1 makes the First FDA-approved Drug for Spinal Muscular Atrophy. *Transl Neurosci*. Jan. 26, 2017;8:1-6. doi: 10.1515/tnsci-2017-0001.
- Ousterout et al., Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat Commun*. Feb. 18, 2015;6:6244. doi: 10.1038/ncomms5724.
- Parente et al., Advances in spinal muscular atrophy therapeutics. *Ther Adv Neurol Disord*. Feb. 5, 2018;11:1756285618754501. doi: 10.1177/1756285618754501. 13 pages.
- Passini et al., Antisense oligonucleotides delivered to the mouse CNS ameliorate symptoms of severe spinal muscular atrophy. *Sci Transl Med*. Mar. 2, 2011;3(72):72ra18. doi: 10.1126/scitranslmed.3001777.
- Pellegrini et al., Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature*. Nov. 21, 2002;420(6913):287-93. doi: 10.1038/nature01230. Epub Nov. 10, 2002.
- Pendse et al., Exon 13-skipped USH2A protein retains functional integrity in mice, suggesting an exo-skipping therapeutic approach

(56)

References Cited**OTHER PUBLICATIONS**

- to treat USH2A-associated disease. *bioRxiv* preprint. Feb. 4, 2020. Retrieved from www.biorxiv.org. doi: 10.1101/2020.02.04.934240. 34 pages.
- Pendse et al., In Vivo Assessment of Potential Therapeutic Approaches for USH2A-Associated Diseases. *Adv Exp Med Biol.* 2019;1185:91-96. doi: 10.1007/978-3-030-27378-1_15.
- Perez-Palma et al., Simple ClinVar: an interactive web server to explore and retrieve gene and disease variants aggregated in ClinVar database. *Nucleic Acids Res.* Jul. 2, 2019;47(W1):W99-W105. doi: 10.1093/nar/gkz411.
- Porensky et al., A single administration of morpholino antisense oligomer rescues spinal muscular atrophy in mouse. *Hum Mol Genet.* Apr. 1, 2012;21(7):1625-38. doi: 10.1093/hmg/ddr600. Epub Dec. 20, 2011.
- Prasad et al., Visualizing the assembly of human Rad51 filaments on double-stranded DNA. *J Mol Biol.* Oct. 27, 2006;363(3):713-28. doi: 10.1016/j.jmb.2006.08.046. Epub Aug. 22, 2006.
- Purcell et al., More than one reason to rethink the use of peptides in vaccine design. *Nat Rev Drug Discov.* May 2007;6(5):404-14. doi: 10.1038/nrd2224.
- Rajagopal et al., High-throughput mapping of regulatory DNA. *Nat Biotechnol.* Feb. 2016;34(2):167-74. doi: 10.1038/nbt.3468. Epub Jan. 25, 2016.
- Ramos et al., Age-dependent SMN expression in disease-relevant tissue and implications for SMA treatment. *J Clin Invest.* Nov. 1, 2019;129(11):4817-4831. doi: 10.1172/JCI124120.
- Reiners et al., Scaffold protein harmonin (USH1C) provides molecular links between Usher syndrome type 1 and type 2. *Hum Mol Genet.* Dec. 15, 2005;14(24):3933-43. doi: 10.1093/hmg/ddi417. Epub Nov. 21, 2005.
- Richardson et al., CRISPR-Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat Genet.* Aug. 2018;50(8):1132-1139. doi: 10.1038/s41588-018-0174-0. Epub Jul. 27, 2018.
- Richardson et al., Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature.* Jun. 8, 2000;405(6787):697-700. doi: 10.1038/35015097.
- Rivoltini et al., A superagonist variant of peptide MART1/Melan A27-35 elicits anti-melanoma CD8+ T cells with enhanced functional characteristics: implication for more effective immunotherapy. *Cancer Res.* Jan. 15, 1999;59(2):301-6. 18 pages.
- Rodriguez-Muela et al., Single-Cell Analysis of SMN Reveals Its Broader Role in Neuromuscular Disease. *Cell Rep.* Feb. 7, 2017;18(6):1484-1498 and Supplemental Information. doi: 10.1016/j.celrep.2017.01.035.
- Saayman et al., The therapeutic application of CRISPR/Cas9 technologies for HIV. *Expert Opin Biol Ther.* Jun. 2015;15(6):819-30. doi: 10.1517/1472598.2015.1036736. Epub Apr. 12, 2015.
- San Filippo et al., Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem.* 2008;77:229-57. doi: 10.1146/annurev.biochem.77.061306.125255.
- Sanjurjo-Soriano et al., Genome Editing in Patient iPSCs Corrects the Most Prevalent USH2A Mutations and Reveals Intriguing Mutant mRNA Expression Profiles. *Mol Ther Methods Clin Dev.* Nov. 27, 2019;17:156-173. doi: 10.1016/j.mtmt.2019.11.016.
- Schlacher et al., Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell.* May 13, 2011;145(4):529-42. doi: 10.1016/j.cell.2011.03.041. Erratum in: *Cell.* Jun. 10, 2011;145(6):993.
- Schrank et al., Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc Natl Acad Sci USA.* Sep. 2, 1997;94(18):9920-5. doi: 10.1073/pnas.94.18.9920.
- Shen et al., Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods.* Apr. 2014;11(4):399-402. doi: 10.1038/nmeth.2857. Epub Mar. 2, 2014.
- Singh et al., Splicing of a critical exon of human Survival Motor Neuron is regulated by a unique silencer element located in the last intron. *Mol Cell Biol.* Feb. 2006;26(4):1333-46. doi: 10.1128/MCB.26.4.1333-1346.2006.
- Song et al., RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat Commun.* Jan. 28, 2016;7:10548. doi: 10.1038/ncomms10548.
- Sorusch et al., Characterization of the ternary Usher syndrome SANS/ush2a/whirlin protein complex. *Hum Mol Genet.* Mar. 15, 2017;26(6):1157-1172. doi: 10.1093/hmg/ddx027.
- Stark et al., ATP hydrolysis by mammalian RAD51 has a key role during homology-directed DNA repair. *J Biol Chem.* Jun. 7, 2002;277(23):20185-94. doi: 10.1074/jbc.M112132200. Epub Mar. 28, 2002.
- Sumner et al., Two breakthrough gene-targeted treatments for spinal muscular atrophy: challenges remain. *J Clin Invest.* Aug. 1, 2018;128(8):3219-3227. doi: 10.1172/JCI121658. Epub Jul. 9, 2018.
- Talbot et al., Spinal muscular atrophy. *Semin Neurol.* Jun. 2001;21(2):189-97. doi: 10.1055/s-2001-15264.
- Tan et al., Engineering of high-precision base editors for site-specific single nucleotide replacement. *Nat Commun.* Jan. 25, 2019;10(1):439. doi: 10.1038/s41467-018-08034-8. Erratum in: *Nat Commun.* May 1, 2019;10(1):2019.
- Vakulskas et al., A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat Med.* Aug. 2018;24(8):1216-1224. doi: 10.1038/s41591-018-0137-0. Epub Aug. 6, 2018.
- Van Den Oord et al., Pixel Recurrent Neural Networks. *Proceedings of the 33rd International Conference on Machine Learning. Journal of Machine Learning Research.* Aug. 19, 2016. vol. 48. 11 pages.
- Vidal et al., Yeast forward and reverse 'n'-hybrid systems. *Nucleic Acids Res.* Feb. 15, 1999;27(4):919-29. doi: 10.1093/nar/27.4.919.
- Vigneron et al., Database of T cell-defined human tumor antigens: the 2013 update. *Cancer Immun.* Jul. 15, 2013;13:15. 6 pages.
- Webb et al., Functional and structural characteristics of NY-ESO-1-related HLA A2-restricted epitopes and the design of a novel immunogenic analogue. *J Biol Chem.* May 28, 2004;279(22):23438-46. doi: 10.1074/jbc.M314066200. Epub Mar. 5, 2004.
- Wirth et al., Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. *Hum Genet.* May 2006;119(4):422-8. doi: 10.1007/s00439-006-0156-7. Epub Mar. 1, 2006.
- Woo et al., Gene activation of SMN by selective disruption of lncRNA-mediated recruitment of PRC2 for the treatment of spinal muscular atrophy. *Proc Natl Acad Sci U S A.* Feb. 21, 2017;114(8):E1509-E1518. doi: 10.1073/pnas.1616521114. Epub Feb. 13, 2017.
- Wu et al., A novel SCN9A mutation responsible for primary erythromelalgia and is resistant to the treatment of sodium channel blockers. *PLoS One.* 2013;8(1):e55212. doi: 10.1371/journal.pone.0055212. Epub Jan. 31, 2013. 15 pages.
- Yamane et al., Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol.* Jan. 2011;12(1):62-9. doi: 10.1038/ni.1964. Epub Nov. 28, 2010.
- Yang et al., BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science.* Sep. 13, 2002;297(5588):1837-48. doi: 10.1126/science.297.5588.1837.
- Yang et al., The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature.* Feb. 10, 2005;433(7026):653-7. doi: 10.1038/nature03234.
- Yu et al., Dynamic control of Rad51 recombinase by self-association and interaction with BRCA2. *Mol Cell.* Oct. 2003;12(4):1029-41. doi: 10.1016/s1097-2765(03)00394-0.
- Zaremba et al., Identification of an enhancer agonist cytotoxic T lymphocyte peptide from human carcinoembryonic antigen. *Cancer Res.* Oct. 15, 1997;57(20):4570-7.
- Zhang et al., Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* Feb. 20, 2017;18(1):35. doi: 10.1186/s13059-017-1164-8.

(56)

References Cited**OTHER PUBLICATIONS**

- Zhang et al., Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9. *PLoS One*. Mar. 24, 2015;10(3):e0120396. doi: 10.1371/journal.pone.0120396. 14 pages.
- Zhu et al., Novel Thrombotic Function of a Human SNP in STXBP5 Revealed by CRISPR/Cas9 Gene Editing in Mice. *Arterioscler Thromb Vasc Biol*. Feb. 2017;37(2):264-270. doi: 10.1161/ATVBAHA.116.308614. Epub Dec. 29, 2016.
- [No Author Listed], MutL homolog 1. UniProtKB Acc. No. F1MPG0. May 3, 2011. Accessible at <https://rest.uniprot.org/unisave/F1MPG0?format=txt&versions=1>. 1 page.
- Acharya et al., hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc Natl Acad Sci U S A*. Nov. 26, 1996;93(24):13629-34. doi: 10.1073/pnas.93.24.13629.
- Basila et al., Minimal 2'-O-methyl phosphorothioate linkage modification pattern of synthetic guide RNAs for increased stability and efficient CRISPR-Cas9 gene editing avoiding cellular toxicity. *PLoS One*. Nov. 27, 2017;12(11):e0188593. doi: 10.1371/journal.pone.0188593.
- Bertsimas et al., Simulated annealing. *Statistical Science*. Feb. 1993;8(1):10-15. doi: 10.1214/ss/1177011077.
- Chen et al., Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*. Dec. 19, 2013;155(7):1479-91. doi: 10.1016/j.cell.2013.12.001. Erratum in: *Cell*. Jan. 16, 2014;156(1-2):373.
- Edraki et al., A Compact, High-Accuracy Cas9 with a Dinucleotide PAM for In Vivo Genome Editing. *Mol Cell*. Feb. 21, 2019;73(4):714-726.e4 and Supplemental Info. doi: 10.1016/j.molcel.2018.12.003. Epub Dec. 20, 2018.
- Fang et al., Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. *J Biol Chem*. Jun. 5, 1993;268(16):11838-44.
- Fang et al., The Menu of Features that Define Primary MicroRNAs and Enable De Novo Design of MicroRNA Genes. *Mol Cell*. Oct. 1, 2015;60(1):131-45. doi: 10.1016/j.molcel.2015.08.015. Epub Sep. 24, 2015.
- Feng et al., Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res*. Oct. 2013;23(10):1229-32. doi: 10.1038/cr.2013.114. Epub Aug. 20, 2013.
- Fishel et al., The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*. Dec. 3, 1993;75(5):1027-38. doi: 10.1016/0092-8674(93)90546-3. Erratum in: *Cell*. Apr. 8, 1994;77(1):1 p following 166.
- Fu et al., Targeted genome editing in human cells using CRISPR/Cas nucleases and truncated guide RNAs. *Methods Enzymol*. 2014;546:21-45. doi: 10.1016/B978-0-12-801185-0.00002-7.
- Geisberg et al., Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. *Cell*. Feb. 13, 2014;156(4):812-24. doi: 10.1016/j.cell.2013.12.026.
- Geng et al., In vitro studies of DNA mismatch repair proteins. *Anal Biochem*. Jun. 15, 2011;413(2):179-84. doi: 10.1016/j.ab.2011.02.017. Epub Feb. 15, 2011.
- Genschel et al., Human exonuclease I is required for 5' and 3' mismatch repair. *J Biol Chem*. Apr. 12, 2002;277(15):13302-11. doi: 10.1074/jbc.M111854200. Epub Jan. 24, 2002.
- Genschel et al., Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J Biol Chem*. Jul. 31, 1998;273(31):19895-901. doi: 10.1074/jbc.273.31.19895. Erratum in: *J Biol Chem* Oct. 9, 1998;273(41):27034.
- Green et al., Characterization of the mechanical unfolding of RNA pseudoknots. *J Mol Biol*. Jan. 11, 2008;375(2):511-28. doi: 10.1016/j.jmb.2007.05.058. Epub May 26, 2007.
- Gueneau et al., Structure of the MutL α C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. *Nat Struct Mol Biol*. Apr. 2013;20(4):461-8. doi: 10.1038/nsmb.2511. Epub Feb. 24, 2013.
- Guerrette et al., The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer. *J Biol Chem*. Mar. 5, 1999;274(10):6336-41. doi: 10.1074/jbc.274.10.6336.
- Gupta et al., Mechanism of mismatch recognition revealed by human MutS β bound to unpaired DNA loops. *Nat Struct Mol Biol*. Dec. 18, 2011;19(1):72-8. doi: 10.1038/nsmb.2175.
- Hänsel-Hertsch et al., DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol*. May 2017;18(5):279-284. doi: 10.1038/nrm.2017.3. Epub Feb. 22, 2017.
- Houck-Loomis et al., An equilibrium-dependent retroviral mRNA switch regulates translational recoding. *Nature*. Nov. 27, 2011;480(7378):561-4. doi: 10.1038/nature10657.
- Houseley et al., The many pathways of RNA degradation. *Cell*. Feb. 20, 2009;136(4):763-76. doi: 10.1016/j.cell.2009.01.019.
- Iaccarino et al., hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutS α . *EMBO J*. May 1, 1998;17(9):2677-86. doi: 10.1093/emboj/17.9.2677.
- Ibrahim et al., RNA recognition by 3'-to-5' exonucleases: the substrate perspective. *Biochim Biophys Acta*. Apr. 2008;1779(4):256-65. doi: 10.1016/j.bbapm.2007.11.004. Epub Dec. 3, 2007.
- Iyer et al., DNA mismatch repair: functions and mechanisms. *Chem Rev*. Feb. 2006;106(2):302-23. doi: 10.1021/cr0404794.
- Jost et al., Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. *Nat Biotechnol*. Mar. 2020;38(3):355-364. doi: 10.1038/s41587-019-0387-5. Epub Jan. 13, 2020.
- Kadyrov et al., Endonucleolytic function of MutL α in human mismatch repair. *Cell*. Jul. 28, 2006;126(2):297-308. doi: 10.1016/j.cell.2006.05.039.
- Ku et al., Nucleic Acid Aptamers: An Emerging Tool for Biotechnology and Biomedical Sensing. *Sensors (Basel)*. Jul. 6, 2015;15(7):16281-313. doi: 10.3390/s150716281.
- Kunkel et al., DNA mismatch repair. *Annu Rev Biochem*. 2005;74:681-710. doi: 10.1146/annurev.biochem.74.082803.133243.
- Kwok et al., G-Quadruplexes: Prediction, Characterization, and Biological Application. *Trends Biotechnol*. Oct. 2017;35(10):997-1013. doi: 10.1016/j.tibtech.2017.06.012. Epub Jul. 26, 2017.
- Lahue et al., DNA mismatch correction in a defined system. *Science*. Jul. 14, 1989;245(4914):160-4. doi: 10.1126/science.2665076.
- Leach et al., Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell*. Dec. 17, 1993;75(6):1215-25. doi: 10.1016/0092-8674(93)90330-s.
- Longsworth, Expanding the Enzymatic Activity of the Programmable Endonuclease Cas9 in Zebrafish. Thesis. Rice University. Houston, TX. May 17, 2019. 41 pages.
- Lujan et al., Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res*. Nov. 2014;24(11):1751-64. doi: 10.1101/gr.178335.114. Epub Sep. 12, 2014.
- MacFadden et al., Mechanism and structural diversity of exoribonuclease-resistant RNA structures in flaviviral RNAs. *Nat Commun*. Jan. 9, 2018;9(1):119. doi: 10.1038/s41467-017-02604-y.
- Mangeot et al., Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nat Commun*. Jan. 3, 2019;10(1):45. doi: 10.1038/s41467-018-07845-z.
- Micozzi et al., Human cytidine deaminase: a biochemical characterization of its naturally occurring variants. *Int J Biol Macromol*. Feb. 2014;63:64-74. doi: 10.1016/j.ijbiomac.2013.10.029. Epub Oct. 29, 2013. Erratum in: *Int J Biol Macromol*. Feb. 2014;63:262.
- Millevoi et al., G-quadruplexes in RNA biology. *Wiley Interdiscip Rev RNA*. Jul.-Aug. 2012;3(4):495-507. doi: 10.1002/wrna.1113. Epub Apr. 4, 2012.
- Pandey et al., Effect of loops and G-quartets on the stability of RNA G-quadruplexes. *J Phys Chem B*. Jun. 13, 2013;117(23):6896-905. doi: 10.1021/jp401739m. Epub May 29, 2013. Supplementary Information, 21 pages.
- Parsons et al., Hypermutability and mismatch repair deficiency in RER+ tumor cells. *Cell*. Dec. 17, 1993;75(6):1227-36. doi: 10.1016/0092-8674(93)90331-j.
- Petit et al., Powerful mutators lurking in the genome. *Philos Trans R Soc Lond B Biol Sci*. Mar. 12, 2009;364(1517):705-15. doi: 10.1098/rstb.2008.0272.

(56)

References Cited**OTHER PUBLICATIONS**

- Pijlman et al., A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. *Cell Host Microbe.* Dec. 11, 2008;4(6):579-91. doi: 10.1016/j.chom.2008.10.007.
- Plotukh et al., Directed evolution of sortase A mutants with altered substrate selectivity profiles. *J Am Chem Soc.* Nov. 9, 2011;133(44):17536-9. doi: 10.1021/ja205630g. Epub Oct. 13, 2011.
- Plotz et al., N-terminus of hMLH1 confers interaction of hMut λ lpha and hMut λ beta with hMut σ lpha. *Nucleic Acids Res.* Jun. 15, 2003;31(12):3217-26. doi: 10.1093/nar/gkg420.
- Räschle et al., Mutations within the hMLH1 and hPMS2 subunits of the human Mut λ lpha mismatch repair factor affect its ATPase activity, but not its ability to interact with hMut σ lpha. *J Biol Chem.* Jun. 14, 2002;277(24):21810-20. doi: 10.1074/jbc.M108787200. Epub Apr. 10, 2002.
- Robert et al., Virus-Like Particles Derived from HIV-1 for Delivery of Nuclear Proteins: Improvement of Production and Activity by Protein Engineering. *Mol Biotechnol.* Jan. 2017;59(1):9-23. doi: 10.1007/s12033-016-9987-1.
- Shcherbakova et al., Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. *Mol Cell Biol.* Apr. 1999;19(4):3177-83. doi: 10.1128/MCB.19.4.3177.
- Steckelberg et al., A folded viral noncoding RNA blocks host cell exoribonucleases through a conformationally dynamic RNA structure. *Proc Natl Acad Sci U S A.* Jun. 19, 2018;115(25):6404-6409. doi: 10.1073/pnas.1802429115. Epub Jun. 4, 2018.
- Strand et al., Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature.* Sep. 16, 1993;365(6443):274-6. doi: 10.1038/365274a0. Erratum in: *Nature* Apr. 7, 1994;368(6471):569.
- Su et al., Mispair specificity of methyl-directed DNA mismatch correction in vitro. *J Biol Chem.* May 15, 1988;263(14):6829-35. Erratum in: *J Biol Chem* Aug. 5, 1988;263(22):11015.
- Sugawara et al., Heteroduplex rejection during single-strand annealing requires Sgs1 helicase and mismatch repair proteins Msh2 and Msh6 but not Pms1. *Proc Natl Acad Sci U S A.* Jun. 22, 2004;101(25):9315-20. doi: 10.1073/pnas.0305749101. Epub Jun. 15, 2004.
- Supek et al., Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature.* May 7, 2015;521(7550):81-4. doi: 10.1038/nature14173. Epub Feb. 23, 2015.
- Svitashov et al., Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol.* Oct. 2015;169(2):931-45. doi: 10.1104/pp15.00793. Epub Aug. 12, 2015.
- Thomas et al., Heteroduplex repair in extracts of human HeLa cells. *J Biol Chem.* Feb. 25, 1991;266(6):3744-51.
- Tomer et al., Contribution of human mlh1 and pms2 ATPase activities to DNA mismatch repair. *J Biol Chem.* Jun. 14, 2002;277(24):21801-9. doi: 10.1074/jbc.M111342200. Epub Mar. 15, 2002.
- Tran et al., Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Mol Cell Biol.* May 1997;17(5):2859-65. doi: 10.1128/MCB.17.5.2859.
- Umar et al., DNA loop repair by human cell extracts. *Science.* Nov. 4, 1994;266(5186):814-6. doi: 10.1126/science.797367.
- Warren et al., Structure of the human Mut σ lpha DNA lesion recognition complex. *Mol Cell.* May 25, 2007;26(4):579-92. doi: 10.1016/j.molcel.2007.04.018.
- Wu et al., MLV based viral-like-particles for delivery of toxic proteins and nuclear transcription factors. *Biomaterials.* Sep. 2014;35(29):8416-26. doi: 10.1016/j.biomaterials.2014.06.006. Epub Jul. 3, 2014.
- Wu et al., Widespread Influence of 3'-End Structures on Mammalian mRNA Processing and Stability. *Cell.* May 18, 2017;169(5):905-917.e11. doi: 10.1016/j.cell.2017.04.036.
- Xi et al., C-terminal Loop Mutations Determine Folding and Secretion Properties of PCSK9. *Biochem Mol Biol J.* 2016;2(3):17. doi: 10.21767/2471-8084.100026. 12 pages.
- Yi et al., Engineering of TEV protease variants by yeast ER sequestration screening (YESSION) of combinatorial libraries. *Proc Natl Acad Sci U S A.* Apr. 30, 2013;110(18):7229-34. doi: 10.1073/pnas.1215994110. Epub Apr. 15, 2013.
- Zhang et al., Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell.* Sep. 9, 2005;122(5):693-705. doi: 10.1016/j.cell.2005.06.027.
- [No. Author Listed] Ncbi Reference Sequence: WP_032188360.1. Apr. 6, 2015. 1 page.
- [No Author Listed], tRNA-specific adenosine deaminase [*Escherichia coli*]. GenBank Acc. No. CTS26096.1. Accessible at <https://www.ncbi.nlm.nih.gov/protein/CTS26096.1>. Aug. 22, 2015. 1 page.
- [No Author Listed], tumor-specific antigen. Retrieved from <http://www.cancer.gov/publications/dictionaries/cancer-terms/def/tumor-specific-antigen>. Retrieved on Oct. 7, 2022. 1 page.
- Alizadeh et al., HR9: An Important Cell Penetrating Peptide for Delivery of HCV NS3 DNA into HEK-293T Cells. *Avicenna J Med Biotechnol.* Jan.-Mar. 2020;12(1):44-51.
- Avidan et al., Expression and characterization of a recombinant novel reverse transcriptase of a porcine endogenous retrovirus. *Virology.* Mar. 15, 2003;307(2):341-57. doi: 10.1016/s0042-6822(02)00131-9.
- Bae et al., Heteroclitic CD33 peptide with enhanced anti-acute myeloid leukemic immunogenicity. *Clin Cancer Res.* Oct. 15, 2004;10(20):7043-52. doi: 10.1158/1078-0432.CCR-04-0322.
- Baños-Sanz et al., Crystal structure and functional insights into uracil-DNA glycosylase inhibition by phage Φ 29 DNA mimic protein p56. *Nucleic Acids Res.* Jul. 2013;41(13):6761-73. doi: 10.1093/nar/gkt395. Epub May 13, 2013.
- Celluzzi et al., Peptide-pulsed dendritic cells induce antigen-specific CTL-mediated protective tumor immunity. *J Exp Med.* Jan. 1, 1996;183(1):283-7. doi: 10.1084/jem.183.1.283.
- Cherian et al., Faster protein splicing with the Nostoc punctiforme DnaE intein using non-native extein residues. *J Biol Chem.* Mar. 1, 2013;288(9):6202-11. doi: 10.1074/jbc.M112.433094. Epub Jan. 10, 2013.
- Choi et al., Optimization of AAV expression cassettes to improve packaging capacity and transgene expression in neurons. *Mol Brain.* Mar. 11, 2014;7:17. doi: 10.1186/1756-6606-7-17.
- Damdinordj et al., A comparative analysis of constitutive promoters located in adeno-associated viral vectors. *PLoS One.* Aug. 29, 2014;9(8):e106472. doi: 10.1371/journal.pone.0106472.
- Ding et al., Gene therapy for cardiovascular disease. *Journal of Shanghai University (Natural Science Edition)*. 2016;3:270-9. DOI: 10.3969/j.issn.1007-2861.2016.03.013.
- Ekman et al., CRISPR-Cas9-Mediated Genome Editing Increases Lifespan and Improves Motor Deficits in a Huntington's Disease Mouse Model. *Mol Ther Nucleic Acids.* Sep. 6, 2019;17:829-839. doi: 10.1016/j.omtn.2019.07.009. Epub Jul. 26, 2019.
- Eriksen et al., Occlusion of the Ribosome Binding Site Connects the Translational Initiation Frequency, mRNA Stability and Premature Transcription Termination. *Front Microbiol.* Mar. 14, 2017;8:362. doi: 10.3389/fmicb.2017.00362.
- Fikes et al., Design of multi-epitope, analogue-based cancer vaccines. *Expert Opin Biol Ther.* Sep. 2003;3(6):985-93. doi: 10.1517/14712598.3.6.985.
- GenBank Submission; NIH/NCBI, Accession No. NP_060228.2. Bi et al., Dec. 21, 2005. 1 page.
- GenBank Submission; NIH/NCBI, Accession No. NP_062826.2. Bokar et al., Sep. 18, 2004. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. NP_066012.1. Ota et al., Apr. 3, 2005. 2 pages.
- GenBank Submission; NIH/NCBI, Accession No. WP_042518169. 1. No Author, Feb. 10, 2015. 1 page.
- Hizi et al., Retroviral reverse transcriptases (other than those of HIV-1 and murine leukemia virus): a comparison of their molecular and biochemical properties. *Virus Res.* Jun. 2008;134(1-2):203-20. doi: 10.1016/j.virusres.2007.12.008. Epub Mar. 3, 2008.

(56)

References Cited**OTHER PUBLICATIONS**

- Houghton et al., Immunological validation of the EpitOptimizer program for streamlined design of heteroclitic epitopes. *Vaccine*. Jul. 20, 2007;25(29):5330-42. doi: 10.1016/j.vaccine.2007.05.008. Epub Jun. 4, 2007.
- Hwang et al., Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS One*. Jul. 9, 2013;8(7):e68708. doi: 10.1371/journal.pone.0068708.
- Kirshenboim et al., Expression and characterization of a novel reverse transcriptase of the LTR retrotransposon Tf1. *Virology*. Sep. 30, 2007;366(2):263-76. doi: 10.1016/j.virol.2007.04.002. Epub May 23, 2007.
- Kwon et al., Precision targeting tumor cells using cancer-specific InDel mutations with CRISPR-Cas9. *Proc Natl Acad Sci U S A*. Mar. 1, 2022;119(9):e2103532119. doi: 10.1073/pnas.2103532119.
- Men et al., Assessment of immunogenicity of human Melan-A peptide analogues in HLA-A*0201/Kb transgenic mice. *J Immunol*. Mar. 15, 1999;162(6):3566-73.
- Misra et al., An enzymatically active chimeric HIV-1 reverse transcriptase (RT) with the RNase-H domain of murine leukemia virus RT exists as a monomer. *J Biol Chem*. Apr. 17, 1998;273(16):9785-9. doi: 10.1074/jbc.273.16.9785.
- Nicholson et al., Tuning T cell activation threshold and effector function with cross-reactive peptide ligands. *Int Immunopharmacol*. Feb. 2000;12(2):205-13. doi: 10.1093/intimm.12.2.205.
- Niemeyer, C.M., Semisynthetic DNA-protein conjugates for biosensing and nanofabrication. *Angew Chem Int Ed Engl*. Feb. 8, 2010;49(7):1200-16. doi: 10.1002/anie.200904930.
- Nowak et al., Ty3 reverse transcriptase complexed with an RNA-DNA hybrid shows structural and functional asymmetry. *Nat Struct Mol Biol*. Apr. 2014;21(4):389-96. doi: 10.1038/nsmb.2785. Epub Mar. 9, 2014. Author Manuscript, 22 pages.
- Pfeiffer et al., Altered peptide ligands can control CD4 T lymphocyte differentiation in vivo. *J Exp Med*. Apr. 1, 1995;181(4):1569-74. doi: 10.1084/jem.181.4.1569.
- Raaijmakers et al., CRISPR/Cas Applications in Myotonic Dystrophy: Expanding Opportunities. *Int J Mol Sci*. Jul. 27, 2019;20(15):3689. doi: 10.3390/ijms20153689.
- Riddle et al., Frameshift suppression: a nucleotide addition in the anticodon of a glycine transfer RNA. *Nat New Biol*. Apr. 25, 1973;242(121):230-4. doi: 10.1038/newbio242230a0.
- Riddle et al., Frameshift suppressors. II. Genetic mapping and dominance studies. *J Mol Biol*. May 28, 1972;66(3):483-93. doi: 10.1016/0022-2836(72)90428-7.
- Riddle et al., Suppressors of frameshift mutations in *Salmonella typhimurium*. *J Mol Biol*. Nov. 28, 1970;54(1):131-44. doi: 10.1016/0022-2836(70)90451-1.
- Romagnani, Lymphokine production by human T cells in disease states. *Annu Rev Immunol*. 1994;12:227-57. doi: 10.1146/annurev.iy.12.040194.001303.
- Sadowski et al., The sequence-structure relationship and protein function prediction. *Curr Opin Struct Biol*. Jun. 2009;19(3):357-62. doi: 10.1016/j.sbi.2009.03.008. Epub May 4, 2009.
- Salazar et al., Agonist peptide from a cytotoxic t-lymphocyte epitope of human carcinoembryonic antigen stimulates production of tc1-type cytokines and increases tyrosine phosphorylation more efficiently than cognate peptide. *Int J Cancer*. Mar. 15, 2000;85(6):829-38. doi: 10.1002/(sici)1097-0215(20000315)85:6<829::aid-ijc16>3.0.co;2-k.
- SCORE Results for US 2014-0186919 A1 to Zhang et al. Aug. 28, 2014. 3 pages.
- Seffernick et al., Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. *J Bacteriol*. Apr. 2001;183(8):2405-10. doi: 10.1128/JB.183.8.2405-2410.2001.
- Selby et al., Hepatitis C virus envelope glycoprotein E1 originates in the endoplasmic reticulum and requires cytoplasmic processing for presentation by class I MHC molecules. *J Immunol*. Jan. 15, 1999;162(2):669-76.
- Simon et al., Retrons and their applications in genome engineering. *Nucleic Acids Res*. Dec. 2, 2019;47(21):11007-11019. doi: 10.1093/nar/gkz865.
- Singh et al., Protein Engineering Approaches in the Post-Genomic Era. *Curr Protein Pept Sci*. 2018;19(1):5-15. doi: 10.2174/1389203718666161117114243.
- Studebaker et al., Depletion of uracil-DNA glycosylase activity is associated with decreased cell proliferation. *Biochem Biophys Res Commun*. Aug. 26, 2005;334(2):509-15. doi: 10.1016/j.bbrc.2005.06.118.
- Tang et al., Identification of *Dehalobacter* reductive dehalogenases that catalyse dechlorination of chloroform, 1,1,1-trichloroethane and 1,1-dichloroethane. *Philos Trans R Soc Lond B Biol Sci*. Mar. 11, 2013;368(1616):20120318. doi: 10.1098/rstb.2012.0318.
- Tao et al., Induction of IL-4-producing CD4+ T cells by antigenic peptides altered for TCR binding. *J Immunol*. May 1, 1997;158(9):4237-44.
- Toro et al., Comprehensive phylogenetic analysis of bacterial reverse transcriptases. *PLoS One*. Nov. 25, 2014;9(11):e114083. doi: 10.1371/journal.pone.0114083.
- Wang et al., The stimulation of low-affinity, nontolerized clones by heteroclitic antigen analogues causes the breaking of tolerance established to an immunodominant T cell epitope. *J Exp Med*. Oct. 4, 1999;190(7):983-94. doi: 10.1084/jem.190.7.983.
- Witkowski et al., Conversion of a beta-ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine. *Biochemistry*. Sep. 7, 1999;38(36):11643-50. doi: 10.1021/bi990993h.
- Yin et al., Optimizing genome editing strategy by primer-extension-mediated sequencing. *Cell Discov*. Mar. 26, 2019;5:18. doi: 10.1038/s41421-019-0088-8.
- Zhang et al., Propagated Perturbations from a Peripheral Mutation Show Interactions Supporting WW Domain Thermostability. *Structure*. Nov. 6, 2018;26(11):1474-1485.e5. doi: 10.1016/j.str.2018.07.014. Epub Sep. 6, 2018.
- Zitvogel et al., Therapy of murine tumors with tumor peptide-pulsed dendritic cells: dependence on T cells, B7 costimulation, and T helper cell 1-associated cytokines. *J Exp Med*. Jan. 1, 1996;183(1):87-97. doi: 10.1084/jem.183.1.87.
- Zügel et al., Termination of peripheral tolerance to a T cell epitope by heteroclitic antigen analogues. *J Immunol*. Aug. 15, 1998;161(4):1705-9.
- Abed et al., The Gag protein PEG10 binds to RNA and regulates trophoblast stem cell lineage specification. *PLoS One*. Apr. 5, 2019;14(4):e0214110. doi: 10.1371/journal.pone.0214110.
- Abifadel et al., Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet*. Jun. 2003;34(2):154-6. doi: 10.1038/ng1161.
- Adikusuma et al., Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. *PLoS One*. Dec. 6, 2017;12(12):e0187236. doi: 10.1371/journal.pone.0187236.
- Addgene Plasmid #42234. pMJ920. 2013. Retrieved Jan. 22, 2025. 3 pages.
- Ashley et al., Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. *Cell*. Jan. 11, 2018;172(1-2):262-274.e11. doi: 10.1016/j.cell.2017.12.022.
- Ayala-Ramirez et al., A new autosomal recessive syndrome consisting of posterior microphthalmos, retinitis pigmentosa, foveoschisis, and optic disc drusen is caused by a MFRP gene mutation. *Mol Vis*. Dec. 4, 2006;12:1483-9.
- Bender et al., Receptor-Targeted Nipah Virus Glycoproteins Improve Cell-Type Selective Gene Delivery and Reveal a Preference for Membrane-Proximal Cell Attachment. *PLoS Pathog*. Jun. 9, 2016;12(6):e1005641. doi: 10.1371/journal.ppat.1005641.
- Bernardi et al., Nucleotide sequence at the binding site for coat protein on RNA of bacteriophage R17. *Proc Natl Acad Sci U S A*. Oct. 1972;69(10):3033-7. doi: 10.1073/pnas.69.10.3033.
- Bikard et al., Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res*. Aug. 2013;41(15):7429-37. doi: 10.1093/nar/gkt520. Epub Jun. 12, 2013.

(56)

References Cited**OTHER PUBLICATIONS**

- Bolukbasi et al., DNA-binding-domain fusions enhance the targeting range and precision of Cas9. *Nat Methods*. Dec. 2015;12(12):1150-6. doi: 10.1038/nmeth.3624. Epub Oct. 19, 2015.
- Cai et al. Targeted genome editing by lentiviral protein transduction of zinc-finger and TAL-effector nucleases. *Elife*. Apr. 24, 2014;3:e01911. doi: 10.7554/elife.01911.
- Cai et al., Abstract OR021: Targeted Genome Editing by Lentiviral Protein Transduction of ZFN and Cas9 Proteins Abstract, Presented at Proceedings of the ESGCT and NVGCT Collaborative Congress: The Hague. *Human Gene Therapy*. 2014. 15 pages.
- Cai, Protein Transduction Using Lentiviral Vectors for Transposition and Site-directed Gene Editing. Thesis for the degree of Doctor of Philosophy, Aarhus University, Department of Biomedicine. 2014. 74 pages.
- Cameron et al., Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat Methods*. Jun. 2017;14(6):600-606 with Erratum. doi: 10.1038/nmeth.4284. Epub May 1, 2017. Erratum in: *Nat Methods*. Dec. 2023;20(12):2068. doi: 10.1038/s41592-023-02114-4. 8 pages.
- Chandler et al., Recombinant Adeno-Associated Viral Integration and Genotoxicity: Insights from Animal Models. *Hum Gene Ther.* Apr. 2017;28(4):314-322. doi: 10.1089/hum.2017.009.
- Chang et al., Functional characterization of the placental fusogenic membrane protein syncytin. *Biol Reprod.* Dec. 2004;71(6):1956-62. doi: 10.1095/biolreprod.104.03340. Epub Jul. 21, 2004.
- Chen et al., DNA methylation and demethylation in mammals. *J Biol Chem.* May 27, 2011;286(21):18347-53. doi: 10.1074/jbc.R110.205286. Epub Mar. 24, 2011.
- Cho et al., Heritable gene knockout in *Caenorhabditis elegans* by direct injection of Cas9-sgRNA ribonucleoproteins. *Genetics*. Nov. 2013;195(3):1177-80. doi: 10.1534/genetics.113.155853. Epub Aug. 26, 2013.
- Chylinski et al., The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* May 2013;10(5):Supplementary Material. doi: 10.4161/rna.24321. Epub Apr. 5, 2013. 12 pages.
- Cideciyan, Leber congenital amaurosis due to RPE65 mutations and its treatment with gene therapy. *Prog Retin Eye Res.* Sep. 2010;29(5):398-427. doi: 10.1016/j.preteyeres.2010.04.002. Epub Apr. 24, 2010.
- Coey, Sumoylation of thymine DNA glycosylase occurs efficiently and weakens DNA binding but does not regulate enzymatic turnover. Dissertation. 2017. 178 pages.
- Cohen et al., Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet.* Feb. 2005;37(2):161-5. doi: 10.1038/ng1509. Epub Jan. 16, 2005. Erratum in: *Nat Genet.* Mar. 2005;37(3):328.
- Cohen et al., Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med.* Mar. 23, 2006;354(12):1264-72. doi: 10.1056/NEJMoa054013.
- Contreras-Galindo et al., Human Endogenous Retrovirus Type K (HERV-K) Particles Package and Transmit HERV-K-Related Sequences. *J Virol.* Jul. 2015;89(14):7187-201. doi: 10.1128/JVI.00544-15. Epub Apr. 29, 2015.
- Coquin, Characterization of lentiviral vectors pseudotyped with murine syncytins and their cellular targets in vitro and in vivo. Thesis. Université de Évry Val d'Essonne. Defense on Dec. 10, 2019. 238 pages.
- Den Hollander et al., Leber congenital amaurosis: genes, proteins and disease mechanisms. *Prog Retin Eye Res.* Jul. 2008;27(4):391-419. doi: 10.1016/j.preteyeres.2008.05.003. Epub Jun. 1, 2008.
- Farhy-Tselnicker et al., Astrocytes, neurons, synapses: a tripartite view on cortical circuit development. *Neural Dev.* May 1, 2018;13(1):7. doi: 10.1186/s13064-018-0104-y.
- Fehér et al., Characterization of the murine leukemia virus protease and its comparison with the human immunodeficiency virus type 1 protease. *J Gen Virol.* May 2006;87(Pt 5):1321-1330. doi: 10.1099/vir.0.81382-0.
- Fitzgerald et al., Effect of an RNA interference drug on the synthesis of proprotein convertase subtilisin/kexin type 9 (PCSK9) and the concentration of serum LDL cholesterol in healthy volunteers: a randomised, single-blind, placebo-controlled, phase 1 trial. *Lancet*. Jan. 4, 2014;383(9911):60-68. doi: 10.1016/S0140-6736(13)61914-5. Epub Oct. 3, 2013.
- Fonfara et al., Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* Feb. 2014;42(4):2577-90. doi: 10.1093/nar/gkt1074. Epub Nov. 22, 2013. Supplementary Information. 67 pages.
- Gaidukov et al., A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Res.* May 4, 2018;46(8):4072-4086. doi: 10.1093/nar/gky216.
- Gao et al., Delineation of the Exact Transcription Termination Signal for Type 3 Polymerase III. *Mol Ther Nucleic Acids*. Mar. 2, 2018;10:36-44. doi: 10.1016/j.omtn.2017.11.006. Epub Nov. 21, 2017.
- Garnier et al., WW domains and retrovirus budding. *Nature*. Jun. 27, 1996;381(6585):744-5. doi: 10.1038/381744a0.
- GenBank Accession No. AAH57574.1 2009. 2 pages.
- Giannoukos et al., UDiTaSTM, a genome editing detection method for indels and genome rearrangements. *BMC Genomics*. Mar. 21, 2018;19(1):212. doi: 10.1186/s12864-018-4561-9.
- Golczak et al., Importance of membrane structural integrity for RPE65 retinoid isomerization activity. *J Biol Chem.* Mar. 26, 2010;285(13):9667-9682. doi: 10.1074/jbc.M109.063941. Epub Jan. 25, 2010.
- Gusel'nikova et al., NeuN As a Neuronal Nuclear Antigen and Neuron Differentiation Marker. *Acta Naturae*. Apr.-Jun. 2015;7(2):42-7.
- Hamilton et al., Knocking out barriers to engineered cell activity. *Science*. Feb. 28, 2020;367(6481):976-977. doi: 10.1126/science.aba9844. Epub Feb. 6, 2020.
- Heins et al., Designing Automated, High-throughput, Continuous Cell Growth Experiments Using e Volver. *J Vis Exp.* May 19, 2019;(147):10.3791/59652. doi: 10.3791/59652.
- Heintze et al., A CRISPR CASe for high-throughput silencing. *Front Genet.* Oct. 7, 2013;4:193. doi: 10.3389/fgen.2013.00193.
- Hooper et al., The C679X mutation in PCSK9 is present and lowers blood cholesterol in a Southern African population. *Atherosclerosis*. Aug. 2007;193(2):445-8. doi: 10.1016/j.atherosclerosis.2006.08.039. Epub Sep. 20, 2006.
- Jacobs et al., DNA glycosylases: in DNA repair and beyond. *Chromosoma*. Feb. 2012;121(1):1-20. doi: 10.1007/s00412-011-0347-4. Epub Nov. 3, 2011. 20 pages.
- Jinek et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. Aug. 17, 2012;337(6096):Supplementary Material. doi: 10.1126/science.1225829. Epub Jun. 28, 2012. 37 pages.
- Johnson et al., Mass spectrometry analysis reveals differences in the host cell protein species found in pseudotyped lentiviral vectors. *Biologicals*. Mar. 2018;52:59-66. doi: 10.1016/j.biologicals.2017.12.005. Epub Feb. 1, 2018.
- Johnson, Origins and evolutionary consequences of ancient endogenous retroviruses. *Nat Rev Microbiol.* Jun. 2019;17(6):355-370. doi: 10.1038/s41579-019-0189-2.
- Kameya et al., Mfrp, a gene encoding a frizzled related protein, is mutated in the mouse retinal degeneration 6. *Hum Mol Genet*. Aug. 1, 2002;11(16):1879-86. doi: 10.1093/hmg/11.16.1879.
- Katoh et al., Exploitation of the interaction of measles virus fusogenic envelope proteins with the surface receptor CD46 on human cells for microcell-mediated chromosome transfer. *BMC Biotechnol.* May 6, 2010;10:37. doi: 10.1186/1472-6750-10-37.
- Kleinsteiner et al., Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol.* Dec. 2015;33(12):1293-1298. doi: 10.1038/nbt.3404. Epub Nov. 2, 2015. Author Manuscript, 14 pages.
- Kneissl et al., Measles virus glycoprotein-based lentiviral targeting vectors that avoid neutralizing antibodies. *PLoS One*. 2012;7(10):e46667. doi: 10.1371/journal.pone.0046667. Epub Oct. 10, 2012.

(56)

References Cited**OTHER PUBLICATIONS**

- Kulcsár et al., Blackjack mutations improve the on-target activities of increased fidelity variants of SpCas9 with 5'G-extended sgRNAs. *Nat Commun.* Mar. 6, 2020;11(1):1223. doi: 10.1038/s41467-020-15021-5.
- Lau et al., In vivo epigenome editing and transcriptional modulation using CRISPR technology. *Transgenic Res.* Dec. 2018;27(6):489-509. doi: 10.1007/s11248-018-0096-8. Epub Oct. 4, 2018.
- Lebar et al., A tunable orthogonal coiled-coil interaction toolbox for engineering mammalian cells. *Nat Chem Biol.* May 2020;16(5):513-519. doi: 10.1038/s41589-019-0443-y. Epub Jan. 6, 2020.
- Lee et al., Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog.* Jan. 2007;3(1):e10. doi: 10.1371/journal.ppat.0030010.
- Leenay et al., Identifying and Visualizing Functional PAM Diversity across CRISPR-Cas Systems. *Mol Cell.* Apr. 7, 2016;62(1):137-47. doi: 10.1016/j.molcel.2016.02.031. Epub Mar. 31, 2016.
- Leibundgut-Landmann et al., Mini-review: Specificity and expression of CIITA, the master regulator of MHC class II genes. *Eur J Immunol.* Jun. 2004;34(6):1513-25. doi: 10.1002/eji.200424964.
- Li et al., A dominant-negative form of mouse SOX2 induces trophectoderm differentiation and progressive polyploidy in mouse embryonic stem cells. *J Biol Chem.* Jul. 6, 2007;282(27):19481-92. doi: 10.1074/jbc.M702056200. Epub May 15, 2007.
- Li et al., Base-Resolution Mapping Reveals Distinct m1A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. *Mol Cell.* Dec. 7, 2017;68(5):993-1005.e9. doi: 10.1016/j.molcel.2017.10.019. Epub Nov. 5, 2017.
- Lim et al., Specific insertions of zinc finger domains into Gag-Pol yield engineered retroviral vectors with selective integration properties. *Proc Natl Acad Sci U S A.* Jul. 13, 2010;107(28):12475-80. doi: 10.1073/pnas.1001402107. Epub Jun. 28, 2010.
- Liu et al., Delivery methods for site-specific nucleases: Achieving the full potential of therapeutic gene editing. *J Control Release.* Dec. 28, 2016;244(Pt A):83-97. doi: 10.1016/j.jconrel.2016.11.014. Epub Nov. 16, 2016.
- Ma et al., Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. *Mol Ther Nucleic Acids.* May 6, 2014;3(5):e161. doi: 10.1038/mtna.2014.12.
- Martín et al., Envelope-targeted retrovirus vectors transduce melanoma xenografts but not spleen or liver. *Mol Ther.* Mar. 2002;5(3):269-74. doi: 10.1006/mthe.2002.0550.
- Mason et al., Coiled coil domains: stability, specificity, and biological implications. *Chembiochem.* Feb. 6, 2004;5(2):170-6. doi: 10.1002/cbic.200300781.
- Mercuri et al., Nusinersen versus Sham Control in Later-Onset Spinal Muscular Atrophy. *N Engl J Med.* Feb. 15, 2018;378(7):625-635. doi: 10.1056/NEJMoa1710504.
- Meunier et al., Drug-Induced Liver Injury: Biomarkers, Requirements, Candidates, and Validation. *Front Pharmacol.* Dec. 11, 2019;10:1482. doi: 10.3389/fphar.2019.01482.
- Nawaz et al., Extracellular Vesicles, Tunneling Nanotubes, and Cellular Interplay: Synergies and Missing Links. *Front Mol Biosci.* Jul. 18, 2017;4:50. doi: 10.3389/fmolsb.2017.00050.
- Nesbitt, Targeted Intracellular Therapeutic Delivery Using Liposomes Formulated with Multifunctional FAST proteins. Electronic Thesis and Dissertation Repository. The University of Western Ontario. 2012. 126 pages.
- Pan et al., Biodistribution and toxicity studies of VSVG-pseudotyped lentiviral vector after intravenous administration in mice with the observation of in vivo transduction of bone marrow. *Mol Ther.* Jul. 2002;6(1):19-29. doi: 10.1006/mthe.2002.0630.
- Pan et al., Identification of a nuclear localization signal in OCT4 and generation of a dominant negative mutant by its ablation. *J Biol Chem.* Aug. 27, 2004;279(35):37013-20. doi: 10.1074/jbc.M405117200. Epub Jun. 24, 2004.
- Pang et al., Retinal degeneration 12 (rd12): a new, spontaneously arising mouse model for human Leber congenital amaurosis (LCA). *Mol Vis.* Feb. 28, 2005;11:152-62.
- Parr-Brownlie et al., Lentiviral vectors as tools to understand central nervous system biology in mammalian model organisms. *Front Mol Neurosci.* May 18, 2015;8:14. doi: 10.3389/fnmol.2015.00014.
- Pastuzyn et al., The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. *Cell.* Jan. 11, 2018;172(1-2):275-288.e18. doi: 10.1016/j.cell.2017.12.024. Erratum in: *Cell.* Mar. 22, 2018;173(1):275. doi: 10.1016/j.cell.2018.03.024.
- Pavlov et al., Roles of DNA polymerases in replication, repair, and recombination in eukaryotes. *Int Rev Cytol.* 2006;255:41-132. doi: 10.1016/S0074-7696(06)55002-8.
- Podbilewicz, Virus and cell fusion mechanisms. *Annu Rev Cell Dev Biol.* 2014;30:111-39. doi: 10.1146/annurev-cellbio-101512-122422. Epub Jun. 27, 2014.
- Puppo et al., Retinal transduction profiles by high-capacity viral vectors. *Gene Ther.* Oct. 2014;21(10):855-65. doi: 10.1038/gt.2014.57. Epub Jul. 3, 2014.
- Qi et al., Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell.* Feb. 28, 2013;152(5):Supplementary Material. doi: 10.1016/j.cell.2013.02.022. 4 pages.
- Ramiro et al., Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat Immunol.* May 2003;4(5):452-6. doi: 10.1038/ni920.
- Rao et al., Large-Scale Phenome-Wide Association Study of PCSK9 Variants Demonstrates Protection Against Ischemic Stroke. *Circ Genom Precis Med.* Jul. 2018;11(7):e002162. doi: 10.1161/CIRCPGEN.118.002162.
- Remington et al., Complete nucleotide sequence of a neuropathogenic variant of Friend murine leukemia virus PVC-211. *Nucleic Acids Res.* Jun. 25, 1992;20(12):3249. doi: 10.1093/nar/20.12.3249.
- Romero et al., Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Biol.* Dec. 2009;10(12):866-76. doi: 10.1038/nrm2805.
- Sanjana et al., Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods.* Aug. 2014;11(8):783-784. doi: 10.1038/nmeth.3047.
- Sapir et al., Viral and developmental cell fusion mechanisms: conservation and divergence. *Dev Cell.* Jan. 2008;14(1):11-21. doi: 10.1016/j.devcel.2007.12.008.
- Schellekens, Bioequivalence and the immunogenicity of biopharmaceuticals. *Nat Rev Drug Discov.* Jun. 2002;1(6):457-62. doi: 10.1038/nrd818.
- Schneider et al., MuLV IN mutants responsive to HDAC inhibitors enhance transcription from unintegrated retroviral DNA. *Virology.* May 10, 2012;426(2):188-96. doi: 10.1016/j.virol.2012.01.034. Epub Feb. 23, 2012.
- Semple et al., Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Prog Neurobiol.* Jul.-Aug. 2013;106-107:1-16. doi: 10.1016/j.pneurobio.2013.04.001. Epub Apr. 11, 2013.
- Serreze et al., Major histocompatibility complex class I-deficient NOD-B2mnull mice are diabetes and insulitis resistant. *Diabetes.* Mar. 1994;43(3):505-9. doi: 10.2337/diab.43.3.505.
- Shen et al., Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled. *Proc Natl Acad Sci U S A.* Aug. 31, 2004;101(35):12997-3002. doi: 10.1073/pnas.0404974101. Epub Aug. 24, 2004.
- Skipper et al., Delivering the Goods for Genome Engineering and Editing. *Hum Gene Ther.* Aug. 2015;26(8):486-97. doi: 10.1089/hum.2015.063.
- Stadtmauer et al., CRISPR-engineered T cells in patients with refractory cancer. *Science.* Feb. 28, 2020;367(6481):eaba7365. doi: 10.1126/science.aba7365. Epub Feb. 6, 2020.
- Stevens et al., Design of a Split Intein with Exceptional Protein Splicing Activity. *J Am Chem Soc.* Feb. 24, 2016;138(7):2162-5. doi: 10.1021/jacs.5b13528. Epub Feb. 8, 2016. Abstract Only. 1 page.
- Swiech et al., In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol.* Jan. 2015;33(1):102-6. doi: 10.1038/nbt.3055. Epub Oct. 19, 2014. Author Manuscript. 22 pages.

(56)

References Cited**OTHER PUBLICATIONS**

- Taylor, Ocular immune privilege. *Eye (Lond)*. Oct. 2009;23(10):1885-9. doi: 10.1038/eye.2008.382. Epub Jan. 9, 2009.
- Thakore et al., Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*. Dec. 2015;12(12):1143-9. doi: 10.1038/nmeth.3630. Epub Oct. 26, 2015.
- Thorne et al., In vivo diffusion analysis with quantum dots and dextran predicts the width of brain extracellular space. *Proc Natl Acad Sci U S A*. Apr. 4, 2006;103(14):5567-72. doi: 10.1073/pnas.0509425103. Epub Mar. 27, 2006.
- Tokuriki et al., Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol*. Oct. 2009;19(5):596-604. doi: 10.1016/j.sbi.2009.08.003. Epub Sep. 16, 2009.
- Tőzsér, Comparative studies on retroviral proteases: substrate specificity. *Viruses*. Jan. 2010;2(1):147-165. doi: 10.3390/v2010147. Epub Jan. 14, 2010.
- Urano et al., Substitution of the myristoylation signal of human immunodeficiency virus type 1 Pr55Gag with the phospholipase C-delta1 pleckstrin homology domain results in infectious pseudovirion production. *J Gen Virol*. Dec. 2008;89(Pt 12):3144-3149. doi: 10.1099/vir.0.2008/004820-0.
- Voisset et al., Phylogeny of a novel family of human endogenous retrovirus sequences, HERV-W, in humans and other primates. *AIDS Res Hum Retroviruses*. Nov. 20, 1999;15(17):1529-33. doi: 10.1089/088922299309810.
- Wang et al., Characterization of an Mps I-H knock-in mouse that carries a nonsense mutation analogous to the human IDUA-W402X mutation. *Mol Genet Metab*. Jan. 2010;99(1):62-71. doi: 10.1016/j.ymgme.2009.08.002. Erratum in: *Mol Genet Metab*. Apr. 2010;99(4):439.
- Wang et al., Influence of the polyanion on the physico-chemical properties and biological activities of polyanion/DNA/polycation ternary polyplexes. *Acta Biomater*. Aug. 2012;8(8):3014-26. doi: 10.1016/j.actbio.2012.04.034. Epub Apr. 27, 2012.
- Webber et al., Highly efficient multiplex human T cell engineering without double-strand breaks using Cas9 base editors. *Nat Commun*. Nov. 19, 2019;10(1):5222. doi: 10.1038/s41467-019-13007-6. Erratum in: *Nat Commun*. Dec. 6, 2019;10(1):5659. doi: 10.1038/s41467-019-13778-y.
- Wheeler et al., Proteomics analysis of cellular components in lentiviral vector production using Gel-LC-MS/MS. *Proteomics Clin Appl*. Feb. 2007;1(2):224-30. doi: 10.1002/pcra.200600522. Epub Jan. 22, 2007.
- Wu et al., Effect of genome size on AAV vector packaging. *Mol Ther*. Jan. 2010;18(1):80-6. doi: 10.1038/mt.2009.255. Epub Nov. 10, 2009.
- Xu et al., Cas9-based tools for targeted genome editing and transcriptional control. *Appl Environ Microbiol*. Mar. 2014;80(5):1544-52. doi: 10.1128/AEM.03786-13. Epub Jan. 3, 2014.
- Xu et al., Sequence and structural analyses of nuclear export signals in the NESdb database. *Mol Biol Cell*. Sep. 2012;23(18):3677-93. doi: 10.1091/mbc.E12-01-0046. Epub Jul. 25, 2012.
- Zhang et al., Morphology and ultrastructure of retrovirus particles. *AIMS Biophys*. 2015;2(3):343-369. doi: 10.3934/biophy.2015.3.343. Epub Aug. 18, 2015. Author Manuscript. 33 pages.
- Zhong et al., Seven novel variants expand the spectrum of RPE65-related Leber congenital amaurosis in the Chinese population. *Mol Vis*. Mar. 18, 2019;25:204-214.
- [No Author Listed], CMP/dCMP-type deaminase domain-containing protein. Uniprot Accession No. A0A2Z6RZE9. Oct. 10, 2018. Accessible at <https://www.uniprot.org/uniprotkb/A0A2Z6RZE9>. entry. 8 pages.
- [No Author Listed], dCas9-5xPlat2AfID-P2A-scFvGCN4sfGFPTET1CD [Cloning vector pPlatTET-gRNA2]. GenBank No. BAV54124. Apr. 18, 2017. 5 pages.
- [No Author Listed], tRNA-specific adenosine deaminase [Candidatus Moranella endobia PCVAL]. GenBank Acc. No. AGJ61179.1. Accessible at <https://www.ncbi.nlm.nih.gov/protein/AGJ61179>. Jan. 3, 2014. 3 pages.
- [No Author Listed], tRNA-specific adenosine deaminase 2 [Terrapene triunguis]. GenBank Acc. No. XP_024075810.1. Accessible at https://www.ncbi.nlm.nih.gov/protein/XP_024075810. Jul. 15, 2019. 2 pages.
- [No Author Listed], tRNA-specific adenosine deaminase TAD2 isoform X2 [Panicum hallii]. GenBank Acc. No. XP_025793740.1. Accessible at <https://www.ncbi.nlm.nih.gov/protein/025793740>. Jul. 27, 2018. 1 page.
- [No Author Listed], tRNA-specific adenosine deaminase TAD2 isoform X1 [*Oryza sativa* Japonica Group]. GenBank Acc. No. XP_15631651.1. Accessible at <https://www.ncbi.nlm.nih.gov/protein/1002254769?sat=58&satkey=133677684>. Aug. 7, 2018. 2 pages.
- [No Author Listed], tRNA-specific adenosine deaminase. Uniprot Accession No. A0A1U7M801. May 10, 2017. Accessible at <https://www.uniprot.org/uniprotkb/A0A1U7M801/history>. 3 pages.
- [No Author Listed], tRNA-specific adenosine deaminase. Uniprot Accession No. A0A1Z4VPW4. Sep. 27, 2017. Accessible at <https://www.uniprot.org/uniprotkb/A0A1Z4VPW4/history>. 3 pages.
- [No Author Listed], tRNA-specific adenosine deaminase. Uniprot Accession No. A0A1Z9LY19. Oct. 25, 2017. Accessible at <https://www.uniprot.org/uniprotkb/A0A1Z9LY19/entry>. 12 pages.
- [No Author Listed], tRNA-specific adenosine deaminase. Uniprot Accession No. A0A2P5TOZ9. May 23, 2018. Accessible at <https://www.uniprot.org/uniprotkb/A0A2P5TOZ9/entry>. 10 pages.
- [No Author Listed], tRNA-specific adenosine deaminase. Uniprot Accession No. A0A4P6PH16. Jul. 31, 2019. Accessible at <https://www.uniprot.org/uniprotkb/A0A4P6PH16/entry>. 12 pages.
- [No Author Listed], tRNA-specific adenosine deaminase. Uniprot Accession No. A0A520SVM3. Oct. 16, 2019. Accessible at <https://www.uniprot.org/uniprotkb/A0A520SVM3/entry>. 10 pages.
- [No Author Listed], tRNA-specific adenosine deaminase. Uniprot Accession No. U2JUU0. Nov. 13, 2013. Accessible at <https://www.uniprot.org/uniprotkb/U2JUU0/entry>. 11 pages.
- Alves et al., Immunogenicity of the carcinoembryonic antigen derived peptide 694 in HLA-A2 healthy donors and colorectal carcinoma patients. *Cancer Immunol Immunother*. Nov. 2007;56(11):1795-805. doi: 10.1007/s00262-007-0323-2. Epub Apr. 20, 2007.
- Asemissem et al., Identification of a highly immunogenic HLA-A*01-binding T cell epitope of WT1. *Clin Cancer Res*. Dec. 15, 2006;12(24):7476-82. doi: 10.1158/1078-0432.CCR-06-1337.
- Attia et al., Autoimmunity correlates with tumor regression in patients with metastatic melanoma treated with anti-cytotoxic T-lymphocyte antigen-4. *J Clin Oncol*. Sep. 1, 2005;23(25):6043-53. doi: 10.1200/JCO.2005.06.205. Epub Aug. 8, 2005.
- Aurisicchio et al., A novel minigene scaffold for therapeutic cancer vaccines. *Oncoimmunology*. Jan. 1, 2014;3(1):e27529. doi: 10.4161/onci.27529. Epub Jan. 16, 2014.
- Bae et al., Identification of novel CD33 antigen-specific peptides for the generation of cytotoxic T lymphocytes against acute myeloid leukemia. *Cell Immunol*. Jan. 2004;227(1):38-50. doi: 10.1016/j.cellimm.2004.01.002.
- Bakker et al., Analogs of CTL epitopes with improved MHC class-I binding capacity elicit anti-melanoma CTL recognizing the wild-type epitope. *Int J Cancer*. Jan. 27, 1997;70(3):302-9. doi: 10.1002/(sici)1097-0215(19970127)70:3<302::aid-ijc10>3.0.co;2-h.
- Banerjee et al., Viral glycoproteins: biological role and application in diagnosis. *Virusdisease*. Mar. 2016;27(1):1-11. doi: 10.1007/s13337-015-0293-5. Epub Jan. 18, 2016.
- Barve et al., Induction of immune responses and clinical efficacy in a phase II trial of IDM-2101, a 10-epitope cytotoxic T-lymphocyte vaccine, in metastatic non-small-cell lung cancer. *J Clin Oncol*. Sep. 20, 2008;26(27):4418-25. doi: 10.1200/JCO.2008.16.6462.
- Benlalam et al., Identification of five new HLA-B*3501-restricted epitopes derived from common melanoma-associated antigens, spontaneously recognized by tumor-infiltrating lymphocytes. *J Immunol*. Dec. 1, 2003;171(11):6283-9. doi: 10.4049/jimmunol.171.11.6283.
- Bernatchez et al., Altered decamer and nonamer from an HLA-A0201-restricted epitope of Survivin differentially stimulate T-cell responses in different individuals. *Vaccine*. Apr. 5, 2011;29(16):3021-30. doi: 10.1016/j.vaccine.2011.01.115. Epub Feb. 12, 2011.

(56)

References Cited**OTHER PUBLICATIONS**

- Bioley et al., Melan-A/MART-1-specific CD4 T cells in melanoma patients: identification of new epitopes and ex vivo visualization of specific T cells by MHC class II tetramers. *J Immunol.* Nov. 15, 2006;177(10):6769-79. doi: 10.4049/jimmunol.177.10.6769.
- Blanchet et al., A new generation of Melan-A/MART-1 peptides that fulfill both increased immunogenicity and high resistance to biodegradation: implication for molecular anti-melanoma immunotherapy. *J Immunol.* Nov. 15, 2001;167(10):5852-61. doi: 10.4049/jimmunol.167.10.5852.
- Borbulevych et al., Increased immunogenicity of an anchor-modified tumor-associated antigen is due to the enhanced stability of the peptide/MHC complex: implications for vaccine design. *J Immunol.* Apr. 15, 2005;174(8):4812-20. doi: 10.4049/jimmunol.174.8.4812.
- Brichard et al., A tyrosinase nonapeptide presented by HLA-B44 is recognized on a human melanoma by autologous cytolytic T lymphocytes. *Eur J Immunol.* Jan. 1996;26(1):224-30. doi: 10.1002/eji.1830260135.
- Cacabatos et al., Chapter 1—The Epigenetic Machinery in the Life Cycle and Pharmacogenetics. *Pharmacogenetics.* vol. 10 in *Translational Epigenetics.* 2019:1-100. doi: <https://doi.org/10.1016/B978-0-12-813939-4.00001-2>. 7 pages.
- Campbell et al., Gesicle-Mediated Delivery of CRISPR/Cas9 Ribonucleoprotein Complex for Inactivating the HIV Provirus. *Mol Ther.* Jan. 2, 2019;27(1):151-163. doi: 10.1016/j.ymthe.2018.10.002. Epub Oct. 11, 2018.
- Campi et al., CD4(+) T cells from healthy subjects and colon cancer patients recognize a carcinoembryonic antigen-specific immunodominant epitope. *Cancer Res.* Dec. 1, 2003;63(23):8481-6.
- Casnici et al., Immunologic evaluation of peptides derived from BCR/ABL-out-of-frame fusion protein in HLA A2.1 transgenic mice. *J Immunother.* May 2012;35(4):321-8. doi: 10.1097/CJI.0b013e3182562d37.
- Casnici et al., Out of frame peptides from BCR/ABL alternative splicing are immunogenic in HLA A2.1 transgenic mice. *Cancer Lett.* Apr. 8, 2009;276(1):61-7. doi: 10.1016/j.canlet.2008.10.032. Epub Dec. 4, 2008.
- Castelli et al., Mass spectrometric identification of a naturally processed melanoma peptide recognized by CD8+ cytotoxic T lymphocytes. *J Exp Med.* Jan. 1, 1995;181(1):363-8. doi: 10.1084/jem.181.1.363.
- Castelli et al., Novel HLA-Cw8-restricted T cell epitopes derived from tyrosinase-related protein-2 and gp100 melanoma antigens. *J Immunol.* Feb. 1, 1999;162(3):1739-48.
- Castle et al., Exploiting the mutanome for tumor vaccination. *Cancer Res.* Mar. 1, 2012;72(5):1081-91. doi: 10.1158/0008-5472.CAN-11-3722. Epub Jan. 11, 2012.
- Cervera et al., Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J Biotechnol.* Jul. 20, 2013;166(4):152-65. doi: 10.1016/j.jbiotec.2013.05.001. Epub May 17, 2013.
- Chen et al., Identification of NY-ESO-1 peptide analogues capable of improved stimulation of tumor-reactive Ctl. *J Immunol.* Jul. 15, 2000;165(2):948-55. doi: 10.4049/jimmunol.165.2.948.
- Cho et al., Optimized peptide vaccines eliciting extensive CD8 T-cell responses with therapeutic antitumor effects. *Cancer Res.* Dec. 1, 2009;69(23):9012-9. doi: 10.1158/0008-5472.CAN-09-2019. Epub Nov. 10, 2009.
- Choi et al., Lentivirus pre-packed with Cas9 protein for safer gene editing. *Gene Ther.* Jul. 2016;23(7):627-33. doi: 10.1038/gt.2016.27. Epub Apr. 7, 2016.
- Christensen et al., Melan-A/MART1 analog peptide triggers anti-myeloma T-cells through crossreactivity with HM1.24. *J Immunother.* Jul.-Aug. 2009;32(6):613-21. doi: 10.1097/CJI.0b013e3181a95198.
- Correale et al., In vitro generation of human cytotoxic T lymphocytes specific for peptides derived from prostate-specific antigen. *J Natl Cancer Inst.* Feb. 19, 1997;89(4):293-300. doi: 10.1093/jnci/89.4.293.
- Courtney et al., CRISPR/Cas9 DNA cleavage at SNP-derived PAM enables both in vitro and in vivo KRT12 mutation-specific targeting. *Gene Ther.* Jan. 2016;23(1):108-12. doi: 10.1038/gt.2015.82. Epub Aug. 20, 2015.
- Cox et al., Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. *Science.* Apr. 29, 1994;264(5159):716-9. doi: 10.1126/science.7513441.
- Cronin et al., Altering the tropism of lentiviral vectors through pseudotyping. *Curr Gene Ther.* Aug. 2005;5(4):387-98. doi: 10.2174/1566523054546224. Erratum in: *Curr Gene Ther.* Oct. 2005;5(5):531. Author Manuscript, 19 pages.
- Crosti et al., Identification of novel subdominant epitopes on the carcinoembryonic antigen recognized by CD4+ T cells of lung cancer patients. *J Immunol.* Apr. 15, 2006;176(8):5093-9. doi: 10.4049/jimmunol.176.8.5093.
- Dalet et al., An antigenic peptide produced by reverse splicing and double asparagine deamidation. *Proc Natl Acad Sci U S A.* Jul. 19, 2011;108(29):E323-31. doi: 10.1073/pnas.1101892108. Epub Jun. 13, 2011.
- David et al., Viral Vectors: The Road to Reducing Genotoxicity. *Toxicol Sci.* Feb. 2017;155(2):315-325. doi: 10.1093/toxsci/kfw220. Epub Nov. 1, 2016.
- Deponteu et al., Identification of tumor-associated, MHC class II-restricted phosphopeptides as targets for immunotherapy. *Proc Natl Acad Sci U S A.* Jul. 21, 2009;106(29):12073-8. doi: 10.1073/pnas.0903852106. Epub Jul. 6, 2009.
- Di Stasi et al., Review of the Results of WT1 Peptide Vaccination Strategies for Myelodysplastic Syndromes and Acute Myeloid Leukemia from Nine Different Studies. *Front Immunol.* Feb. 4, 2015;6:36. doi: 10.3389/fimmu.2015.00036.
- Duan et al., Immune rejection of mouse tumors expressing mutated self. *Cancer Res.* Apr. 15, 2009;69(8):3545-53. doi: 10.1158/0008-5472.CAN-08-2779. Epub Apr. 7, 2009. Author Manuscript. 18 pages.
- Duportet et al., A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Res.* Dec. 1, 2014;42(21):13440-51. doi: 10.1093/nar/gku1082. Epub Nov. 5, 2014.
- Fontana et al., Rabies virus-like particles expressed in HEK293 cells. *Vaccine.* May 19, 2014;32(24):2799-804. doi: 10.1016/j.vaccine.2014.02.031. Epub Mar. 12, 2014.
- Fonteneau et al., The Tumor Antigen NY-ESO-1 Mediates Direct Recognition of Melanoma Cells by CD4+ T Cells after Intercellular Antigen Transfer. *J Immunol.* Jan. 1, 2016;196(1):64-71. doi: 10.4049/jimmunol.1402664. Epub Nov. 25, 2015.
- Fourcade et al., PD-1 and Tim-3 regulate the expansion of tumor antigen-specific CD8⁺ T cells induced by melanoma vaccines. *Cancer Res.* Feb. 15, 2014;74(4):1045-55. doi: 10.1158/0008-5472.CAN-13-2908. Epub Dec. 16, 2013.
- Fridman et al., An efficient T-cell epitope discovery strategy using in silico prediction and the iTopia assay platform. *Oncimmunology.* Nov. 1, 2012;1(8):1258-1270. doi: 10.4161/onci.21355.
- Fujiki et al., Identification and characterization of a WT1 (Wilms Tumor Gene) protein-derived HLA-DRB1*0405-restricted 16-mer helper peptide that promotes the induction and activation of WT1-specific cytotoxic T lymphocytes. *J Immunother.* Apr. 2007;30(3):282-93. doi: 10.1097/01.cji.0000211337.91513.94.
- Gee et al., Extracellular nanovesicles for packaging of CRISPR-Cas9 protein and sgRNA to induce therapeutic exon skipping. *Nat Commun.* Mar. 13, 2020;11(1):1334. doi: 10.1038/s41467-020-14957-y.
- GenBank Access No. BAP64357. Aug. 1, 2013. 1 page.
- Geynisman et al., A randomized pilot phase I study of modified carcinoembryonic antigen (CEA) peptide (CAP1-6D)/montanide/GM-CSF-vaccine in patients with pancreatic adenocarcinoma. *J Immunother Cancer.* Jun. 27, 2013;1:8. doi: 10.1186/2051-1426-1-8.
- Ghosh et al., Synapsis in phage Bxb1 integration: selection mechanism for the correct pair of recombination sites. *J Mol Biol.* Jun. 3, 2005;349(2):331-48. doi: 10.1016/j.jmb.2005.03.043. Epub Apr. 7, 2005.
- Girard-Gagnepain et al., Baboon envelope pseudotyped LVs outperform VSV-G-LVs for gene transfer into early-cytokine-

(56)

References Cited**OTHER PUBLICATIONS**

- stimulated and resting HSCs. *Blood*. Aug. 21, 2014;124(8):1221-31. doi: 10.1182/blood-2014-02-558163. Epub Jun. 20, 2014.
- Godefroy et al., Identification of two Melan-A CD4+ T cell epitopes presented by frequently expressed MHC class II alleles. *Clin Immunol*. Oct. 2006;121(1):54-62. doi: 10.1016/j.clim.2006.05.007. Epub Jun. 30, 2006.
- Graff-Dubois et al., Generation of CTL recognizing an HLA-A*0201-restricted epitope shared by MAGE-A1, -A2, -A3, -A4, -A6, -A10, and -A12 tumor antigens: implication in a broad-spectrum tumor immunotherapy. *J Immunol*. Jul. 1, 2002;169(1):575-80. doi: 10.4049/jimmunol.169.1.575.
- Gross et al., High vaccination efficiency of low-affinity epitopes in antitumor immunotherapy. *J Clin Invest*. Feb. 2004;113(3):425-33. doi: 10.1172/JCI19418.
- Guevara-Patiño et al., Optimization of a self antigen for presentation of multiple epitopes in cancer immunity. *J Clin Invest*. May 2006;116(5):1382-90. doi: 10.1172/JCI25591. Epub Apr. 13, 2006.
- Guibinga et al., Cell surface heparan sulfate is a receptor for attachment of envelope protein-free retrovirus-like particles and VSV-G pseudotyped MLV-derived retrovirus vectors to target cells. *Mol Ther*. May 2002;5(Pt 1):538-46. doi: 10.1006/mthe.2002.0578.
- Gulley et al., Combining a Recombinant Cancer Vaccine with Standard Definitive Radiotherapy in Patients with Localized Prostate Cancer. *Clin Cancer Res*. May 2, 2005;11(9):3353-62. doi: 10.1158/1078-0432.CCR-04-2062.
- Guo et al., Direct recognition and lysis of leukemia cells by WT1-specific CD4+ T lymphocytes in an HLA class II-restricted manner. *Blood*. Aug. 15, 2005;106(4): 1415-8. doi: 10.1182/blood-2005-01-0413. Epub Apr. 21, 2005.
- Haeussler et al., Genome Editing with CRISPR-Cas9: Can It Get Any Better? *J Genet Genomics*. May 20, 2016;43(5):239-50. doi: 10.1016/j.jgg.2016.04.008. Epub Apr. 24, 2016. Author Manuscript. 22 pages.
- Herbst-Kralovetz et al., Norwalk virus-like particles as vaccines. *Expert Rev Vaccines*. Mar. 2010;9(3):299-307. doi: 10.1586/erv.09.163. Author Manuscript, 16 pages.
- Hirohashi et al., An HLA-A24-restricted cytotoxic T lymphocyte epitope of a tumor-associated protein, survivin. *Clin Cancer Res*. Jun. 2002;8(6):1731-9.
- Hong et al., Novel recombinant hepatitis B virus vectors efficiently deliver protein and RNA encoding genes into primary hepatocytes. *J Virol*. Jun. 2013;87(12):6615-24. doi: 10.1128/JVI.03328-12. Epub Apr. 3, 2013.
- Jalaguier et al., Efficient production of HIV-1 virus-like particles from a mammalian expression vector requires the N-terminal capsid domain. *PLoS One*. 2011;6(11):e28314. doi: 10.1371/journal.pone.0028314. Epub Nov. 30, 2011.
- Jaramillo et al., Identification of HLA-A3-restricted CD8+ T cell epitopes derived from gammaglobin-A, a tumor-associated antigen of human breast cancer. *Int J Cancer*. Dec. 10, 2002;102(5):499-506. doi: 10.1002/ijc.10736.
- Joglekar et al., Pseudotyped Lentiviral Vectors: One Vector, Many Guises. *Hum Gene Ther Methods*. Dec. 2017;28(6):291-301. doi: 10.1089/hgtb.2017.084. Epub Sep. 4, 2017.
- Kaczmarczyk et al., Protein delivery using engineered virus-like particles. *Proc Natl Acad Sci U S A*. Oct. 11, 2011;108(41):16998-7003. doi: 10.1073/pnas.1101874108. Epub Sep. 26, 2011.
- Kang et al., Chimeric rabies virus-like particles containing membrane-anchored GM-CSF enhances the immune response against rabies virus. *Viruses*. Mar. 11, 2015;7(3):1134-52. doi: 10.3390/v7031134.
- Kang et al., Identification of a tyrosinase epitope recognized by HLA-A24-restricted, tumor-infiltrating lymphocytes. *J Immunol*. Aug. 1, 1995;155(3):1343-8.
- Karbach et al., Long-term complete remission following radiosurgery and immunotherapy in a melanoma patient with brain metastasis: immunologic correlates. *Cancer Immunol Res*. May 2014;2(5):404-9. doi: 10.1158/2326-6066.CIR-13-0200. Epub Feb. 5, 2014.
- Kato et al., A lentiviral strategy for highly efficient retrograde gene transfer by pseudotyping with fusion envelope glycoprotein. *Hum Gene Ther*. Feb. 2011;22(2):197-206. doi: 10.1089/hum.2009.179. Epub Jan. 27, 2011.
- Kato et al., Selective neural pathway targeting reveals key roles of thalamostriatal projection in the control of visual discrimination. *J Neurosci*. Nov. 23, 2011;31(47):17169-79. doi: 10.1523/JNEUROSCI.4005-11.2011.
- Kawakami et al., Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection. *Proc Natl Acad Sci U S A*. Jul. 5, 1994;91(14):6458-62. doi: 10.1073/pnas.91.14.6458.
- Kawakami et al., Identification of new melanoma epitopes on melanosomal proteins recognized by tumor infiltrating T lymphocytes restricted by HLA-A1, -A2, and -A3 alleles. *J Immunol*. Dec. 15, 1998;161(12):6985-92.
- Kawakami et al., Identification of the immunodominant peptides of the MART-1 human melanoma antigen recognized by the majority of HLA-A2-restricted tumor infiltrating lymphocytes. *J Exp Med*. Jul. 1, 1994;180(1):347-52. doi: 10.1084/jem.180.1.347.
- Kawakami et al., Recognition of multiple epitopes in the human melanoma antigen gp100 by tumor-infiltrating T lymphocytes associated with in vivo tumor regression. *J Immunol*. Apr. 15, 1995;154(8):3961-8.
- Kawashima et al., Identification of gp100-derived, melanoma-specific cytotoxic T-lymphocyte epitopes restricted by HLA-A3 supertype molecules by primary in vitro immunization with peptide-pulsed dendritic cells. *Int J Cancer*. Nov. 9, 1998;78(4):518-24. doi: 10.1002/(sici)1097-0215(19981109)78:4<518::aid-ijc20>3.0.co;2-0.
- Kawashima et al., Identification of HLA-A3-restricted cytotoxic T lymphocyte epitopes from carcinoembryonic antigen and HER-2/neu by primary in vitro immunization with peptide-pulsed dendritic cells. *Cancer Res*. Jan. 15, 1999;59(2):431-5.
- Kawashima et al., The multi-epitope approach for immunotherapy for cancer: identification of several CTL epitopes from various tumor-associated antigens expressed on solid epithelial tumors. *Hum Immunol*. Jan. 1998;59(1):1-14. doi: 10.1016/s0198-8859(97)00255-3.
- Kemmler et al., Elevated tumor-associated antigen expression suppresses variant peptide vaccine responses. *J Immunol*. Nov. 1, 2011;187(9):4431-9. doi: 10.4049/jimmunol.1101555. Epub Sep. 21, 2011.
- Kittlesen et al., Human melanoma patients recognize an HLA-A1-restricted CTL epitope from tyrosinase containing two cysteine residues: implications for tumor vaccine development. *J Immunol*. Mar. 1, 1998;160(5):2099-106. Erratum in: *J Immunol* Mar. 1, 1999;162(5):3106. Shabanowitz JA [corrected to Shabanowitz J].
- Kizer et al., Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl Environ Microbiol*. May 2008;74(10):3229-41. doi: 10.1128/AEM.02750-07. Epub Mar. 14, 2008.
- Kobayashi et al., CD4+ T cells from peripheral blood of a melanoma patient recognize peptides derived from nonmutated tyrosinase. *Cancer Res*. Jan. 15, 1998;58(2):296-301.
- Kobayashi et al., Identification of an antigenic epitope for helper T lymphocytes from carcinoembryonic antigen. *Clin Cancer Res*. Oct. 2002;8(10):3219-25.
- Kobayashi et al., Identification of helper T-cell epitopes that encompass or lie proximal to cytotoxic T-cell epitopes in the gp100 melanoma tumor antigen. *Cancer Res*. Oct. 15, 2001;61(20):7577-84.
- Kotterman et al., Engineering adeno-associated viruses for clinical gene therapy. *Nat Rev Genet*. Jul. 2014;15(7):445-51. doi: 10.1038/nrg3742. Epub May 20, 2014.
- Kueh et al., The new editor-targeted genome engineering in the absence of homology-directed repair. *Cell Death Discov*. Jun. 13, 2016;2:16042. doi: 10.1038/cddiscovery.2016.42.
- Kushnir et al., Virus-like particles as a highly efficient vaccine platform: diversity of targets and production systems and advances in clinical development. *Vaccine*. Dec. 17, 2012;31(1):58-83. doi: 10.1016/j.vaccine.2012.10.083. Epub Nov. 6, 2012.

(56)

References Cited**OTHER PUBLICATIONS**

- Lally et al., Unmasking cryptic epitopes after loss of immunodominant tumor antigen expression through epitope spreading. *Int J Cancer.* Sep. 2001;93(6):841-7. doi: 10.1002/ijc.1420.
- Lapointe et al., Retrovirally transduced human dendritic cells can generate T cells recognizing multiple MHC class I and class II epitopes from the melanoma antigen glycoprotein 100. *J Immunol.* Oct. 15, 2001;167(8):4758-64. doi: 10.4049/jimmunol.167.8.4758.
- Larrieu et al., A HLA-Cw*0701 restricted Melan-A/MART1 epitope presented by melanoma tumor cells to CD8+ tumor infiltrating lymphocytes. *Cancer Immunol Immunother.* May 2008;57(5):745-52. doi: 10.1007/s00262-007-0436-7. Epub Dec. 21, 2007.
- Latham et al., Formation of wild-type and chimeric influenza virus-like particles following simultaneous expression of only four structural proteins. *J Virol.* Jul. 2001;75(13):6154-65. doi: 10.1128/JVI.75.13.6154-6165.2001.
- Lennerz et al., The response of autologous T cells to a human melanoma is dominated by mutated neoantigens. *Proc Natl Acad Sci U S A.* Nov. 1, 2005;102(44):16013-8. doi: 10.1073/pnas.050090102. Epub Oct. 24, 2005.
- Li et al., Expression and self-assembly of empty virus-like particles of hepatitis E virus. *J Virol.* Oct. 1997;71(10):7207-13. doi: 10.1128/JVI.71.10.7207-7213.1997.
- Lin et al., HLA-DPB1*05: 01-restricted WT1332-specific TCR-transduced CD4+ T lymphocytes display a helper activity for WT1-specific CTL induction and a cytotoxicity against leukemia cells. *J Immunother.* Apr. 2013;36(3):159-70. doi: 10.1097/CJI.0b013e318273581.
- Lu, Periodic Chart of Amino Acid PDF. Accessed on the internet at https://figshare.com/articles/figure/periodic_chart_of_amino_acid_pdf/3445001/1. Posted Jun. 21, 2016. www.bachem.com. 1 page.
- Ludwig et al., Virus-like particles-universal molecular toolboxes. *Curr Opin Biotechnol.* Dec. 2007;18(6):537-45. doi: 10.1016/j.copbio.2007.10.013.
- Lueck et al., Engineered transfer RNAs for suppression of premature termination codons. *Nat Commun.* Feb. 18, 2019;10(1):822. doi: 10.1038/s41467-019-108329-4.
- Lupetti et al., Translation of a retained intron in tyrosinase-related protein (TRP) 2 mRNA generates a new cytotoxic T lymphocyte (CTL)-defined and shared human melanoma antigen not expressed in normal cells of the melanocytic lineage. *J Exp Med.* Sep. 21, 1998;188(6):1005-16. doi: 10.1084/jem.188.6.1005.
- Lyu et al., Delivering Cas9/sgRNA ribonucleoprotein (RNP) by lentiviral capsid-based bionanoparticles for efficient 'hit-and-run' genome editing. *Nucleic Acids Res.* Sep. 26, 2019;47(17):e99. doi: 10.1093/nar/gkz605.
- Maetzig et al., Retroviral protein transfer: falling apart to make an impact. *Curr Gene Ther.* Oct. 2012;12(5):389-409. doi: 10.2174/156652312802762581.
- Mandic et al., The alternative open reading frame of LAGE-1 gives rise to multiple promiscuous HLA-DR-restricted epitopes recognized by T-helper 1-type tumor-reactive CD4+ T cells. *Cancer Res.* Oct. 1, 2003;63(19):6506-15.
- Mangeot et al., A universal transgene silencing method based on RNA interference. *Nucleic Acids Res.* Jul. 12, 2004;32(12):e102. doi: 10.1093/nar/gnh105.
- Mangeot et al., Development of minimal lentivirus vectors derived from simian immunodeficiency virus (SIVmac251) and their use for gene transfer into human dendritic cells. *J Virol.* Sep. 2000;74(18):8307-15. doi: 10.1128/jvi.74.18.8307-8315.2000.
- Mangeot et al., Protein transfer into human cells by VSV-G-induced nanovesicles. *Mol Ther.* Sep. 2011;19(9):1656-66. doi: 10.1038/mt.2011.138. Epub Jul. 12, 2011.
- Mariani et al., Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell.* Jul. 11, 2003;114(1):21-31. doi: 10.1016/s0092-8674(03)00515-4.
- Meng et al., Identification of an HLA-DPB1*0501 restricted Melan-A/MART-1 epitope recognized by CD4+ T lymphocytes: prevalence for immunotherapy in Asian populations. *J Immunother.* Sep. 2011;34(7):525-34. doi: 10.1097/CJI.0b013e318226bd45. Author Manuscript. 16 pages.
- Michaux et al., A spliced antigenic peptide comprising a single spliced amino acid is produced in the proteasome by reverse splicing of a longer peptide fragment followed by trimming. *J Immunol.* Feb. 15, 2014;192(4):1962-71. doi: 10.4049/jimmunol.1302032. Epub Jan. 22, 2014.
- Milone et al., Clinical use of lentiviral vectors. *Leukemia.* Jul. 2018;32(7):1529-1541. doi: 10.1038/s41375-018-0106-0. Epub Mar. 22, 2018.
- Momose et al., Diving into marine genomics with CRISPR/Cas9 systems. *Mar Genomics.* Dec. 2016;30:55-65. doi: 10.1016/j.margen.2016.10.003. Epub Oct. 12, 2016.
- Morel et al., A tyrosinase peptide presented by HLA-B35 is recognized on a human melanoma by autologous cytotoxic T lymphocytes. *Int J Cancer.* Dec. 10, 1999;83(6):755-9. doi: 10.1002/(sici)1097-0215(19991210)83:6<755::aid-ijc10>3.0.co;2-s.
- Mselli-Lakhal et al., Gene transfer system derived from the caprine arthritis-encephalitis lentivirus. *J Virol Methods.* Sep. 2006;136(1-2):177-84. doi: 10.1016/j.jviromet.2006.05.006. Epub Jun. 21, 2006.
- Murawski et al., Newcastle disease virus-like particles containing respiratory syncytial virus G protein induced protection in BALB/c mice, with no evidence of immunopathology. *J Virol.* Jan. 2010;84(2):1110-23. doi: 10.1128/JVI.01709-09. Epub Nov. 4, 2009.
- Naskalska et al., Virus Like Particles as Immunogens and Universal Nanocarriers. *Pol J Microbiol.* 2015;64(1):3-13.
- Negre et al., Characterization of novel safe lentiviral vectors derived from simian immunodeficiency virus (SIVmac251) that efficiently transduce mature human dendritic cells. *Gene Ther.* Oct. 2000;7(19):1613-23. doi: 10.1038/sj.gt.3301292.
- Nuppen et al., Naturally processed and concealed HLA-A2.1-restricted epitopes from tumor-associated antigen tyrosinase-related protein-2. *Int J Cancer.* Jul. 15, 2000;87(2):241-6.
- Nukaya et al., Identification of HLA-A24 epitope peptides of carcinoembryonic antigen which induce tumor-reactive cytotoxic T lymphocyte. *Int J Cancer.* Jan. 5, 1999;80(1):92-7. doi: 10.1002/(sici)1097-0215(19990105)80:1<92::aid-ijc18>3.0.co;2-m.
- Ogasawara et al., Recombinant viral-like particles of parvovirus B19 as antigen carriers of anthrax protective antigen. *In Vivo.* May-Jun. 2006;20(3):319-24.
- Ohminami et al., HLA class I-restricted lysis of leukemia cells by a CD8(+) cytotoxic T-lymphocyte clone specific for WT1 peptide. *Blood.* Jan. 1, 2000;95(1):286-93.
- Oka et al., WT1 peptide vaccine for the treatment of cancer. *Curr Opin Immunol.* Apr. 2008;20(2):211-20. doi: 10.1016/j.co.2008.04.009. Epub May 24, 2008.
- Olsen, J.C., Gene transfer vectors derived from equine infectious anemia virus. *Gene Ther.* Nov. 1998;5(11):1481-7. doi: 10.1038/sj.gt.3300768.
- Olson et al., HLA-A2-restricted T-cell epitopes specific for prostatic acid phosphatase. *Cancer Immunol Immunother.* Jun. 2010;59(6):943-53. doi: 10.1007/s00262-010-0820-6. Epub Feb. 6, 2010.
- Osen et al., Screening of human tumor antigens for CD4 T cell epitopes by combination of HLA-transgenic mice, recombinant adenovirus and antigen peptide libraries. *PLoS One.* Nov. 30, 2010;5(11):e14137. doi: 10.1371/journal.pone.0014137.
- Parkhurst et al., Identification of a shared HLA-A*0201-restricted T-cell epitope from the melanoma antigen tyrosinase-related protein 2 (TRP2). *Cancer Res.* Nov. 1, 1998;58(21):4895-901.
- Parkhurst et al., Induction of CD4+ Th1 lymphocytes that recognize known and novel class II MHC restricted epitopes from the melanoma antigen gp100 by stimulation with recombinant protein. *J Immunother.* Mar.-Apr. 2004;27(2):79-91. doi: 10.1097/00002371-200403000-00001. Author Manuscript. 22 pages.
- Paschen et al., Detection of spontaneous CD4+ T-cell responses in melanoma patients against a tyrosinase-related protein-2-derived epitope identified in HLA-DRB1*0301 transgenic mice. *Clin Cancer Res.* Jul. 15, 2005;11(14):5241-7. doi: 10.1158/1078-0432.CCR-05-0170.
- Pinilla et al., Combinatorial peptide libraries as an alternative approach to the identification of ligands for tumor-reactive cytolytic T lymphocytes. *Cancer Res.* Jul. 1, 2001;61(13):5153-60.

(56)

References Cited

OTHER PUBLICATIONS

- Pinilla-Ibarz et al., Improved human T-cell responses against synthetic HLA-0201 analog peptides derived from the WT1 oncoprotein. *Leukemia*. Nov. 2006;20(11):2025-33. doi: 10.1038/sj.leu.2404380. Epub Aug. 31, 2006.
- Prather et al., De novo biosynthetic pathways: rational design of microbial chemical factories. *Curr Opin Biotechnol*. Oct. 2008;19(5):468-74. doi: 10.1016/j.copbio.2008.07.009. Epub Sep. 5, 2008.
- Quan et al., Influenza M1 VLPs containing neuraminidase induce heterosubtypic cross-protection. *Virology*. Sep. 1, 2012;430(2):127-35. doi: 10.1016/j.virol.2012.05.006. Epub Jun. 2, 2012.
- Rasmussen et al., Characterization of virus-like particles produced by a recombinant baculovirus containing the gag gene of the bovine immunodeficiency-like virus. *Virology*. Oct. 1990;178(2):435-51. doi: 10.1016/0042-6822(90)90341-n.
- Renner et al., Intact Viral Particle Counts Measured by Flow Virometry Provide Insight into the Infectivity and Genome Packaging Efficiency of Moloney Murine Leukemia Virus. *J Virol*. Jan. 6, 2020;94(2):e01600-19. doi: 10.1128/JVI.01600-19.
- Riley et al., Identification of a new shared HLA-A2.1 restricted epitope from the melanoma antigen tyrosinase. *J Immunother*. May-Jun. 2001;24(3):212-20.
- Rimoldi et al., Efficient simultaneous presentation of NY-ESO-1/LAGE-1 primary and nonprimary open reading frame-derived CTL epitopes in melanoma. *J Immunol*. Dec. 15, 2000;165(12):7253-61. doi: 10.4049/jimmunol.165.12.7253.
- Robbins et al., Multiple HLA class II-restricted melanocyte differentiation antigens are recognized by tumor-infiltrating lymphocytes from a patient with melanoma. *J Immunol*. Nov. 15, 2002;169(10):6036-47. doi: 10.4049/jimmunol.169.10.6036.
- Robbins et al., The intronic region of an incompletely spliced gp100 gene transcript encodes an epitope recognized by melanoma-reactive tumor-infiltrating lymphocytes. *J Immunol*. Jul. 1, 1997;159(1):303-8.
- Rohovie et al., Virus-like particles: Next-generation nanoparticles for targeted therapeutic delivery. *Bioeng Transl Med*. Jan. 19, 2017;2(1):43-57. doi: 10.1002/btm2.10049.
- Rosenberg et al., Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nat Med*. Mar. 1998;4(3):321-7. doi: 10.1038/nm0398-321.
- Rubio-Godoy et al., Toward synthetic combinatorial peptide libraries in positional scanning format (PS-SCL)-based identification of CD8+ Tumor-reactive T-Cell Ligands: a comparative analysis of PS-SCL recognition by a single tumor-reactive CD8+ cytolytic T-lymphocyte clone. *Cancer Res*. Apr. 1, 2002;62(7):2058-63.
- Ruiz et al., Identification and characterization of a T-helper peptide from carcinoembryonic antigen. *Clin Cancer Res*. Apr. 15, 2004;10(8):2860-7. doi: 10.1158/1078-0432.ccr-03-0476.
- Rusk, Cas9 and the importance of asymmetry. *Nat Methods*. Apr. 2016;13(4):286-7. doi: 10.1038/nmeth.3826.
- Saenger et al., Improved tumor immunity using anti-tyrosinase related protein-1 monoclonal antibody combined with DNA vaccines in murine melanoma. *Cancer Res*. Dec. 1, 2008;68(23):9884-91. doi: 10.1158/0008-5472.CAN-08-2233. Author Manuscript. 19 pages.
- Saenz et al., Feline immunodeficiency virus-based lentiviral vectors. *Cold Spring Harb Protoc*. Jan. 1, 2012;2012(1):71-6. doi: 10.1101/pdb.ip067579.
- Saenz et al., Production, harvest, and concentration of feline immunodeficiency virus-based lentiviral vector from cells grown in CF10 or CF2 devices. *Cold Spring Harb Protoc*. Jan. 1, 2012;2012(1):118-23. doi: 10.1101/pdb.prot067546.
- Sakuma et al., Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci Rep*. Jun. 23, 2014;4:5400. doi: 10.1038/srep05400.
- Schneider et al., Overlapping peptides of melanocyte differentiation antigen Melan-A/MART-1 recognized by autologous cytolytic T lymphocytes in association with HLA-B45.1 and HLA-A2.1. *Int J Cancer*. Jan. 30, 1998;75(3):451-8. doi: 10.1002/(sici)1097-0215(19980130)75:3<451::aid-ijc20>3.0.co;2-a.
- Sensi et al., Identification of a novel gp100/pMel17 peptide presented by HLA-A*6801 and recognized on human melanoma by cytolytic T cell clones. *Tissue Antigens*. Apr. 2002;59(4):273-9. doi: 10.1034/j.1399-0039.2002.590404.x.
- Shang et al., The spontaneous CD8+ T-cell response to HLA-A2-restricted NY-ESO-1b peptide in hepatocellular carcinoma patients. *Clin Cancer Res*. Oct. 15, 2004;10(20):6946-55. doi: 10.1158/1078-0432.CCR-04-0502.
- Sharma et al., Noninfectious virus-like particles produced by Moloney murine leukemia virus-based retrovirus packaging cells deficient in viral envelope become infectious in the presence of lipofection reagents. *Proc Natl Acad Sci U S A*. Sep. 30, 1997;94(20):10803-8. doi: 10.1073/pnas.94.20.10803.
- Shellenberger et al., A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. *Nat Biotechnol*. Dec. 2009;27(12):1186-90. doi: 10.1038/nbt.1588.
- Shen et al., Identification of a MHC class-II restricted epitope in carcinoembryonic antigen. *Cancer Immunol Immunother*. May 2004;53(5):391-403. doi: 10.1007/s00262-003-0455-y. Epub Nov. 18, 2003.
- Shim et al., Nonviral Delivery Systems for Cancer Gene Therapy: Strategies and Challenges. *Curr Gene Ther*. 2018;18(1):3-20. doi: 10.2174/156652321866180119121949.
- Skipper et al., An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. *J Exp Med*. Feb. 1, 1996;183(2):527-34. doi: 10.1084/jem.183.2.527.
- Skipper et al., Shared epitopes for HLA-A3-restricted melanoma-reactive human CTL include a naturally processed epitope from Pmel-17/gp100. *J Immunol*. Dec. 1, 1996;157(11):5027-33.
- Slansky et al., Enhanced antigen-specific antitumor immunity with altered peptide ligands that stabilize the MHC-peptide-TCR complex. *Immunity*. Oct. 2000;13(4):529-38. doi: 10.1016/s1074-7613(00)00052-2.
- Slingluff et al., Clinical and immunologic results of a randomized phase II trial of vaccination using four melanoma peptides either administered in granulocyte-macrophage colony-stimulating factor in adjuvant or pulsed on dendritic cells. *J Clin Oncol*. Nov. 1, 2003;21(21):4016-26. doi: 10.1200/JCO.2003.10.005.
- Slingluff et al., Immunologic and clinical outcomes of vaccination with a multiepitope melanoma peptide vaccine plus low-dose interleukin-2 administered either concurrently or on a delayed schedule. *J Clin Oncol*. Nov. 15, 2004;22(22):4474-85. doi: 10.1200/JCO.2004.10.212.
- Tangri et al., Structural features of peptide analogs of human histocompatibility leukocyte antigen class I epitopes that are more potent and immunogenic than wild-type peptide. *J Exp Med*. Sep. 17, 2001;194(6):833-46. doi: 10.1084/jem.194.6.833.
- Tomé-Amat et al., Secreted production of assembled Norovirus virus-like particles from *Pichia pastoris*. *Microb Cell Fact*. Sep. 10, 2014;13:134. doi: 10.1186/s12934-014-0134-z.
- Topalian et al., Melanoma-specific CD4+ T cells recognize nonmutated HLA-DR-restricted tyrosinase epitopes. *J Exp Med*. May 1, 1996;183(5):1965-71. doi: 10.1084/jem.183.5.1965.
- Touloukian et al., Expression of a "self"-antigen by human tumor cells enhances tumor antigen-specific CD4(+) T-cell function. *Cancer Res*. Sep. 15, 2002;62(18):5144-7. Author Manuscript. 11 pages.
- Touloukian et al., Identification of a MHC class II-restricted human gp100 epitope using DR4-IE transgenic mice. *J Immunol*. Apr. 1, 2000;164(7):3535-42. doi: 10.4049/jimmunol.164.7.3535.
- Touloukian et al., Normal tissue depresses while tumor tissue enhances human T cell responses in vivo to a novel self/tumor melanoma antigen, OA1. *J Immunol*. Feb. 1, 2003;170(3): 1579-85. doi: 10.4049/jimmunol.170.3.1579.
- Trojan et al., Generation of cytotoxic T lymphocytes against native and altered peptides of human leukocyte antigen-A*0201 restricted epitopes from the human epithelial cell adhesion molecule. *Cancer Res*. Jun. 15, 2001;61(12):4761-5.

(56)

References Cited**OTHER PUBLICATIONS**

- Tsai et al., Identification of subdominant CTL epitopes of the GP100 melanoma-associated tumor antigen by primary in vitro immunization with peptide-pulsed dendritic cells. *J Immunol.* Feb. 15, 1997;158(4):1796-802.
- Tsang et al., A human cytotoxic T-lymphocyte epitope and its agonist epitope from the nonvariable No. of tandem repeat sequence of MUC-1. *Clin Cancer Res.* Mar. 15, 2004;10(6):2139-49. doi: 10.1158/1078-0432.ccr-1011-03.
- Tsang et al., Generation of human cytotoxic T cells specific for human carcinoembryonic antigen epitopes from patients immunized with recombinant vaccinia-CEA vaccine. *J Natl Cancer Inst.* Jul. 5, 1995;87(13):982-90. doi: 10.1093/jnci/87.13.982.
- Tsuboi et al., Enhanced induction of human WT1-specific cytotoxic T lymphocytes with a 9-mer WT1 peptide modified at HLA-A*2402-binding residues. *Cancer Immunol Immunother.* Dec. 2002;51(11-12):614-20. doi: 10.1007/s00262-002-0328-9. Epub Oct. 18, 2002.
- Tuorto et al., Genome recoding by tRNA modifications. *Open Biol.* Dec. 2016;6(12):160287. doi: 10.1098/rsob.160287.
- Tycko et al., Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. *Mol Cell.* Aug. 4, 2016;63(3):355-70. doi: 10.1016/j.molcel.2016.07.004.
- Valmori et al., Analysis of the cytolytic T lymphocyte response of melanoma patients to the naturally HLA-A*0201-associated tyrosinase peptide 368-376. *Cancer Res.* Aug. 15, 1999;59(16):4050-5.
- Valmori et al., Enhanced generation of specific tumor-reactive CTL in vitro by selected Melan-A/MART-1 immunodominant peptide analogues. *J Immunol.* Feb. 15, 1998;160(4):1750-8.
- Valmori et al., Naturally occurring human lymphocyte antigen-A2 restricted CD8+ T-cell response to the cancer testis antigen NY-ESO-1 in melanoma patients. *Cancer Res.* Aug. 15, 2000;60(16):4499-506.
- Vigneron et al., A peptide derived from melanocytic protein gp100 and presented by HLA-B35 is recognized by autologous cytolytic T lymphocytes on melanoma cells. *Tissue Antigens.* Feb. 2005;65(2):156-62. doi: 10.1111/j.1399-0039.2005.00365.x.
- Vigneron et al., An antigenic peptide produced by peptide splicing in the proteasome. *Science.* Apr. 23, 2004;304(5670):587-90. doi: 10.1126/science.1095522.
- Visseren et al., Affinity, specificity and T-cell-receptor diversity of melanoma-specific CTL generated in vitro against a single tyrosinase epitope. *Int J Cancer.* Sep. 17, 1997;72(6):1122-8. doi: 10.1002/(sici)1097-0215(19970917)72:6<1122::aid-ijc30>3.0.co;2-3.
- Voelkel et al., Protein transduction from retroviral Gag precursors. *Proc Natl Acad Sci U S A.* Apr. 27, 2010;107(17):7805-10. doi: 10.1073/pnas.0914517107. Epub Apr. 12, 2010.
- Volpe et al., Alternative BCR/ABL splice variants in Philadelphia chromosome-positive leukemias result in novel tumor-specific fusion proteins that may represent potential targets for immunotherapy approaches. *Cancer Res.* Jun. 1, 2007;67(11):5300-7. doi: 10.1158/0008-5472.CAN-06-3737.
- Voutev et al., Bxb1 phage recombinase assists genome engineering in *Drosophila melanogaster*. *Biotechniques.* Jan. 1, 2017;62(1):37-38. doi: 10.2144/000114494.
- Walpita et al., Mammalian Cell-Derived Respiratory Syncytial Virus-Like Particles Protect the Lower as well as the Upper Respiratory Tract. *PLoS One.* Jul. 14, 2015;10(7):e0130755. doi: 10.1371/journal.pone.0130755.
- Walton et al., Spontaneous CD8 T cell responses against the melanocyte differentiation antigen RAB38/NY-MEL-1 in melanoma patients. *J Immunol.* Dec. 1, 2006;177(11):8212-8. doi: 10.4049/jimmunol.177.11.8212.
- Wang et al., CRISPR/Cas9 in Genome Editing and Beyond. *Annu Rev Biochem.* Jun. 2, 2016;85:227-64. doi: 10.1146/annurev-biochem-060815-014607. Epub Apr. 25, 2016.
- Wang et al., Recognition of an antigenic peptide derived from tyrosinase-related protein-2 by CTL in the context of HLA-A31 and -A33. *J Immunol.* Jan. 15, 1998;160(2):890-7.
- Wang et al., Recognition of breast cancer cells by CD8+ cytotoxic T-cell clones specific for NY-BR-1. *Cancer Res.* Jul. 1, 2006;66(13):6826-33. doi: 10.1158/0008-5472.CAN-05-3529.
- Wang et al., Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen. *J Exp Med.* Mar. 1, 1996;183(3):1131-40. doi: 10.1084/jem.183.3.1131.
- Wang et al., Virus-like particles for the prevention of human papillomavirus-associated malignancies. *Expert Rev Vaccines.* Feb. 2013;12(2):129-41. doi: 10.1586/erv.12.151. Author Manuscript, 22 pages.
- Wölfel et al., Two tyrosinase nonapeptides recognized on HLA-A2 melanomas by autologous cytolytic T lymphocytes. *Eur J Immunol.* Mar. 1994;24(3):759-64. doi: 10.1002/eji.1830240340.
- Yang et al., HIV-1 virus-like particles produced by stably transfected *Drosophila* S2 cells: a desirable vaccine component. *J Virol.* Jul. 2012;86(14):7662-76. doi: 10.1128/JVI.07164-11. Epub May 2, 2012.
- Yee et al., A general method for the generation of high-titer, pantrropic retroviral vectors: highly efficient infection of primary hepatocytes. *Proc Natl Acad Sci U S A.* Sep. 27, 1994;91(20):9564-8. doi: 10.1073/pnas.91.20.9564.
- Yu et al., Poor immunogenicity of a self/tumor antigen derives from peptide-MHC-I instability and is independent of tolerance. *J Clin Invest.* Aug. 2004;114(4):551-9. doi: 10.1172/JCI21695.
- Zarour et al., Melan-A/MART-1(51-73) represents an immunogenic HLA-DR4-restricted epitope recognized by melanoma-reactive CD4(+) T cells. *Proc Natl Acad Sci U S A.* Jan. 4, 2000;97(1):400-5. doi: 10.1073/pnas.97.1.400.
- Zeltins, A., Construction and characterization of virus-like particles: a review. *Mol Biotechnol.* Jan. 2013;53(1):92-107. doi: 10.1007/s12033-012-9598-4.
- Zhang et al., Cell-specific targeting of lentiviral vectors mediated by fusion proteins derived from Sindbis virus, vesicular stomatitis virus, or avian sarcoma/leukosis virus. *Retrovirology.* Jan. 25, 2010;7:3. doi: 10.1186/1742-4690-7-3.
- Zhang et al., CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice. *Sci Adv.* Apr. 12, 2017;3(4):e1602814. doi: 10.1126/sciadv.1602814.
- Zhao et al., Study on p21 gene knock out in G401 cell line by using CRISPR/Cas9 system. *Tianjin Med J.* Oct. 2016;44(10):1190-1194.

* cited by examiner

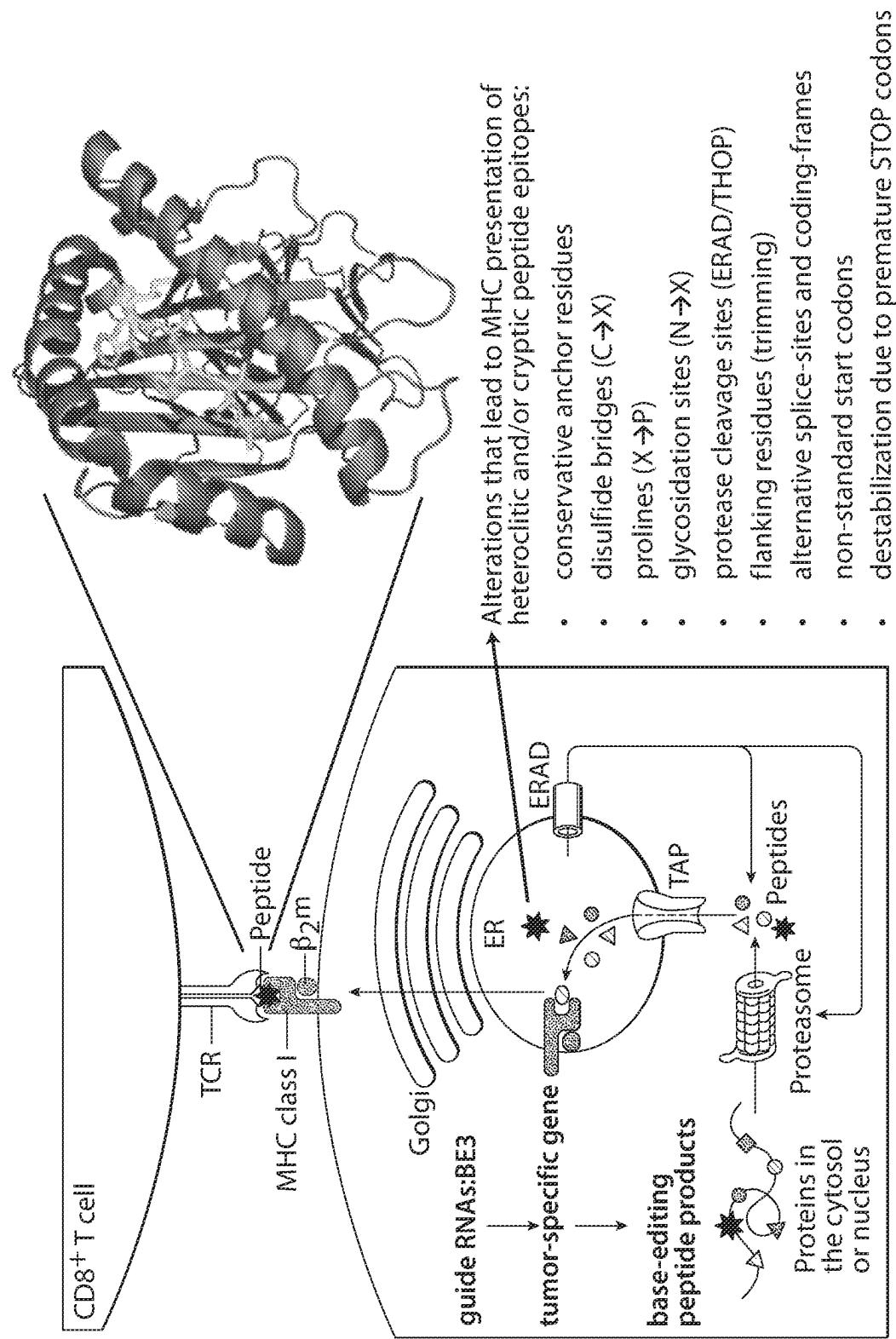


Figure 1

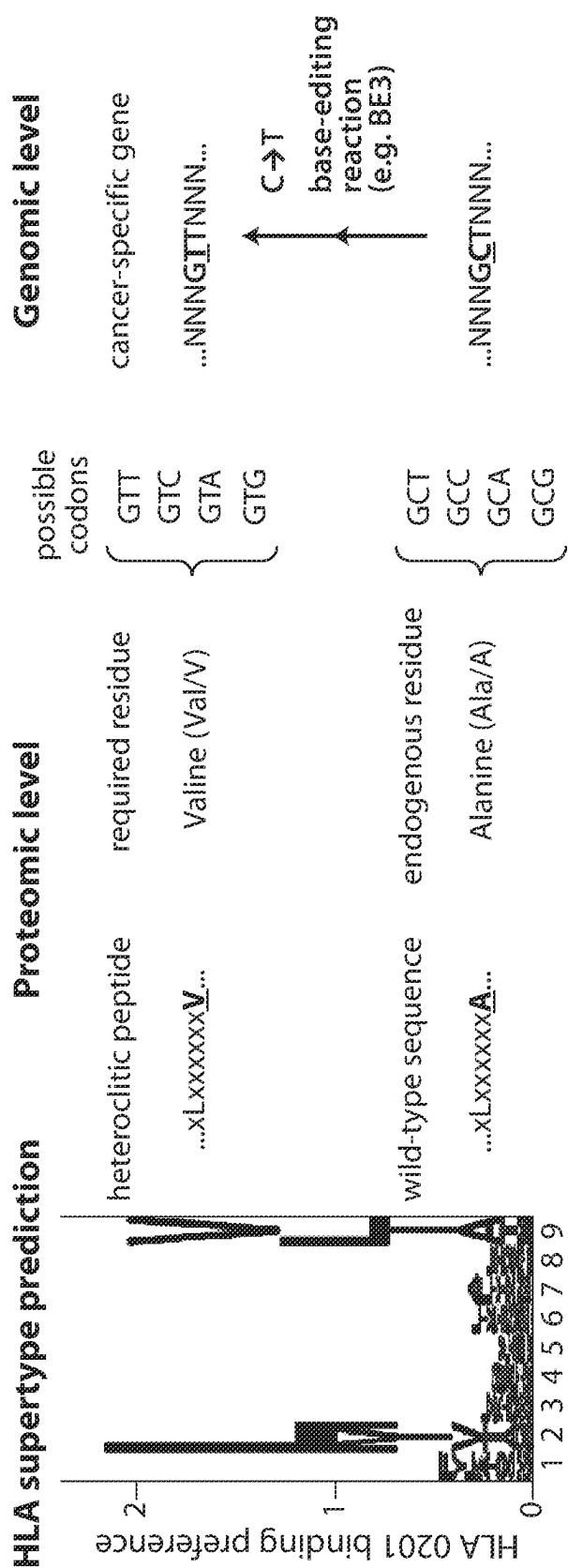


Figure 2A

HLA	Common alleles that are members of the super-type	Peptide-binding specificity	Population coverage	HLA	Common alleles that are members of the super-type	Peptide-binding specificity	Population coverage
A1	A*0101-02, A*2501, A*2601, A*2602, A*2604, A*3201, A*3601, A*4301, A*8001	X _T XXXXXX, Y _I W _V F _M	14.7-47.1%	B7	B*0702-05, B*1508, B*3501-03, B*51, B*5301, B*5401, B*5501-02, B*5601-02, B*6701, B*7801	X _P XXXXXX A _L I _V M _Y F _W	43.0-57.1%
A2	A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*6802, A*6901	X _L XXXXXX, L _I V _V M _M A _A T _T	39.0-45.9%	B27	B*1401-02, B*1503, B*1509, B*1510, B*1518, B*3801-02, B*3901-04, B*4801-02, B*7301, B*2701-08	X _R XXXXXX, F _L H _K Y _W	13.3-35.3%
A3	A*0301, A*01101, A*3101, A*3303, A*6801	X _A XXXXXX, F _L Y _I W _V M _M V _V	37.5-52.7%	B44	B*3701, B*4402-05, B*4001, B*4006	X _E XXXXXX, Y _F W _W	21.2-43.0%
A24	A*2402-04, A*3001-03	X _F XXXXXX, F _I Y _V	23.9-58.6%	B58	B*1516-17, B*5701-02, B*58	X _A XXXXXX, F _S T _W Y _Y	1.6-25.1%
				B62	B*1501-02, B*1513	X _Q XXXXXX, F _L W _W Y _Y	4.8-36.5%

Figure 2B

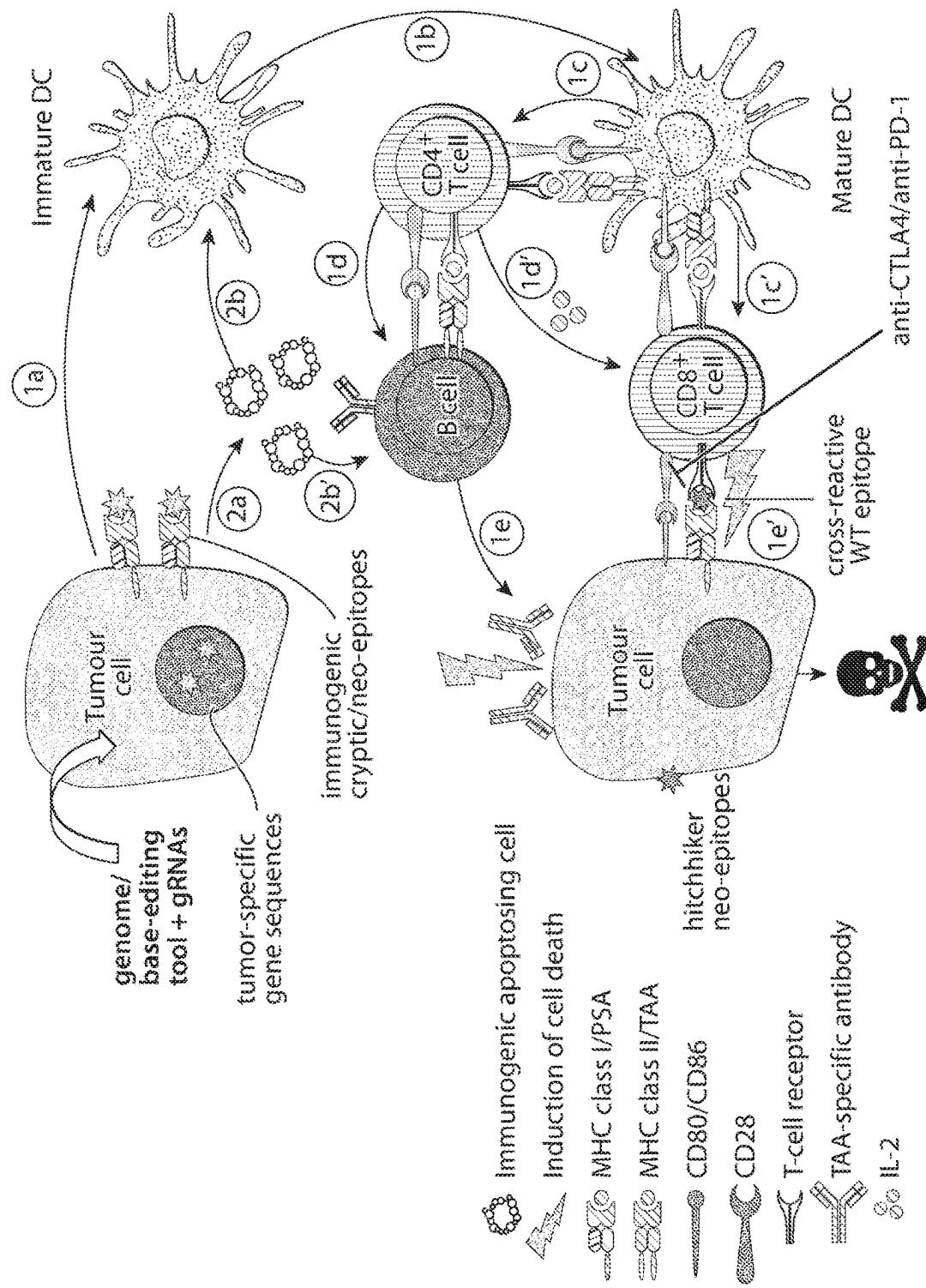


Figure 3

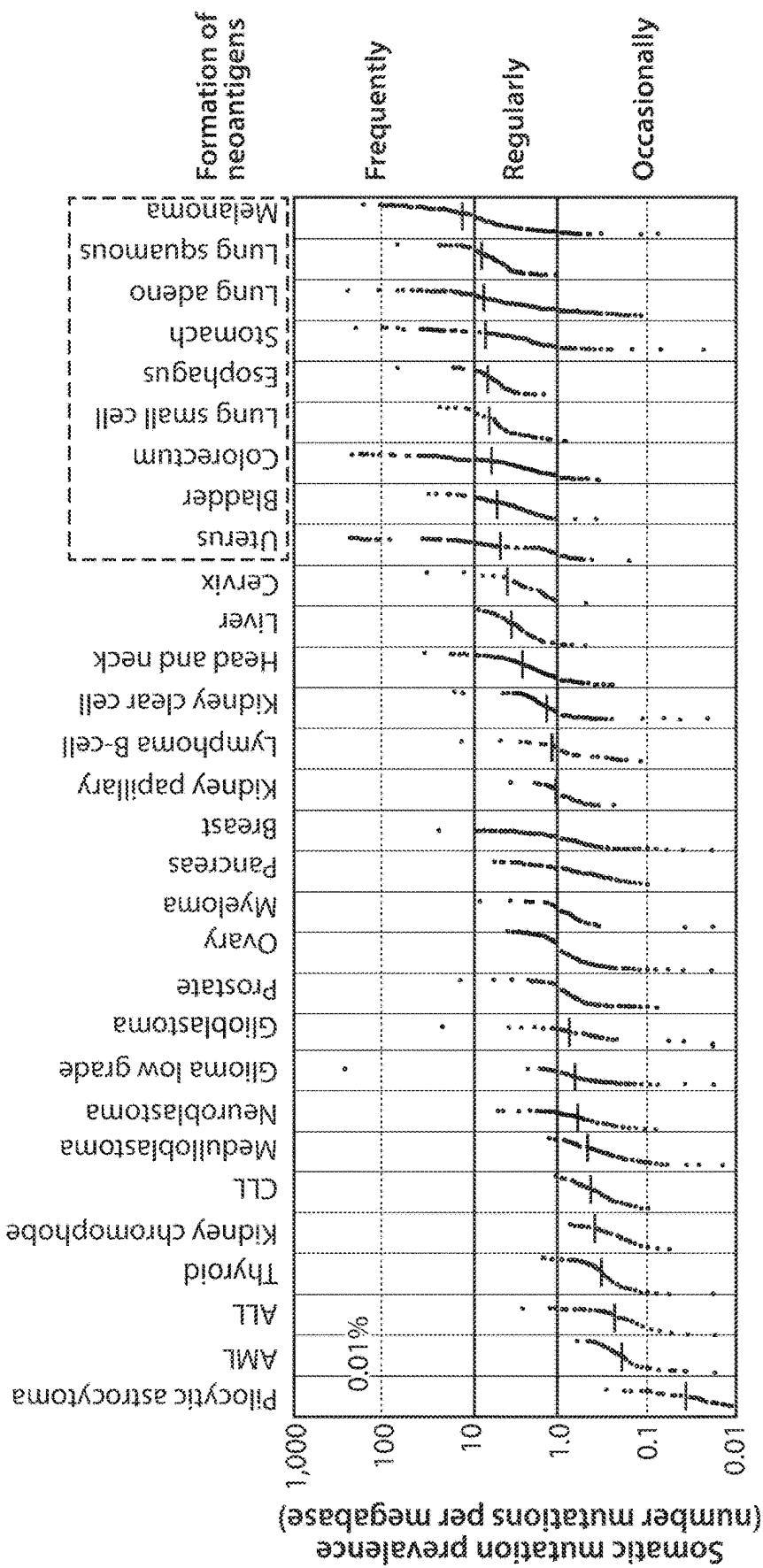


Figure 4

1
CANCER VACCINE

RELATED APPLICATIONS

The present application is a national stage filing under 35 U.S.C. § 371 of international PCT application, PCT/US2018/021880, filed Mar. 9, 2018, which claims priority under 35 U.S.C. § 119(e) to U.S. provisional patent application, U.S. Ser. No. 62/469,219, filed Mar. 9, 2017, the entire contents of each of which are incorporated herein by reference.

REFERENCE TO A SEQUENCE LISTING
SUBMITTED AS A TEXT FILE VIA EFS-WEB

The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Oct. 27, 2023, is named H082470241US01-SUBSEQ-AZW and is 3,827,234 bytes in size.

BACKGROUND OF THE INVENTION

Tumor-specific immune responses may be elicited by peptides generated from proteins expressed in tumor cells or on tumor cell surface (e.g., tumor-specific antigens). Native peptides derived from tumor-specific antigens are tolerated as “self” by the immune system and do not elicit strong immune response against the tumor-specific antigen. Altered versions of the native peptides derived from tumor-specific antigens (e.g., heteroclitic peptides or cryptic peptides) may be engineered to elicit potent immune reactions through the MHC-I and MHC-II antigen presentation pathways, which also produce cross-reactive responses towards the native tumor-specific antigen sequences.

It is well established that the immune system can function to kill tumor cells, including both primary and metastatic cancer cells. Indeed, evidence that the immune system recognizes the presence of neoplastic cancerous cells is supported by the existence of infiltrating lymphocytes in tumor tissues (Haskill et al., 1978, *Contemp. Top. Immunobiol.* 8: 107-170; Vose and Moore, 1985, *Semin. Hematol.* 22: 27-40). Yet, for reasons that are not completely clear, despite the presence of immune cells, tumors often prevail and not only survive but metastasize to distant sites with unrestricted growth. Recent advances in the understanding of T cell activation and recognition of target cells have begun to permit some progress in development of T cell mediated cancer immunotherapy (Schwartz, 1992, *Cell* 71: 1065-1068; Pardoll, 1992, *Curr. Opin. Immunol.* 4: 619-623).

SUMMARY OF THE INVENTION

Described herein are systems, methods, compositions, and kits for producing immunogenic peptides derived from tumor specific antigens (e.g., heteroclitic epitopes or cryptic epitopes) that may be used as cancer vaccines *in vivo* or *ex vivo*. Targeted mutations are introduced into tumor-specific antigens using gene editing agents, e.g., a nucleobase editor comprising a programmable DNA binding domain (e.g., catalytically-inactive Cas9 or a Cas9 nickase) fused to a cytosine deaminase, to generate altered versions of peptides arising from the tumor-specific antigens (heteroclitic epitopes) or peptides arising from normally untranslated regions of the tumor-specific antigen genes (cryptic peptides). The heteroclitic peptides or cryptic peptides may be

generated *in vivo* in a subject (e.g., a subject who has cancer) and presented to the adaptive immune system via the MHC class I or MHC class II pathway, which in turn induces a strong adaptive immune response, e.g., T cell response and B cell response. Such an adaptive immune response is antigen specific and is effective in reducing tumor growth and preventing metastasis.

Some aspects of the present disclosure provide methods of eliciting a tumor-specific immune response in a subject in need thereof, the methods including administering to the subject a therapeutically effective amount of a composition comprising: (i) a fusion protein comprising (a) a guide nucleotide sequence-programmable DNA-binding protein domain; and (b) a cytosine deaminase domain; and (ii) a guide nucleotide sequence, wherein the guide nucleotide sequence of (ii) targets the fusion protein of (i) to a polynucleotide encoding a tumor-specific antigen in a tumor cell, wherein the fusion protein changes a target cytosine (C) base to a thymine (T) base via deamination.

In some embodiments, the polynucleotide comprises a coding strand and a complementary strand. In some embodiments, the polynucleotide comprises a coding region and a non-coding region. In some embodiments, the polynucleotide encoding the tumor-specific antigen is located in the genome of the tumor cell. In some embodiments, deamination of the target C base results in a C-G base-pair to thymine-adenine (T-A) base-pair change.

In some embodiments, the guide nucleotide sequence-programmable DNA binding protein domain is selected from the group consisting of: nuclease inactive Cas9 (dCas9) domains, nuclease inactive Cpf1 domains, nuclease inactive Argonaute domains, and variants thereof.

In some embodiments, the guide nucleotide sequence-programmable DNA-binding protein domain is a nuclease inactive Cas9 (dCas9) domain. In some embodiments, the amino acid sequence of the dCas9 domain comprises mutations corresponding to a D10A and/or H840A mutation in SEQ ID NO: 1. In some embodiments, the amino acid sequence of the dCas9 domain comprises a mutation corresponding to a D10A mutation in SEQ ID NO: 1, and wherein the dCas9 domain comprises a histidine at the position corresponding to amino acid 840 of SEQ ID NO: 1.

In some embodiments, the guide nucleotide sequence-programmable DNA-binding protein domain comprises a nuclease inactive Cpf1 (dCpf1) domain. In some embodiments, the dCpf1 domain is from a species of *Acidaminococcus* or *Lachnospiraceae*. In some embodiments, the guide nucleotide sequence-programmable DNA-binding protein domain comprises a nuclease inactive Argonaute (dAgo) domain. In some embodiments, the (dAgo) domain is from *Natronobacterium gregoryi* (dNgAgo).

In some embodiments, the cytosine deaminase domain comprises an apolipoprotein B mRNA-editing complex (APOBEC) family deaminase. In some embodiments, the cytosine deaminase is selected from the group consisting of APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G deaminase, APOBEC3H deaminase, APOBEC4 deaminase, and activation-induced deaminase (AID). In some embodiments, the cytosine deaminase comprises an amino acid sequence of any of SEQ ID NOS: 27-292, 303, and 1072-1083.

In some embodiments, the fusion protein of (a) further comprises a uracil glycosylase inhibitor (UGI) domain. In some embodiments, the cytosine deaminase domain is fused to the N-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain. In some embodi-

ments, the UGI domain is fused to the C-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain. In some embodiments, the cytosine deaminase and the guide nucleotide sequence-programmable DNA-binding protein domain is fused via an optional linker. In some embodiments, the UGI domain is fused to the guide nucleotide sequence-programmable DNA-binding protein domain via an optional linker.

In some embodiments, the fusion protein comprises the structure NH₂-[cytosine deaminase domain]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA-binding protein domain]-[optional linker sequence]-[UGI domain]-COOH.

In some embodiments, the optional linker comprises (GGGS)_n, (SEQ ID NO: 337) (GGGGS)_n (SEQ ID NO: 308), (G)_n (SEQ ID NO: 783), (EAAAK)_n (SEQ ID NO: 309), (GGS)_n (SEQ ID NO: 784), SGSETPGTSESATPES (SEQ ID NO: 310), or (XP)_n (SEQ ID NO: 785) motif, or a combination of any of these, wherein n is independently an integer between 1 and 30 and wherein X is any amino acid. In some embodiments, the linker comprises the amino acid sequence of SGSETPGTSESATPES (SEQ ID NO: 310). In some embodiments, the linker is (GGS)_n (SEQ ID NO: 784), and wherein n is 1, 3, or 7.

In some embodiments, the fusion protein comprises the amino acid sequence of any one of SEQ ID NOS: 293-302, 1071, and 1084.

In some embodiments, the tumor specific antigen is selected from the group consisting of: CEA; gp100; Pmel17; gammaglobin-A; Melan-A; MART-1; NY-BR-1; ERBB2; OA1; PAP; PSA; RAB38; NY-MEL-1; TRP-1; gp75; TRP-2; tyrosinase; WT1; CD33; BAGE-1; D393-CD20n; Cyclin-A1; GAGE-1,2,8; GAGE-3,4,5,6,7; GnTVf; HERV-K-MEL; KK-LC-1; KM-HN-1; LAGE-1; LY6K; MAGE-A1; MAGE-A2; MAGE-A3; MAGE-A4; MAGE-A6; MAGE-A9; MAGE-A10; MAGE-A12m; MAGE-C1; MAGE-C2; mucink; NA88-A; NY-ESO-1; LAGE-2; SAGE; Sp17; SSX-2; SSX-4; survivin; BIRC5; TAG-1; TAG-2; TRAG-3; TRP2-INT2g; XAGE-1b; GAGED2a; BCR-ABL (b3a2); adipophilin; AIM-2; ALDH1A1; BCLX(L); BING-4; CALCA; CD45; CD274; CPSF; cyclin D1; DKK1; ENAH (hMena); EpCAM; EphA3; EZH2; FGF5; glyican-3; G250; MN; CAIX; HER-2; neu; HLA-DOB; Hepsin; IDO1; IGF2B3; IL13Ralpha2; Intestinal carboxyl esterase; alpha-fetoprotein; Kallikrein 4; KIF20A; Lengsin; M-CSF; MCSP; mdm-2; Meloe; Midkine; MMP-2; MMP-7; MUC1; MUC5AC; p53; PAX5; PBF; PRAME; PSMA; RAGE-1; RGS5; RhoC; RNF43; RU2AS; secerin 1; SOX10; STEAP1; Telomerase; TPBG; and VEGF.

In some embodiments, the target C base is in a target codon located in a coding region of the polynucleotide encoding the tumor-specific antigen. In some embodiments, the target codon is any one of the target codons in Tables 4 and 8.

In some embodiments, the target codon is converted to a modified codon selected from any one of the modified codons in Table 4. In some embodiments, the target C base is located in a non-coding region of the polynucleotide encoding the tumor specific antigen. In some embodiments, the target C base is located in an intron in the polynucleotide encoding the tumor specific antigen.

In some embodiments, the methods described herein further comprising generating an immunogenic peptide from the tumor-specific antigen. In some embodiments, the immunogenic peptide is a heteroclitic epitope. In some embodiments, the heteroclitic epitope is at least 2 fold, at least 5 fold, at least 10 fold, at least 20 fold, at least 30 fold,

at least 40 fold, at least 50 fold, at least 60 fold, at least 70 fold, at least 80 fold, at least 90 fold, at least 100 fold, or more immunogenic than a native epitope from the tumor specific antigen. In some embodiments, the immunogenic peptide is a cryptic epitope. In some embodiments, the cryptic epitope is at least 2 fold, at least 5 fold, at least 10 fold, at least 20 fold, at least 30 fold, at least 40 fold, at least 50 fold, at least 60 fold, at least 70 fold, at least 80 fold, at least 90 fold, at least 100 fold, or more immunogenic than a native epitope from the tumor specific antigen.

In some embodiments, the immunogenic peptide is displayed on the surface of the tumor cell via the MHC class I antigen presentation pathway. In some embodiments, the immunogenic peptide is displayed on the surface of an antigen presenting cell (APC) via the MHC class II antigen presentation pathway.

In some embodiments, the method is carried out *in vivo*. In some embodiments, the method is carried out *ex vivo*.

In some embodiments, the APC is selected from the group consisting of: tumor cells, dendritic cells, mononuclear phagocytes, thymic epithelial cells, and B cells.

In some embodiments, the immunogenic peptide elicits adaptive immune response against the tumor-specific antigen. In some embodiments, the adaptive immune response comprises promoting the maturation of dendritic cells, activation of CD4+ T lymphocytes, activation of CD8+ T lymphocytes, activation and maturation of B lymphocytes, and/or production of tumor antigen-specific antibodies. In some embodiments, the adaptive immune response kills tumor cells, reduces tumor size, and/or prevents metastasis.

In some embodiments, the guide nucleotide sequence is an RNA. In some embodiments, the RNA is chemically modified.

In some embodiments, the guide nucleotide sequence is a single strand DNA (ssDNA).

In some embodiments, the tumor specific antigen is gp100. In some embodiments, the gp100 is from melanoma. In some embodiments, the deamination of the target C base in codon T210 of gp100 results in a T210I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of IIDQVPFSV (SEQ ID NO: 786) is generated, and wherein the I at position 2 corresponds to the T210I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 724 and 870-888.

In some embodiments, the deamination of the target C base in codon A288 of gp100 results in a A288V mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of YLEPGPVTV (SEQ ID NO: 818) is generated, and wherein the V at position 7 corresponds to the A288V mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 725 and 889.

In some embodiments, the deamination of the target C base in codon T155 of gp100 results in a T155I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of KIWGQQYWQV (SEQ ID NO: 787) is generated, and wherein the I at position 2 corresponds to the T155I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 726 and 890-892.

In some embodiments, the tumor specific antigen is melanoma antigen recognized by T cells 1 (MART-1). In some embodiments, the MART-1 antigen is from melanoma. In some embodiments, the deamination of the target C base in codon A27 of MART-1 results in a A27V mutation. In

some embodiments, a heteroclitic epitope comprising the amino acid sequence of EVAGIGILTV (SEQ ID NO: 819) is generated, and wherein the V at position 2 corresponds to the A27V mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 727 and 893-896.

In some embodiments, the tumor specific antigen is cancer/testis antigen 1B (NY-ESO-1). In some embodiments, the NY-ESO-1 antigen is from melanoma or breast cancer. In some embodiments, the deamination of the target C base in codon C165 of NY-ESO-1 results in a C165Y mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of SLLMWITQY (SEQ ID NO: 788) is generated, and wherein the C at position 9 corresponds to the C165Y mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 728 and 897.

In some embodiments, the tumor specific antigen is Tyrosinase (TYR). In some embodiments, the TYR antigen is from melanoma. In some embodiments, the deamination of the target C base in codon T373 of TYR results in a T373I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of YMNGIMSQV (SEQ ID NO: 789) is generated, and wherein the I at position 5 corresponds to the T373I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 729 and 898-901.

In some embodiments, the tumor specific antigen is tyrosinase-related protein 1 (TyRP1). In some embodiments, the TyRP1 antigen is from melanoma. In some embodiments, the deamination of the target C base in codon C244 of TyRP1 results in a C244Y mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of DAEKYDICTDEY (SEQ ID NO: 790) is generated, and wherein the Y at position 5 corresponds to the C244Y mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 730 and 902.

In some embodiments, the tumor specific antigen is Survivin. In some embodiments, the Survivin is from melanoma, breast cancer, or leukemia. In some embodiments, the deamination of the target C base in codon T97 of Survivin results in a T97I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of ELILGEFLKL (SEQ ID NO: 791) is generated, and wherein the I at position 3 corresponds to the T97I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 731 and 903.

In some embodiments, the tumor specific antigen is telomerase reverse transcriptase (hTERT). In some embodiments, the hTERT is from breast cancer. In some embodiments, the deamination of the target C base in codon M549 of hTERT results in a M549I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of ILAKFLHWL (SEQ ID NO: 792) is generated, and wherein the I at position 10 corresponds to the M549I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 735 and 916-920.

In some embodiments, the tumor specific antigen is human epidermal growth factor receptor 2 (HER2). In some embodiments, the HER2 is from breast cancer. In some embodiments, the deamination of the target C base in codon

V658 of HER2 results in a V658M mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of AMVGILLVVV (SEQ ID NO: 793) is generated, and wherein the M at position 2 corresponds to the V658M mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 732 and 904-909.

In some embodiments, the deamination of the target C base in codon T912 of HER2 results in a T912I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of IIWELMTFGA (SEQ ID NO: 794) is generated, and wherein the V at position 2 corresponds to the T912I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 733 and 910-912.

In some embodiments, the deamination of the target C base in codon A920 of HER2 results in a A920V mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of ITWELMTFGV (SEQ ID NO: 795) is generated, and wherein the V at position 10 corresponds to the A920V mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 734 and 913-915.

In some embodiments, the tumor specific antigen is CD33. In some embodiments, the CD33 is from leukemia. In some embodiments, the deamination of the target C base in codon A65 of CD33 results in a A65V mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of VIISGDSPV (SEQ ID NO: 796) is generated, and wherein the V at position 1 corresponds to the A65V mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 740 and 930-932.

In some embodiments, the tumor specific antigen is Synovial Sarcoma X Breakpoint 2 (SSX2). In some embodiments, the deamination of the target C base in codon A42 of SSX2 results in a A42V mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of KVSEKIFYV (SEQ ID NO: 797) is generated, and wherein the V at position 2 corresponds to the A42V mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 737 and 921.

In some embodiments, the tumor specific antigen is Wilm's tumor 1 (WT1) protein. In some embodiments, the WT1 is from leukemia. In some embodiments, the deamination of the target C base in codon C235 of WT1 results in a C235Y mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of YMTWNQMNL (SEQ ID NO: 798) is generated, and wherein the Y at position 1 corresponds to the C235Y mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 738 and 922-925.

In some embodiments, the deamination of the target C base in codon M236 of WT1 results in a M236I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of CITWNQMNL (SEQ ID NO: 799) is generated, and wherein the I at position 2 corresponds to the M236I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 739 and 926-929.

In some embodiments, the tumor specific antigen is Epithelial cell adhesion molecule precursor (EpCAM). In some embodiments, the deamination of the target C base in codon T192 of EpCAM results in a T192I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of ILYENNVI (SEQ ID NO: 800) is generated, and wherein the I at position 9 corresponds to the T192I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 741 and 933-934.

In some embodiments, the tumor specific antigen is carcinoembryonic antigen-related cell adhesion molecules (CEA-CAM). In some embodiments, the CEA-CAM is from colorectal cancer, lung cancer, or breast cancer. In some embodiments, the deamination of the target C base in codon T314 of CEA-CAM results in a T314I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of LLTFWNPP (SEQ ID NO: 801) is generated, and wherein the I at position 9 corresponds to the T314I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 742 and 935-936.

In some embodiments, the deamination of the target C base in codon T311 of CEA-CAM results in a T311I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of RITVTTITV (SEQ ID NO: 802) is generated, and wherein the V at position 2 corresponds to the T311I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 743 and 937-940.

In some embodiments, the deamination of the target C base in codon T688 of CEA-CAM results in a T688V mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of AVVGIMIVG (SEQ ID NO: 803) is generated, and wherein the V at position 2 corresponds to the T688V mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 744 and 941-947.

In some embodiments, the deamination of the target C base in codon V695 of CEA-CAM results in a V695M mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of IMIGMLVGV (SEQ ID NO: 804) is generated, and wherein the M at position 5 corresponds to the V695M mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 745 and 948-953.

In some embodiments, the tumor specific antigen is melanoma-associated antigen A3 (MAGEA3). In some embodiments, the deamination of the target C base in codon H118 of MAGEA3 results in a H118Y mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of KVAELVYFL (SEQ ID NO: 805) is generated, and wherein the Y at position 7 corresponds to the H118Y mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 746 and 954.

In some embodiments, the tumor specific antigen is melanoma-associated antigen (MAGE) common antigen A3, A1, A4, A2, or A12. In some embodiments, the deamination of the target C base in codon C181 of MAGE common antigen A3, A1, A4, A2, or A12 results in a C181Y mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of YLGLSYDGLL (SEQ ID NO: 806) is generated, and wherein the Y at position 1

corresponds to the C181Y mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 747-750 and 955-983.

In some embodiments, the tumor specific antigen is MUC-1. In some embodiments, the deamination of the target C base in codon T93 of MUC-1 results in a T93I mutation. In some embodiments, a heteroclitic epitope comprising the amino acid sequence of AIWGQDVTSV (SEQ ID NO: 807) is generated, and wherein the I at position 2 corresponds to the T93I mutation. In some embodiments, the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 751 and 984-985.

In some embodiments, the target C base is located in intron 4 of the premelanosome protein (PMEL) gene. In some embodiments, the deamination of the target C base results in a cryptic peptide comprising the amino acid sequence of VYFFLPDHL (SEQ ID NO: 808). In some embodiments, the guide nucleotide sequence comprises a nucleotide selected from the group consisting of SEQ ID NOs: 752-753 and 986-998.

In some embodiments, the target C base is located on the complementary strand of open reading frame 1 (ORF1) of TYRP1 gene. In some embodiments, the target C base is located in the complementary strand of the first start codon (ATG) of ORF1 of the TYRP1 gene. In some embodiments, the deamination of the target C base results in a cryptic peptide comprising the amino acid sequence of MSLQRQFLR (SEQ ID NO: 809). In some embodiments, the guide nucleotide sequence comprises a nucleotide selected from the group consisting of SEQ ID NOs: 754 and 999-1005.

In some embodiments, the target C base is located on the complementary strand of the last base of intron 2 of the mannosey (alpha-1,6)-glycoprotein beta-1,6-N-acetyl glucosaminyltransferase (MGAT5) gene. In some embodiments, the deamination of the target C base results in a cryptic peptide comprising the amino acid sequence of VLPDVFIRCV (SEQ ID NO: 810). In some embodiments, the cryptic peptide is translated from exon 3 of the MGAT5 gene. In some embodiments, the guide nucleotide sequence comprises a nucleotide selected from the group consisting of SEQ ID NOs: 755 and 1006-1008.

In some embodiments, the target C base is located in open reading frame 1 (ORF1) of cancer/testis antigen 2 (LAGE-1) gene. In some embodiments, the target C base is located in the complementary strand of the first start codon of ORF1 of the LAGE-1 gene. In some embodiments, the deamination of the target C base results in a cryptic peptide comprising the amino acid sequence of selected from the group consisting of: MLMAQEALAF (SEQ ID NO: 811), LAAQERRVPR (SEQ ID NO: 812), APRGVRMVA (SEQ ID NO: 813), QGAMLAQERRVPRAEVPR (SEQ ID NO: 814), and CLSRRPWKRWSAGSCPGMPHL (SEQ ID NO: 815). In some embodiments, the guide nucleotide sequence comprises a nucleotide selected from the group consisting of SEQ ID NOs: 756 and 1009-1014.

In some embodiments, the target C base is located in intron 2 of tyrosinase-related protein 2 (TRP-2) gene. In some embodiments, the target C base is located on the complementary strand of the first base of intron 2 of the TRP-2 gene. In some embodiments, the deamination of the target C base results in a cryptic peptide comprising the amino acid sequence of EVISCKLIKR (SEQ ID NO: 816). In some embodiments, the guide nucleotide sequence com-

prises a nucleotide selected from the group consisting of SEQ ID NOs: 757-758 and 1015-1023.

In some embodiments, the target C base is located in intron 2 of baculoviral IAP repeat containing 5 (BIRC5) gene. In some embodiments, the target C base is located on the spliceosome branch site of intron 2 of the BIRC5 gene. In some embodiments, the target C base is located in the complementary strand of the last base of intron 2 of the BIRC5 gene. In some embodiments, the deamination of the target C base results in a cryptic peptide comprising the amino acid sequence of AYACNTSTL (SEQ ID NO: 817). In some embodiments, the guide nucleotide sequence comprises a nucleotide selected from the group consisting of SEQ ID NOs: 759 and 1024-1029.

In some embodiments, the target C base is located in intron 1 acceptor site of BCR/ABL fusion proteins (BCR/ABL-OOF) gene. In some embodiments, the target C base is located in intron 2 acceptor site of BCR/ABL fusion proteins (BCR/ABL-OOF) gene. In some embodiments, the deamination of the target C base results in a cryptic peptide comprising the amino acid sequence of any one of SSKA-LQRPV (SEQ ID NO: 603), GFKQSSKAL (SEQ ID NO: 604), and ATGFKQSSKALQRPVAS (SEQ ID NO: 605). In some embodiments, the guide nucleotide sequence comprises a nucleotide selected from the group consisting of SEQ ID NOs: 761 and 1032-1045. In some embodiments, the guide nucleotide sequence comprises a nucleotide selected from the group consisting of SEQ ID NOs: 762 and 1046-1056.

In some embodiments, the methods further comprising administering to the subject a therapeutically effective amount of an immune checkpoint inhibitor. In some embodiments, the immune checkpoint inhibitor inhibits binding of CTLA-4, PD-1, PD-L1, TIM3, LAG3, B7-H3, B7-H4, BTLA, GAL9, Chk1, or A2aR to a cognate binding partner. In some embodiments, the immune checkpoint inhibitor is an antibody or a fragment thereof. In some embodiments, the antibody is selected from anti-CTLA-4 antibodies, anti-PD-1 antibodies, anti-PD-L1 antibodies, anti-TIM3 antibodies, anti-LAG3 antibodies, anti-B7-H3 antibodies, anti-B7-H4 antibodies, anti-BTLA antibodies, anti-GAL9 antibodies, anti-Chk1 antibodies, and anti-A2aR antibodies. In some embodiments, the antibody is selected from pembrolizumab, nivolumab, and ipilimumab.

In some embodiments, the immune checkpoint inhibitor is a small molecule.

In some embodiments, the immune checkpoint inhibitor is a recombinant protein.

In some embodiments, the immune checkpoint inhibitor is a nucleic acid aptamer.

In some embodiments, the immune checkpoint inhibition is performed by genome editing of a gene selected from the group consisting of: CTLA-4, PD-1, PD-L1, TIM3, LAG3, B7-H3, B7-H4, BTLA, GAL9, Chk1, or A2aR.

Other aspects of the present disclosure provide methods of treating cancer, the methods including administering to a subject in need thereof a therapeutically effective amount of a composition comprising: (i) a fusion protein comprising (a) a guide nucleotide sequence-programmable DNA-binding protein domain; and (b) a cytosine deaminase domain; and (ii) a guide nucleotide sequence; wherein the fusion protein of (i) and the guide nucleotide sequence of (ii) enters a tumor cell, and wherein the guide nucleotide sequence targets the fusion protein of (i) to a polynucleotide encoding a tumor-specific antigen, wherein the fusion protein changes a target cytosine (C) residue to a (T) residue in the polynucleotide.

In some embodiments, the methods include administering to the subject a therapeutically effective amount of an immune checkpoint inhibitor.

Further provided herein are methods of inducing a tumor-specific immune response in a subject in need thereof, the method comprising administering to the subject a therapeutically effective amount of a composition comprising: (i) a fusion protein comprising (a) a guide nucleotide sequence-programmable DNA-binding protein domain; and (b) a nuclease domain; and (ii) a guide nucleotide sequence; wherein the fusion protein of (i) and the guide nucleotide sequence of (ii) enters the tumor cell, and wherein the guide nucleotide sequence targets the fusion protein of (i) to a polynucleotide encoding a tumor-specific antigen, wherein the fusion protein introduces an indel in the polynucleotide. In some embodiments, the nuclease is a FokI nuclease.

Further provided herein are methods of inducing a tumor-specific immune response in a subject in need thereof, the methods including administering to the subject a therapeutically effective amount of a composition comprising: (i) a guide nucleotide sequence-programmable nuclease; and (ii) a guide nucleotide sequence; wherein the fusion protein of (i) and the guide nucleotide sequence of (ii) enters the tumor cell, and wherein the guide nucleotide sequence targets the fusion protein of (i) to a polynucleotide encoding a tumor-specific antigen, wherein the guide nucleotide sequence-programmable nuclease introduces an indel in the polynucleotide.

In some embodiments, the guide nucleotide sequence-programmable nuclease comprises a Cas9, a Cpf1, an Argonaute, or a variant thereof. In some embodiments, the indel causes a mutation or frame shift.

Method of inducing a tumor-specific immune response in a subject in need thereof are also provided, the methods including administering to a subject in need thereof a therapeutically effective amount of a composition comprising a fusion protein comprising (a) a programmable DNA-binding protein domain; and (b) a deaminase domain; wherein the fusion protein enters the tumor cell and changes a target base in the polynucleotide via deamination.

In some embodiments, the deaminase domain comprises a cytosine deaminase and the target base is a cytosine (C) base. In some embodiments, the programmable DNA-binding domain comprises a zinc finger nuclease (ZFN). In some embodiments, the programmable DNA-binding domain comprises a transcription activator-like effector (TALE).

In some embodiments, the programmable DNA-binding domain is a guide nucleotide sequence-programmable DNA binding protein domain. In some embodiments, the programmable DNA-binding domain is selected from the group consisting of: nuclease-inactive Cas9 domains, nuclease inactive Cpf1 domains, nuclease inactive Argonaute domains, and variants thereof. In some embodiments, the programmable DNA-binding domain is associated with a guide nucleotide sequence. In some embodiments, the deamination of the target C base results in a C to thymine (T) change. In some embodiments, the deamination of the target C base results in a C-G base pair to thymine-adenine (T-A) change in a translated codon, resulting in the incorporation of a different amino acid in an immunogenic or heteroclitic peptide. In some embodiments, the deamination of the target C base results in a C-G basepair to thymine-adenine (T-A) change in an non-coding intron region of a gene, resulting in alternative splicing and translation of immunogenic or cryptic peptide sequences. In some embodiments, the deamination of the target C base results in a C-G basepair to thymine-adenine (T-A) change in the start (Met) codon of

11

the open reading frame of a gene, resulting in the translation of an alternative open reading frame comprising immunogenic or cryptic peptide sequences.

Other aspects of the present disclosure provide compositions comprising: (i) a fusion protein comprising (a) a guide nucleotide sequence-programmable DNA-binding protein domain; and (b) a cytosine deaminase domain; and (ii) a guide nucleotide sequence targeting the fusion protein of (i) to a polynucleotide encoding a tumor specific antigen.

Yet other aspects of the present disclosure provide compositions comprising a polynucleotide encoding a fusion protein and a guide nucleotide sequence, wherein the fusion protein comprises (a) a guide nucleotide sequence-programmable DNA-binding protein domain; and (b) a cytosine deaminase domain, and wherein the guide nucleotide sequence targets the fusion protein to a polynucleotide encoding a tumor specific antigen.

Yet other aspects of the present disclosure provide cancer vaccines comprising: (i) a fusion protein comprising (a) a guide nucleotide sequence-programmable DNA-binding protein domain; and (b) a cytosine deaminase domain; and (ii) a guide nucleotide sequence targeting the fusion protein of (i) to a polynucleotide encoding a tumor specific antigen.

Further provided herein are cancer vaccine comprising a polynucleotide encoding a fusion protein and a guide nucleotide sequence, wherein the fusion protein comprises (a) a guide nucleotide sequence-programmable DNA-binding protein domain; and (b) a cytosine deaminase domain, and wherein the guide nucleotide sequence targets the fusion protein to a polynucleotide encoding a tumor specific antigen.

Kits comprising the cancer vaccines described herein are also provided.

The details of certain embodiments of the disclosure are set forth in the Detailed Description of Certain Embodiments, as described below. Other features, objects, and advantages of the disclosure will be apparent from the Definitions, Examples, Figures, and Claims.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which constitute a part of this specification, illustrate several embodiments of the disclosure and together with the description, serve to explain the principles of the disclosure.

FIG. 1 shows strategies to engineer heteroclitic and cryptic epitopes using genome base editing.

FIG. 2A shows strategies to introduce immunogenic heteroclitic epitopes by editing conservative anchor residues to match the binding preference of the main HLA allele supertypes. The example shows a base-editing reaction that turns an alanine residue at anchor position 9 of a weakly immunogenic peptide epitope into a preferred valine residue for binding HLA-A2.

FIG. 2B shows anchor-residue binding preference and population coverage of the main HLA allele supertypes (MHC-I pathway). The peptides in FIG. 2B are as follows: HLA A1, A2, A3, and A24 (SEQ ID NOS: 878-881) and HLA B7, B27, B44, B58, and B62 (SEQ ID NOS: 882-886).

FIG. 3 shows a proposed mechanism for anti-cancer vaccination by heteroclitic/cryptic epitopes introduced by genome base-editing reactions programmed by guide-RNAs. The edited tumor cells produce heteroclitic and cryptic epitopes in cancer-specific genes, which chemotactically attract immature dendritic cells (DCs) (1a), inducing DC maturation (1b). Edited tumor cells produce apoptotic bodies (2a) that are taken up by DCs (2b), contributing to

12

maturity of DCs (1b) and B cells (2b'). Mature DCs activate CD4+ T lymphocytes (1c) and CD8+ T lymphocytes (1c'). Activated CD4+ T lymphocytes further stimulate B-lymphocyte activation (1d) and provide IL-2 for CD8+ T lymphocytes (1d'). B lymphocytes produce TAA-specific antibodies to cell-surface proteins that result in antibody-dependent cell-mediated cytotoxicity or complement-mediated tumor cell death (1e). Activated CD8+ T lymphocytes then kill tumor cells via recognition of MHC class I molecules in association with TAA epitopes (1e').

FIG. 4 shows comparison of cancer lineages that display high frequency of mutagenesis, which may harbor non-synonymous hitchhiker mutations and “neo-epitopes”.

DEFINITIONS

As used herein and in the claims, the singular forms “a,” “an,” and “the” include the singular and the plural reference unless the context clearly indicates otherwise. Thus, for example, a reference to “an agent” includes a single agent and a plurality of such agents.

An “immunogenic peptide” or “antigenic peptide” is a peptide or epitope that can be recognized by the immune system and elicit an immune response. Immunogenic peptides or antigenic peptide may comprise a motif such that the peptide will bind an MHC molecule and induce a T cell response, or can be recognized by the B cell receptor on the B cell to induce antibody production. These terms are used interchangeably herein.

An “immunogenic epitope” or “antigenic epitope” refers to a part of an antigen is recognized by the immune system, e.g., by antibodies, B cells, or T cells. In some embodiments, the epitope is the specific piece of the antigen to which an antibody binds. Although epitopes are usually non-self proteins, sequences derived from the host can, in some instances, be recognized.

“Immune response” is how your body recognizes and defends itself against bacteria, viruses, and substances that appear foreign and harmful to the body. In its general form, the immune response begins with the sensitization of helper (TH, CD4+) and cytotoxic (CD8+) T cell subsets through their interaction with antigen presenting cells (APC) that express major histocompatibility (MHC)-class I or class II molecules associated with antigenic fragments (i.e., specific amino acid sequences derived from the antigen which bind to MHC I and/or MHC II for presentation on the cell surface). The sensitized or primed CD4+ T cells produce lymphokines that participate in the activation of B cells as well as various T cell subsets. The sensitized CD8+ T cells increase in numbers in response to lymphokines and are capable of destroying any cells that express the specific antigenic fragments associated with matching MHC-encoded class I molecules. Thus, in the course of a cancerous tumor, CTL eradicate cells expressing cancer associated or cancer specific antigens, thereby limiting the progression of tumor spread and disease development.

The “adaptive immune system,” also known as the acquired immune system, is a subsystem of the overall immune system that is composed of highly specialized, systemic cells and processes that eliminate or prevent pathogen growth. The adaptive immune system is one of the two main immunity strategies found in vertebrates (the other being the innate immune system). Adaptive immunity creates immunological memory after an initial response to a specific pathogen, and leads to an enhanced response to subsequent encounters with that pathogen. This process of acquired immunity is the basis of vaccination. Like the

13

innate system, the adaptive system includes both humoral immunity components and cell-mediated immunity components.

Unlike the innate immune system, the adaptive immune system is highly specific to a particular pathogen or antigen. Adaptive immunity can also provide long-lasting protection. The adaptive system response destroys invading pathogens and any toxic molecules they produce. In accordance with the present disclosure, the adaptive immune system response destroys tumor or cancer cells. Sometimes the adaptive system is unable to distinguish harmful from harmless foreign molecules. The cells that carry out the adaptive immune response are white blood cells known as lymphocytes. Two main broad classes—antibody responses and cell mediated immune response—are also carried by two different lymphocytes (B cells and T cells). In antibody responses, B cells are activated to secrete antibodies, which are proteins also known as immunoglobulins. Antibodies travel through the bloodstream and bind to the foreign antigen causing it to inactivate, which does not allow the antigen to bind to the host.

In adaptive immunity, pathogen-specific receptors are “acquired” during the lifetime of the organism (whereas in innate immunity pathogen-specific receptors are already encoded in the germline). The acquired response is called “adaptive” because it prepares the body’s immune system for future challenges (though it can actually also be maladaptive when it results in autoimmunity).

The immune system is highly adaptable because of somatic hypermutation (a process of accelerated somatic mutations), and V(D)J recombination (an irreversible genetic recombination of antigen receptor gene segments). This mechanism allows a small number of genes to generate a vast number of different antigen receptors, which are then uniquely expressed on each individual lymphocyte. Since the gene rearrangement leads to an irreversible change in the DNA of each cell, all progeny (offspring) of that cell inherit genes that encode the same receptor specificity, including the memory B cells and memory T cells that are the keys to long-lived specific immunity.

A “T cell” or “T lymphocyte” is a type of lymphocyte (a subtype of white blood cell) that plays a central role in cell-mediated immunity. T cells can be distinguished from other lymphocytes, such as B cells and natural killer cells, by the presence of a T-cell receptor on the cell surface. They are called T cells because they mature in the thymus from thymocytes. The several subsets of T cells each have a distinct function. The majority of human T cells rearrange their alpha and beta chains on the cell receptor and are termed alpha beta T cells ($\alpha\beta$ T cells) and are part of the adaptive immune system. Specialized gamma delta T cells, (a small minority of T cells in the human body, more frequent in ruminants), have invariant T cell receptors with limited diversity, that can effectively present antigens to other T cells and are considered to be part of the innate immune system. Effector T cell broadly includes various T cell types that actively respond to a stimulus, such as co-stimulation. This includes helper, killer, regulatory, and potentially other T cell types. One skilled in the art is familiar with different types of T cells and their respective roles in adaptive immune response.

A “human leukocyte antigen (HLA) system” is a gene complex encoding the major histocompatibility complex (MHC) proteins in humans. These cell-surface proteins are responsible for the regulation of the immune system in humans. The HLA gene complex resides on a 3 Mbp stretch within chromosome 6p21. HLA genes are highly polymor-

14

phic, which means that they have many different alleles, allowing them to fine-tune the adaptive immune system. The proteins encoded by certain genes are also known as antigens, as a result of their historic discovery as factors in organ transplants. Different classes have different functions:

HLAs encoding major histocompatibility complex (MHC) class I MHC class I (A, B, and C) molecules, which present peptides from inside the cell. “Major histocompatibility complex (MHC) class I” or “MHC class I” molecules are found on the cell surface of all nucleated cells in the body. Their function is to display peptide fragments of antigens from within the cell to cytotoxic T cells; this will trigger an immediate response from the immune system against a particular non-self antigen displayed with the help of an MHC class I protein. Because MHC class I molecules present peptides derived from cytosolic proteins, the pathway of MHC class I presentation is often called cytosolic or endogenous pathway.

Class I MHC molecules bind peptides generated mainly from degradation of cytosolic proteins by the proteasome. The MHC I peptide complex is then inserted via endoplasmic reticulum into the external plasma membrane of the cell. The epitope peptide is bound on extracellular parts of the class I MHC molecule. Thus, the function of the class I MHC is to display intracellular proteins to cytotoxic T cells (CTLs). However, class I MHC can also present peptides generated from exogenous proteins, in a process known as cross-presentation.

A normal cell will display peptides from normal cellular protein turnover on its class I MHC, and CTLs will not be activated in response to them due to central and peripheral tolerance mechanisms. When a cell expresses foreign proteins, such as after viral infection, a fraction of the class I MHC will display these peptides on the cell surface. Consequently, CTLs specific for the MHC:peptide complex will recognize and kill presenting cells. Alternatively, class I MHC itself can serve as an inhibitory ligand for natural killer cells (NKs). Reduction in the normal levels of surface class I MHC, a mechanism employed by some viruses during immune evasion or in certain tumors, will activate NK cell killing. Antigens or antigenic epitopes presented by MHC class II molecules are recognized by cytotoxic T cells.

HLAs encoding MHC class II (DP, DM, DOA, DOB, DQ, and DR) molecules, which present antigens from outside of the cell to T-lymphocytes. “Major histocompatibility complex class II” or “MHC class II” molecules are a family of molecules normally found only on antigen-presenting cells such as dendritic cells, mononuclear phagocytes, some endothelial cells, thymic epithelial cells, and B cells. The antigens presented by class II peptides are usually derived from extracellular proteins (not cytosolic as in class I); hence, the MHC class II-dependent pathway of antigen presentation is called the endocytic or exogenous pathway. Loading of a MHC class II molecule occurs by phagocytosis; extracellular proteins are endocytosed, digested in lysosomes, and the resulting epitopic peptide fragments are loaded onto MHC class II molecules prior to their migration to the cell surface. Antigens or antigenic epitopes presented by MHC class II molecules are recognized by T helper cells and stimulate the multiplication of T-helper cells, which in turn stimulate antibody-producing B-cells to produce antibodies to that specific antigen. Self-antigens are suppressed by regulatory T cells.

An “antigen-presenting cell (APC)” is a cell that displays antigen complexed with major histocompatibility complexes (MHCs) on their surfaces; this process is known as antigen presentation. T cells may recognize these complexes using

their T cell receptors (TCRs). These cells process antigens and present them to T-cells. Antigen-presenting cells fall into two categories: professional and non-professional. Those that express MHC class II molecules along with co-stimulatory molecules and pattern recognition receptors are often called professional antigen-presenting cells. The non-professional APCs express MHC class I molecules.

Professional APCs specialize in presenting antigen to T cells. They are very efficient at internalizing antigens, either by phagocytosis (macrophages and dendritic cells) or by receptor-mediated endocytosis (B cells), processing the antigen into peptide fragments and then displaying those peptides, bound to a class II MHC molecule, on their membrane. [1] The T cell recognizes and interacts with the antigen-class II MHC molecule complex on the membrane of the antigen-presenting cell. An additional co-stimulatory signal is then produced by the antigen-presenting cell, leading to activation of the T cell. The expression of co-stimulatory molecules and MHC class II are defining features of professional APCs.

Almost all cell types can serve as a non-professional APC. They are found in a variety of tissue types. Professional antigen-presenting cells, including dendritic cells, mononuclear phagocytes, thymic epithelial cells, and B cells, present foreign antigens to helper T cells, while other cell types can present antigens originating inside the cell to cytotoxic T cells. In addition to the MHC family of proteins, antigen presentation relies on other specialized signaling molecules on the surfaces of both APCs and T cells.

A “B lymphocyte” or “B cell” is a type of white blood cell of the lymphocyte subtype. B cells function in the humoral immunity component of the adaptive immune system by secreting antibodies. Additionally, B cells present antigen (they are also classified as professional antigen-presenting cells (APCs)) and secrete cytokines. In mammals, B cells mature in the bone marrow, which is at the core of most bones. B cells express B cell receptors (BCRs) on their cell membrane. BCRs allow the B cell to bind a specific antigen, against which it will initiate an antibody response.

“Cancer immunotherapy” refers to a type of cancer treatment designed to boost the body’s natural defenses to fight the cancer. It uses substances either made by the body or in a laboratory to improve or restore immune system function.

“Tumor specific antigen (TSA)” or “tumor associated antigen (TAA)” refers to a protein that is specifically expressed or upregulated in cells of the respective tumor, as compared to non-cancerous cells of the same origin. A tumor specific antigen, or epitopes derived therefrom, can be recognized by the immune system to induce an immune response. Herein, the terms “tumor associated antigen” and “tumor specific antigen” are used interchangeably. The tumor specific antigen may be from all protein classes, e.g., enzymes, receptors, transcription factors, etc.

A “heteroclitic epitope” or “heteroclitic analog” refers to an altered version of an endogenous peptide sequence (i.e., an analog) engineered to elicit potent immune reactions. Heteroclitic epitopes have increased stimulatory capacity or potency for a specific T cell, as measured by increased responses to a given dose, or by a requirement of lesser amounts to achieve the same response and therefore provide benefit as vaccine components since these epitopes induce T cell responses stronger than those induced by the native epitope.

A “self-antigen” refers to an antigen that originates from within the body. The immune system usually does not react to self-antigens under normal homeostatic conditions. Epitopes from self-antigens (i.e., self-epitopes) are found in

high concentration on the surface of Antigen-presenting cells (APC’s) in association with its major histocompatibility complex (MHC) are known as dominant epitopes. These are stimulants of negative selection mechanisms to remove potentially self-destructing autoreactive T cells. Their “self” antigens are displayed to a developing T-cell and signal those “self-reactive” T-cells to die via programmed cell death (apoptosis) and thereby deletion from the T cell repertoire, preventing autoimmunity.

10 A “cryptic epitope” refers to an epitope derived from a self-antigen that does not necessarily undergo antigen processing/presentation and are ‘hidden’ from immune recognition. Cryptic epitopes usually appear in very low concentration on APC and do not delete auto-reactive T cells. 15 Cryptic epitopes are not presented for recognition by T cells unless they are produced in unusually large concentrations or unless they are freed from the configuration of their native antigen. Cryptic epitopes derived from tumor-specific antigens may be used to break the tolerance of T cells to the 20 tumor and induce potent immune response against the tumor. Such principles have been described in Pardoll, et al., *PNAS*, Vol. 96, pp. 5340-5342 (1999), the entire contents of which are incorporated herein by reference.

A “neoepitope” refers to an antigenic epitope generated 25 via random somatic mutations occurring in tumor cells. Neoepitopes are usually derived from individually specific tumor antigens or unique antigens and is thus specific to the lineage of tumor cells it is derived from. Neoepitopes are regarded in the art to be responsible for the immunogenicity 30 of tumors ((Srivastava et al., 1993, Duan et al., 2009; van der Bruggen et al., 2013), and mathematic modeling has predicted the existence of tens to hundreds of neoepitopes in individual human tumors (Srivastava 2009). The recent revolution in high-throughput DNA sequencing and accompanying bioinformatics approaches has finally made it possible 35 to actually identify the individually specific neoepitopes in individual cancers.

“Cancer vaccine,” as used herein, refers to a composition 40 that induces tumor-specific immunoresponse against a tumor or a tumor-specific antigen. Such immunoresponse is effective in inhibiting tumor growth and/or preventing reoccurrence of tumor.

An “intron” refers to any nucleotide sequence within a 45 gene that is removed by RNA splicing during maturation of the final RNA product. The term intron refers to both the DNA sequence within a gene and the corresponding sequence in RNA transcripts. Sequences that are joined together in the final mature RNA after RNA splicing are exons. Introns are found in the genes of most organisms and many viruses, and can be located in a wide range of genes, including those that generate proteins, ribosomal RNA (rRNA), and transfer RNA (tRNA). When proteins are generated from intron-containing genes, RNA splicing takes place as part of the RNA processing pathway that follows transcription and precedes translation.

An “exon” refers to any part of a gene that will become 50 a part of the final mature RNA produced by that gene after introns have been removed by RNA splicing. The term exon refers to both the DNA sequence within a gene and to the corresponding sequence in RNA transcripts. In RNA splicing, introns are removed and exons are covalently joined to one another as part of generating the mature messenger RNA.

“RNA splicing” refers to the processing of a newly 55 synthesized messenger RNA transcript (also referred to as a primary mRNA transcript). After splicing, introns are removed and exons are joined together (ligated) for form

mature mRNA molecule containing a complete open reading frame that is decoded and translated into a protein. For nuclear-encoded genes, splicing takes place within the nucleus either co-transcriptionally or immediately after transcription. The molecular mechanism of RNA splicing has been extensively described, e.g., in Pagani et al., *Nature Reviews Genetics* 5, 389-396, 2004; Clancy et al., *Nature Education* 1 (1): 31, 2011; Cheng et al., *Molecular Genetics and Genomics* 286 (5-6): 395-410, 2014; Taggart et al., *Nature Structural & Molecular Biology* 19 (7): 719-2, 2012, the contents of each of which are incorporated herein by reference. One skilled in the art is familiar with the mechanism of RNA splicing.

“Alternative splicing” refers to a regulated process during gene expression that results in a single gene coding for multiple proteins. In this process, particular exons of a gene may be included within or excluded from the final, processed messenger RNA (mRNA) produced from that gene. Consequently, the proteins translated from alternatively spliced mRNAs will contain differences in their amino acid sequence and, often, in their biological functions. Notably, alternative splicing allows the human genome to direct the synthesis of many more proteins than would be expected from its 20,000 protein-coding genes. Alternative splicing is sometimes also termed differential splicing. Alternative splicing occurs as a normal phenomenon in eukaryotes, where it greatly increases the biodiversity of proteins that can be encoded by the genome; in humans, ~95% of multi-exonic genes are alternatively spliced. There are numerous modes of alternative splicing observed, of which the most common is exon skipping. In this mode, a particular exon may be included in mRNAs under some conditions or in particular tissues, and omitted from the mRNA in others. Abnormal variations in splicing are also implicated in disease; a large proportion of human genetic disorders result from splicing variants. Abnormal splicing variants are also thought to contribute to the development of cancer, and splicing factor genes are frequently mutated in different types of cancer. The regulation of alternative splicing is also described in the art, e.g., in Douglas et al., *Annual Review of Biochemistry* 72 (1): 291-336, 2003; Pan et al., *Nature Genetics* 40 (12): 1413-1415, 2008; Martin et al., *Nature Reviews* 6 (5): 386-398, 2005; Skotheim et al., *The international journal of biochemistry & cell biology* 39 (7-8): 1432-49, 2007, the entire contents of each of which is incorporated herein by reference.

A “coding frame” or “open reading frame” refers to a stretch of codons that encodes a polypeptide. Since DNA is interpreted in groups of three nucleotides (codons), a DNA strand has three distinct reading frames. The double helix of a DNA molecule has two anti-parallel strands so, with the two strands having three reading frames each, there are six possible frame translations. A functional protein may be produced when translation proceeds in the correct coding frame. An insertion or a deletion of one or two bases in the open reading frame causes a shift in the coding frame that is also referred to as a “frameshift mutation.” A frameshift mutation typical results in premature translation termination and/or truncated or non-functional protein.

The term “proteome” refers to the entire set of proteins expressed by a genome, cell, tissue, or organism at a certain time. More specifically, it is the set of expressed proteins in a given type of cell or organism, at a given time, under certain conditions. The term is a blend of proteins and genome. “Proteome-wide” refers to each and every protein in the proteome without any bias.

The term “genome” refers to the genetic material of a cell or organism. It typically includes DNA (or RNA in the case of RNA viruses). The genome includes both the genes, the coding regions, the noncoding DNA, and the genomes of the mitochondria and chloroplasts. A genome does not typically include genetic material that is artificially introduced into a cell or organism, e.g., a plasmid that is transformed into a bacteria is not a part of the bacterial genome.

A “programmable DNA-binding protein,” as used herein, refers to DNA binding proteins that can be programmed to navigate to any desired target nucleotide sequence within the genome. To program the DNA-binding protein to bind a desired nucleotide sequence, the DNA binding protein may be modified to change its binding specificity, e.g., zinc finger nuclease (ZFN) or transcription activator-like effector proteins (TALE). ZFNs are artificial restriction enzymes generated by fusing a zinc finger DNA-binding domain to a DNA-cleavage domain. Zinc finger domains can be engineered to target specific desired DNA sequences and this enables zinc-finger nucleases to target unique sequences within complex genomes. Transcription activator-like effector nucleases (TALEN) are restriction enzymes that can be engineered to cut specific sequences of DNA. They are made by fusing a TAL effector DNA-binding domain to a DNA cleavage domain (a nuclease which cuts DNA strands). Transcription activator-like effectors (TALEs) can be engineered to bind practically any desired DNA sequence, so when combined with a nuclease, DNA can be cut at specific locations. The restriction enzymes can be introduced into cells, for use in gene editing or for genome editing in situ, Methods of programming ZFNs and TALEs are familiar to one skilled in the art. For example, such methods are described in Maeder, et al., *Mol. Cell* 31 (2): 294-301, 2008; Carroll et al., *Genetics Society of America*, 188 (4): 773-782, 2011; Miller et al., *Nature Biotechnology* 25 (7): 778-785, 2007; Christian et al., *Genetics* 186 (2): 757-61, 2008; Li et al., *Nucleic Acids Res* 39 (1): 359-372, 2010; and Moscou et al., *Science* 326 (5959): 1501, 2009, the entire contents of each of which are incorporated herein by reference.

A “guide nucleotide sequence-programmable DNA-binding protein,” as used herein, refers to a protein, a polypeptide, or a domain that is able to bind DNA, and the binding to its target DNA sequence is mediated by a guide nucleotide sequence. Thus, it is appreciated that the guide nucleotide sequence-programmable DNA-binding protein binds to a guide nucleotide sequence. The “guide nucleotide” may be a RNA molecule or a DNA molecule (e.g., a single-stranded DNA or ssDNA molecule) that is complementary to the target sequence and can guide the DNA binding protein to the target sequence. In some embodiments, the guide nucleotide sequence is an oligonucleotide sequence. As such, a guide nucleotide sequence-programmable DNA-binding protein may be a RNA-programmable DNA-binding protein (e.g., a Cas9 protein), or an ssDNA-programmable DNA-binding protein (e.g., an Argonaute protein). “Programmable” means the DNA-binding protein may be programmed to bind any DNA sequence that the guide nucleotide targets.

In some embodiments, the guide nucleotide sequence exists as a single nucleotide molecule and comprises comprise two domains: (1) a domain that shares homology to a target nucleic acid (e.g., and directs binding of a guide nucleotide sequence-programmable DNA-binding protein to the target); and (2) a domain that binds a guide nucleotide sequence-programmable DNA-binding protein. In some embodiments, the guide nucleotide is a guide RNA (gRNA). In some embodiments, domain (2) of the gRNA corresponds

to a sequence known as a tracrRNA, and comprises a stem-loop structure. For example, in some embodiments, domain (2) is identical or homologous to a tracrRNA as provided in Jinek et al., *Science* 337:816-821(2012), the entire contents of which is incorporated herein by reference. Other examples of gRNAs (e.g., those including domain 2) can be found in U.S. Provisional Patent Application, U.S. Ser. No. 61/874,682, filed Sep. 6, 2013, entitled "Switchable Cas9 Nucleases And Uses Thereof," and U.S. Provisional Patent Application, U.S. Ser. No. 61/874,746, filed Sep. 6, 2013, entitled "Delivery System For Functional Nucleases," the entire contents of each are hereby incorporated by reference in their entirety.

Because the guide nucleotide sequence hybridizes to target DNA sequence, the guide nucleotide sequence-programmable DNA-binding proteins are able to be targeted, in principle, to any sequence specified by the guide nucleotide sequence. Methods of using guide nucleotide sequence-programmable DNA-binding protein, such as Cas9, for site-specific cleavage (e.g., to modify a genome) are known in the art (see e.g., Cong, L. et al. *Science* 339, 819-823 (2013); Mali, P. et al. *Science* 339, 823-826 (2013); Hwang, W. Y. et al. *Nature biotechnology* 31, 227-229 (2013); Jinek, M. et al. *eLife* 2, e00471 (2013); Dicarlo, J. E. et al. *Nucleic acids research* (2013); Jiang, W. et al. *Nature biotechnology* 31, 233-239 (2013); the entire contents of each of which are incorporated herein by reference).

It is to be understood that any DNA binding domain that is programmable by a guide nucleotide sequence may be used in accordance with the present disclosure. For example, in some embodiments, the guide nucleotide sequence-programmable DNA binding protein may be a Cas9 protein, or a variant thereof. One skilled in the art would understand that the present disclosure is not limited to the use of Cas9 as the guide nucleotide sequence-programmable DNA binding protein, but that other DNA binding proteins that adopt similar mechanism of target sequence binding may also be used.

As used herein, the term "Cas9" or "Cas9 nuclease" refers to an RNA-guided nuclease comprising a Cas9 protein, a fragment, or a variant thereof. A Cas9 nuclease is also

referred to sometimes as a casnI nuclease or a CRISPR (clustered regularly interspaced short palindromic repeat)-associated nuclease. CRISPR is an adaptive immune system that provides protection against mobile genetic elements (viruses, transposable elements and conjugative plasmids). CRISPR clusters contain spacers, sequences complementary to antecedent mobile elements, and target invading nucleic acids. CRISPR clusters are transcribed and processed into CRISPR RNA (crRNA). In type II CRISPR systems correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (mc) and a Cas9 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves linear or circular dsDNA target complementary to the spacer. The target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3'-5' exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gRNA") can be engineered so as to incorporate aspects of both the crRNA and tracrRNA into a single RNA species. See, e.g., Jinek et al., *Science* 337:816-821(2012), the entire contents of which is incorporated herein by reference.

Cas9 nuclease sequences and structures are well known to those of skill in the art (see, e.g., Ferretti et al., *Proc. Natl. Acad. Sci.* 98:4658-4663(2001); Deltcheva E. et al., *Nature* 471:602-607(2011); and Jinek et al., *Science* 337:816-821 (2012), the entire contents of each of which are incorporated herein by reference). Cas9 orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski et al., (2013) *RNA Biology* 10:5, 726-737; the entire contents of which are incorporated herein by reference. In some embodiments, wild type Cas9 corresponds to Cas9 from *Streptococcus pyogenes* (*NCBI Reference Sequence*: NC_002737.2, SEQ ID NO: 4 (nucleotide); and Uniport Reference Sequence: Q99ZW2, SEQ ID NO: 1 (amino acid)).

(SEQ ID NO: 4)

```

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGTGATCAC
TGATGAATAATAAGGTTCCGCTAAAAAGTCAAGGTTCTGGAAATACAGACGCCAACAGTATCA
AAAAAAATCTTATAGGGGCTCTTTATTGACAGTGGAGAGACGGGAAGCGACTCGCTCTCAA
CGGACAGCTCTGAGAGGTATACACGTCGGAAAGATCTGTTATCTACAGGAGATTTTTCA
AATGAGATGGCAGAAAGTAGATGATAGTTCTTCATCGACTTGAAGAGCTTTTGGTGGAAAGAA
GACAAGAAAGCATGAACGTCTACCTTATTTGAAATATAGTAGATGAAAGTGTCTTATCATGAGAAA
TATCCAACTATCTATCTGCGAAAAAAATGGTAGATCTACTGTAAAGCGGATTGGCGCTTA
ATCTATTTGGCCTTAGCGCATATGATTAAGTTCTGTTGATTTGATTGAGGGAGATTAAATC
CTGATAATAGTGTGGACAAACTATTTCAGCTGGTACAAACCTACAATCAATTATTGAAAG
AAAACCTATTAACGCAAGTGGAGTAGATGCTAAAGCGATTCTTCGACGATTGAGTAATCAA
GACGATTAGAAAATCTCATGGCTCAGCTCCCCTGGAGAAGAAAATGGCTTATTGGGAAATCTCA
TTGCTTGTCTTGGGGTATGCCCTAAATCAATTGGTGGCAGAAGATGCTAAATT
ACAGCTTCAAAAGATACTCGATGATGTTAGATAATTATGGCCAAATGGAGATCAATA
TGCTGATTGTTTGGCAGCTAAGAATTTCAGATGCTATTTACTTCAGATATCTAAGAGTA
AATACTGAAATAACTAAGGCTCCCTATCAGCTTAAGTAAACGCTACGATGAAACATCATCAA
GACTTGACTCTTTAAAGCTTAGTCAGCAACAACTCCAGAAAAGTAAAGAAATCTTTT
GATCAATCAAAAACCGATAATGCAAGGTTATTGATGGGGGAGCTAGCAAGAAGAATTTATAA
ATTTATCAACCAATTAGAAAAATGGATGGTACTGAGGAATTATTGGTAAACTAAATCGTGA
AGATTTGCTGGCAAGCAACGGACCTTGAACAACGGCTTATTCCCCATCAAATTCACTGGGTGA
GCTGCATGCTATTGGAGAAGACAAGAAGACTTTATCATTCTTAAAGACATCGTGAGAAGAT
TGAAAAAAATCTGACTTTCTGAAATTCTTATATGTTGGTCCATTGGCCGCTGGCAATAGTCGTTT
GCATGGATGACTCGGAAGTCGAAGAAACAACTACCCCATGGATTGGAGAAGTGTGCTGATAA
AGGTGCTTCAGCTCAATCTTATTGAAACGCTGACAACACTTGTATAAAAGTCTCCAAATGAAA
AGTACTACCAAAACATAGTTGCTTTATGAGTATTTCAGGTTATAACGAATTGACAAAGGTCAA
ATATGTTACTGAAGGAATGCGAAAACCAGCATTCTTCAGGTGAACAGAAGAAAGCATTGTT
ATTTACTCTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAGAAGATTATTCAAAAAAA
ATAGAATGTTTGATAGTGTGAAATTTCAGGAGTTGAAGATAGTTAATGCTTCAAGGTACC

```

- continued

TACCATGATTTCTAAAAATTAAAGATAAAGATTTTTGATAATGAAGAAAATGAAGATATC
 TTAGAGGATATTGTTAACATTGACCTTATTGAGATAGGGAGATGATTGAGGAAAGACTAAA
 ACATATGTCACCTCTTGATGATAAGGTGATGAAACAGCTAACGTCGCCGTATACTGGTTGG
 GGACGTTGTCGAAATTGATTAATGGTATTAGGATAACGAACTCGCAAAACAATTAGAT
 TTTTGAAATCAGATGGTTGCAATCGAATTATGCACTGATGATGATAGTTGACAT
 TAAAGAAGACATTCAAAGACAAGTGTCTGACAAGGCAGATTTACAGACTGAAAGTTGATGAAATTG
 AATTAGCTGGTAGGCCGCTTAAAGGAGATCGAGACTGTTACAGACTGAAAGTTGATGAAATTG
 GTCAGAAAGTAACTGGGGCGCATAAGCAGAAAATATCGTATTGAAATGGCACGTCAGAAATCAGAC
 AACTCAAAGGGCAGAAAATTCGCGAGAGCTGAAAGCACTGAAAGAAGGTATCAAAAGAA
 TAGGAAGCTCAGATTCTAAAGAGCATCCTGTGAAAATACTCAATTGCAAAATGAAAAGCTCTAT
 CTCCTATTATCTCCAAATGGAGAGACATGATGTCGACAAAGTTCTTAAAGACGATTCAATAGACAATAAGGTC
 TAAACGCGTCTGATAAAATCGTGGTAAATCGGATAACGTCAGTGAAGAAGTAGTCAGAAA
 GATGAAAACATTGGAGACAATTCTAAACCCAAGTTAATCACTCAACGTAAGTTGATAATT
 AACGAAAGCTGAACGTCGAGGTTGAGTGAACCTTGATAAAAGCTGTTTATCAAACGCCAATTGG
 TTGAAACTCGCCAAATCAACTAGCATGTCGACAAAGTTGAGTGAACGTCGATGAAATAACAAATACG
 ATGAAAATGATAAACTTCTGGAGAGTTAAAGTATTGATCTAAATTAGTTCTGACT
 TCCGAAAAGATTCCAATTCTATAAAAGTACCTGAGATTAAACATTACCATCATGCCATGCGT
 ATCTAAATGCCGCTGGAACTGCTTGATTAAGAAAATATCCTAAACATTGAACTGGAGTTGCT
 ATGGTGATTATAAGTTATGATGTTGTAATGATTGCTAAGTCTGAGCAAGAAATAGGCAA
 GCAACCCAAAATTTCTTACTCTAAATCATGAAACTCTTCAAAACAGAAAATTACACTTGCA
 ATGGAGAGATTCGCAACGCCCTAACTCGAAAACTATGGGAAACTGGAGAAATTGCTGGGA
 TAAAGGGCAGAGTTGCGCACAGTGGCCAAAGTATTGTCATGCCCAAGTCATAATTGTCAGAAA
 AACAGAAGTACAGACAGCGGATTCTCAAGGAGTCATTACCAAAAGAAATTGGACAAGC
 TTATTGCTCGTAAAAAGACTGGGATCCTTAAAGTGGGAAAGGGAAATCGAGAAAGTTAACCGTAAAGAG
 TTACTAGGGATCACAATTATGGAAAGAAGTTCTTGAAGGAAATTCCGATTGACTTTAGAGCT
 AAAGGATATAAGGAAGTAAAGACTTAATCATTAACCTAAATAGTCTTTTGAGTTA
 GAAAACGGTCGTAACGGATGCTGGCTAGTGGCGGAAATTACAAAAGGAAATGAGCTGGCTCT
 GCCAAGAAATATGGTATTGAGCTAGTCATTAGTAAAAGGTTGAAGGGTAGTCAG
 AGATAACGAAACAAAATGGTGTGAGCAGATAAGCATTAGTGGAGATTGAGC
 AAATCAGTGAATTCTAAGCGTGTATTAGCAGATGCCATTAGATAAAAGTTCTAGTGCAT
 ATAACAAACATAGAGACAAACAAACGTGAACAAGCAGAAAATTATTACATTATTCGTT
 ACGAATCTGGAGCTCCGCTGTTAAATATTGATACAACATTGATGCTAACGATATAACG
 TCTACAAAAGAAGTTTAGATGCCACTCTTATCCATCAATCCACTGGCTTATGAAACACGC
 ATTGATTGAGTCAGCTAGGAGGTACTGA

(SEQ ID NO: 1)

MDKKYSIGLDIGTNsvgwaiTDEYKPSKKFVLGNTDRHSIKKNLIGALLEDSGETAEARLKRTR
RRYTRRKNRICYLQEISNEMAKVDDSFHRLIESFLEEDKKHERHPIFGNIVDEVAYHEKPYTIYHLR
KKLVDSTDKDILRLIYLALAHMIFRGHFLIEGDLNPDNSDVDKLFQIQLVQTYNQFEEPINASGVDA
KAILSARLSKSRLLENLIAQLPGEKKNGLFGNLIALSGLTPNFKNFDLAEDAKLQLSKDYYDDLDN
LIAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
YKEIFFDQSNSKGYAGYIDGGASQEEFYKFIKPILEKMDGEELLVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRQEDFPFLKDNREKIEKILTFRIPIYVGPLARGNSRFAMTRKSEETITPWNFEVVDKGA
SQSFIERMTNDPKNLPNEKVKLPHSLLYEFTVYNELTKVKKVTEGMRKPAPFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECPDSVEISVGEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVTLT
LFEDREMIERLKYAHFLDDKVMQLKRRYTGWRSLRKLINGIRDQSGKTIIDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDLSHEHIANLAGSPAIKGILQTVKVDDELVKVMGRHKPENIVI
EMARENQTTQKGOKNSRERMKRIEEGIKELGSQILKEHPVENTQLONEKLYLYLQNGRDMDVQEL
DINRLSDYDVDHIVPQSLKDDIDSIDNKVLTRSDKNRGKSDMPSEEVVKMKNYWQRQLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKROLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSD
FRKDFQFYKVRINNYHHADAYLNAVGTALIKYPKLESEFVYGDYKVYDVRKMIAKSEOEIGKA
TAKYFFYSNIMNNFKTEITLANGEIRKRPLIETNGETGEIVWDKGDFATVRKVL SMPQVNIVKKTEVO
TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKEKGSKKLKSVKELLGITIME
RSSFEKPNIDFLEAKGYKEVKDDLIILKLPKYSLFELENGRKMLASAGELQKGNEALALPSKVVNPLYLA
SHYEKLKGSPEDNEQKQLFVEQHKYLDEIIEQISEFSKRVILADNLKVLSAYNKHRDKPIREQAENI
IHLFTLTNLGAPAFAFKYFDTTIDRKRTSTKEVLDATLHQSTIGLYETRIDLSQLGGD
 (single underline: HNH domain; double underline: RuvC domain)

In some embodiments, wild type Cas9 corresponds to Cas9 from *Streptococcus aureus*. *S. aureus* Cas9 wild type

(SEQ ID NO: 6)

MKRNYILGLDIGITSVGYGIIDYETRVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKK
 LIPDYNLLTDHSELSGINPYEARVKGLSQKLSSEEFSAAALLHLAKRGRVHNVNEVEEDTGNELSTKEQI
 SRNSKALEEKYVAELQLERLKKDGEVGRSINRFKTSVDYVKEAKQQLKVKQAYHOLDQSFIDTYIDLLE
 TRRTYYEGPGEGPSFGWDKPYEMLMGHTYFPEELRSVVKYAYNADLYNALNDLNNLVITRDENE
 KLEYYEKFQIIEENVFKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIEN
 AELLDQIAKILTYQSSEDIQEELTNLSELQEEIEQISNLKGYTGTNLSLKAINTLILDELWHTNDNQIA
 IFNRLKLVPKKVDSLQQKEIPTLWDDFILSPVVKRSFIQSIVNIAIKYGLPNDIIIELAREKNSKDAQK

- continued

```

MINEMQKRNRQTNERIEEI IRTGKENAKYLIIEKIKLHDQEGKCLYSLEAIPLEDLLNNPNEYVDHI IP
RSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLISSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEER
DINRFVSKQDFINRNLVDTRYATRGLMNLLRSYFRVNVLDVKVSINGGFTSFLRRKWKPKKERKNGY
KHAEDALIIANADIFKEWKKLDKAKKVMENQMFEEKQAESPEIETEQEYKEIFITPHQIKHIDKDF
DYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVMNLNLNGLYDKNDKLKLINKSPEKLLMLYHDP
QTYQLKLIMEQYGDEKNPLNLYEETQYGNLTYSKSKDNGPVIKKIYVGKLNNAHLDITDDYPNSR
NKVVKLSLKPYRDVYLDNGVYKFVTVKNLDVICKENYYEVNSCYEEAKLKKISNQAFIASFYNN
DLIKINGELYRVIGVNNNDLNRIEVNMIDITYREYLEMNMDRPPRIIKTIASKTQSICKYSTDILGNLYE
VSKKKHPQIIKKG

```

In some embodiments, wild type Cas9 corresponds to Cas9 from *Streptococcus thermophilus*.

Streptococcus thermophilus wild type CRISPR3 Cas9 (St3Cas9)

(SEQ ID NO: 7)

```

MTKPYSIGLDIGTNSVGWAVITDNYKVPSKKMKVLGNTSKYIKKNLLGVLLFDGSITAEGRRLKRTA
RRRYTRRRNRIRYLQEIFTSTEMATLDDAFFQRLLDSFLVPDDKRDTSKYPFGNLVEEKVYHDFPTIYHL
RKYLAQSTKKAKLRLVYLAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDSLSENSKQ
LEEIVKDCKISLEKKDRKLKPGKEKNNSGIEFLKLIVGNQADFRKCFNLDEKASLHESKESYDEDLETL
LGYIGDDYSDVFLKAKKLYDAILLSGFLVTDNETEAPLSSAMIKRYNEHEDLALLKEYIRNLSKTYN
EVFKDDTKNGYAGYIDGKTNQEDFYVYLNKLLAEFEGADYFLEKIDREFDLRKQRTFDNGSIPYQIHLQ
EMRAILDKQAKYPPFLAKNKEIREKTFPIYYVGPLARGNSDFAWSIRKRNEKITPWNPEDVIDKESS
AEAFINRMRITSFDYLPEEKVPLKHSLLYETPNVYNELTKVPRFIAESMDYQFLDSKQKKDIVRLYFKDK
RKVTDKDIIYEYLHAIYGDIGIELKGIEKQFNSSLSTYHDLLNIINKEFLDDSNNEAIIEBIIHLTIFEDRE
MIKQRLSKFENIFDKSVLKKLSSRRHYTGWGLSAKLINGIRDEKSGNTILDYLIIDDGISNRNFMQLIHDD
ALSFKKKIQKAQIIGGEDKGNIKIVEVVKSLPGSPA1KKGILQSIKIVDELVVKVMGGRKPEIVVEMARENQ
YTNGQKSNSQQLRKLREKSLKGSKILNPAKLSKIDNNALQNDRLLYLYLQNGKDMYTGDDLDI
DRLSNYDIDHIIQPAFLKDNSIDNKVLUVSSASNNGKSDDPFSLEVVKRKTFWYQLLKSKLISQRKFDNL
TKAERGGLLPEDKAGFIQRQLVETRQITKVARLLDEKENNKDEENNRAVRTVIITLKSTLVSQFRKD
FELYKVREINDEPHAAHDAYLNAVIASALLKKYKPKEPEFVYGDYPKYNNSFRERKSATEKVYFYSNIMNI
FFKKSISLADGRVIERPLIEVNEETGESVWNKESDLATVRRVLSYPQVNVVKKVEBQNHGLDRGKPKGL
FNAKLSSKPKPNSENLVGAKEYLDPKVKYGGYAGISNSFAVLVKGTEIKGAKKKITNVLFQOGISILDRI
NYRKDKLNLFLLEKGKYDIELIELPKYSLFELSDGSRRMLASILSTNNKRGIEIHGNQIFLSQPKVKKLYH
AKRISNTINENHRKYVENHKKEFEELFYIILEFNENYVGAKNGKLLNSAFQSWQNHSIDECLCSSFIGPT
GSERKGLFELTSRGSAADFEFLGVKIPRYRDTPSSLKDATALHQSVTGLYETRIDLAKLGE

```

Streptococcus thermophilus CRISPR1 Cas9 wild type (St1Cas9)

(SEQ ID NO: 8)

```

MSDLVGLDIGIGSVGVGILNKTGEI IHKNSRIPPAQAENNLVRRTNRQGRRLTRRKHRRVRLNRL
FEESGLITDFTKISINLNPyQLRVKGTDELSNEELFIALKNMVKHRGISYLDASDDGNSSIGDYAQIVK
ENSKQLETKTPCQIQLERYQTGYQLRGDFTEVKDGKKHRLINVFTSAYRSEALIRLQTOQEFPNQITDE
FINRYLEILTGKRKYHHGPNGNEKSRTDYGRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNL
LNDLNNLTVPTEKKLSEKQKQIINYVKNEKAMGPALKFQYIAKLLSCDVAIKGYRIDKSGKAEIHT
FEAYRKMKTLETDIEQMDRETLKLAYVLTNTEREQIEAHEFADGSFSQKQVDELVQFRKANS
SFIGKGWHNFSVKLMMBELIPELYETSEQMTLTRLGKQTTSSNKTQYIDEKLLTEEI YNPVVAKSVR
QAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKELPHSV
FHGHKOLATKRLWHQGGERCLYTGTISIHDLINNSNQFEDHILPLSITFDDSLANKVLVYATANQE
KGQRTPYQALDSMDDAWSFRELKAFRESKTLSNKKKEYLILTQFDRKKFIERNLVDTTRYASR
VVLNALQEHFRAHKIDTKSVVRGQFTSQLRHWGIEKTRDTYHHHADALIAASSQLNLWKKQKN
TLVSYSEDQLLIDETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYAT
RQAKVGDKADETYVLGKIIYTDQYDAMPKIYKKDKSKFLMYRHDPQTFEKVIEPILLENYPNPKQI
NEKGKEVCPNPFLKYKEEHGYIKYSSKGNGPEIKSLKLYYDSSKLGHNIDITPKDSNNKVVLQSVSPWR
ADVYFNKTTGKYEILGLKYADLQFEGKTGTGKISQEKYNDIKKKCEGVDSDSEFKPTLYKNDLLLVKDT
ETKEQQLFRFLSRMPKQKHYVELKPYDKQKPEGGEALIKVGLNVANSGQCKGLGKSNIISYKVRTD
VLGNQHIIKNEGDKPKLDF

```

In some embodiments, the Cas9 domain of any of the fusion proteins provided herein is a Cas9 from archaea (e.g. nanoarchaea), which constitute a domain and kingdom of single-celled prokaryotic microbes. In some embodiments, the Cas9 domain is CasX or CasY, which have been described in, for example, Burstein et al., "New CRISPR-Cas systems from uncultivated microbes." *Cell Res.* 2017 Feb. 21. doi: 10.1038/cr.2017.21, which is incorporated herein by reference. Using genome-resolved metagenomics, a number of CRISPR-Cas systems were identified, including the first reported Cas9 in the archaeal domain of life. This divergent Cas9 protein was found in nanoarchaea as part of an active CRISPR-Cas system. In bacteria, two previously unknown systems were discovered, CRISPR-CasX and CRISPR-CasY, which are among the most compact systems yet discovered. In some embodiments, Cas9 refers to CasX, or a variant of CasX. In some embodiments, Cas9 refers to a CasY, or a variant of CasY. It should be appreciated that other RNA-guided DNA binding proteins may be used as a

nucleic acid programmable DNA binding protein (napD-NAbp) and are within the scope of this disclosure.

In some embodiments, the Cas9 domain comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring CasX or CasY protein. In some embodiments, the Cas9 domain is a naturally-occurring CasX or CasY protein. In some embodiments, the Cas9 domain comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of SEQ ID NOS: 336-337 or 3000. In some embodiments, the Cas9 domain comprises an amino acid sequence of any one SEQ ID NOS: 336-337 or 3000. It should be appreciated that CasX and CasY from other bacterial species may also be used in accordance with the present disclosure.

In some embodiments, wild-type Cas9 refers to CasX from *Sulfobolus islandicus* (strain REY15A).

(SEQ ID NO: 338)

```
MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEEDAAAEE
RRGKAKKKGEGETTSNIIPLSGNDKNPWETLKCYNPPTTVALSEVFKNFSQV
KECEEVSAPSFSVKPEFYKFRSPGMVERTRRVKLEVEPHYLIMAAAGWVLTRLGKA
KVSEGDYVGVNVPTRGILYSLIQNVNGIVPGIKPETAFLGLWIARKVVSSTPNVUS
VVIYTISDAVGQNPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSNMERYIV
LANYIYEYLTGSKRLEDLLYFANRDLIMMLNSDDGKVRDLKLISAYVNGLIRGEG
```

In some embodiments, wild-type Cas9 refers to CasX from *Sulfobolus islandicus* (strain REY15A).

(SEQ ID NO: 339)

```
MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEEDAAAEE
RRGKAKKKGEGETTSNIIPLSGNDKNPWETLKCYNPPTTVALSEVFKNFSQV
KECEEVSAPSFSVKPEFYEFGRSPGMVERTRRVKLEVEPHYLIAAAAGWVLTRLGKA
VSEGDYVGVNVPTRGILYSLIQNVNGIVPGIKPETAFLGLWIARKVVSSTPNVSV
VRIYTISDAVGQNPTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNMERYIVL
ANYIYEYLTGSKRLEDLLYFANRDLIMMLNSDDGKVRDLKLISAYVNGLIRGEG
```

In some embodiments, wild-type Cas9 refers to CasY from a *Parcubacteria* group bacterium.

CasY (ncbi.nlm.nih.gov/protein/APG80656.1)

>APG80656.1 CRISPR-associated protein CasY [uncultured *Parcubacteria* group bacterium]

(SEQ ID NO: 3000)

```
MSKRHPRISGVGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGRTVPREIVSAINDDYVGLYGLSNFD
DLYNAEKRNEEKVSVLDFWYDCVQYGAWSYTAGPLLKVNAEVRGGSYELTTKLKGSHLYDELQID
KVKFLNKKEISRANGSLDKLKKDIIDCFKAEYERERHKDQCNCNKLADDIKAALKDAGASLGERQKKLFR
DFFGISEQSENDKPSFTNPNLNTCCLLPFDTVNRRNRGEVLENKLKEYAQQLDKNEGSLEMWEYIGIG
NSGTAFSNFLGEGFLGLRRENKI TELKKAMMDITDAWRGQEQQEELEKRLRILAALTIKLREPKFDNHW
GGYRSDINGKLSSWLQNYINQTKVKEKDLGHKKDLKAKEMINRFGESDTKEEAVVSSLLESIEKIVP
DDSDADKEPDIPAIYRFLSDGRLTLNRFVQREDVQEALIKEERLEAEKKKPKRKKKSDAEDEKETI
DFKELFPHLAKPLKLKPVNFYGDSKRELKYKKKNAIYTDAWLKAVEKIKYSAFSSSLKNSFFDTDFDK
DFFIKRLQKIFSVYRPNFTDKWPVCKNSFAPYCDIVS LAENEVLYKPKQSRSRKSAIDKNRVRLPSTE
NTAKAGITALARELSVAGFDWKDLKLKEEHEBYIDLIELHTKALALLAVETEQLDISALDFVENGTVKD
FMKTRDGNLVLFGRFLEMFSQSIVFSELRLLAGLMSRKEFITSRAIQTMNQKQABELLYIPHEFQSAKITT
PKEMSRAFLDLPAPAEFATSLEPELSKSEKSLLKLKQMRYYPHFGYELTRTGQGIDGGVAENALREKSP
VKKREIKCQYKTLGRGNKIVLYVRSSYYQTQFLEWFLHRPKNVQTDVAVSGSFLIDEKKVKTRWN
YDALTALEPVSGSERFVFSQFTIPEKSAEEEGORYLGIDIGEYGIAYTALETGDSAKIILDQNFISDPO
LKTLRREEVKGLKDQRRGTFAMPSTKIAIRESLVHSLRNRIHHHLAKHKAKIVYELEVSRFEEGKQKIK
KVTATLKADVYSEIDADKNLQTTVWGKLVASEIISASYTSQFCACKLWRAEMQVDETITQBLIG
TVRVIKGTLIDAICDFMRPPIFDENDTPFPYKRDPCDKHHISKKMGNCLFICPFCRANADADIQASQ
TIALLRYVKEKKVEDYFERFRKLKNIKVLGQMKK
```

In some embodiments, Cas9 refers to Cas9 from: *Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrphidicola* (NCBI Ref: NC_021284.1); *Prevotella intermedia* (NCBI Ref: NC_017861.1); *Spiroplasma taiwanense* (NCBI Ref: NC_021846.1); *Streptococcus iniae* (NCBI Ref: NC_021314.1); *Belliella baltica* (NCBI Ref: NC_018010.1); *Psychrophlexus torquisl* (NCBI Ref: NC_018721.1); *Listeria innocua* (NCBI Ref: NP_472073.1), *Campylobacter jejuni* (NCBI Ref: YP_002344900.1) or *Neisseria meningitidis* (NCBI Ref: YP_002342100.1) or to a Cas9 from any of the organisms listed in Example 1 (SEQ ID NOS: 11-260).

To be used as in the fusion protein of the present disclosure as the guide nucleotide sequence-programmable DNA binding protein domain, a Cas9 protein needs to be nuclease

inactive. A nuclease-inactive Cas9 protein may interchangeably be referred to as a “dCas9” protein (for nuclease-“dead” Cas9). Methods for generating a Cas9 protein (or a fragment thereof) having an inactive DNA cleavage domain are known (See, e.g., Jinek et al., *Science*. 337:816-821(2012); Qi et al., (2013) *Cell*. 28; 152(5):1173-83, the entire contents of each of which are incorporated herein by reference). For example, the DNA cleavage domain of Cas9 is known to include two subdomains, the HNH nuclease subdomain and the RuvC1 subdomain. The HNH subdomain cleaves the strand complementary to the gRNA, whereas the RuvC1 subdomain cleaves the non-complementary strand. Mutations within these subdomains can silence the nuclease activity of Cas9. For example, the mutations D10A and H840A completely inactivate the nuclease activity of *S. pyogenes* Cas9 (Jinek et al., *Science*. 337:816-821(2012); Qi et al., *Cell*. 28; 152(5):1173-83 (2013)). dCas9 (D10A and H840A)

(SEQ ID NO: 2)

```
MDKKYSIGLAIGTNSVGAWITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDGETAEATRLKRTAR
RRYTRRKRNRCYLCQEIFSNEMAKVDDSFHRLIESFLVEEDKKHERHPFGNIVDEVAYHEKYPTIYHLR
KKLVDSTDADLRLIYLALAHMIKFRGHFLIEGDLNPNDNSDVDFKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNPDLAEDAKLQLSKDTYDDLDN
LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHQDLTLLKALVRQQLPEK
YKEIFFDQSXNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
```

- continued

GELHAILRRQEDFYPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKG
 SAQSFIERMNTFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAPLSGEQKKAIVDLLFK
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT
 LFEDREMIEERLKTYAHLFDDKVMQLKRRRTGWRGLSRKLINGIRDQSGKTILDFLKSDGFANRN
FMQLIHDDSLTPKEDIQKAQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDELVKVMGRHKPENIVI
EMARENQTTOKGQKNSRERMKRIEEGIKELGSQILKEHPVENTOLQNEKLYLYLQNGRDMYVQDQEL
DINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQOLLNAKLITQRK
FDNLTKAERGGLELDKAGFIKRQLVETRQITKHVAQILDLSRMNTKYDENDKLIREVKVITLKSCLVSD
FRKDFQFYKVREINNYHHAAHDAYLNAAVVTALIKKPKLESEFVYGDYKVDVRKMIAKSEQEIGKA
TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGDRFATVRKVLSPQVNIVKKTEVO
TGGFSKESILPKRNSDKLIARKKDWDPKYGGFDSPPTVAYSVLVVAKEKGSKKLKSVKELLGITIME
RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKMLASAGELQKGNEALPSKYVNFLYLA
SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENI
IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIGHQSITGLYETRIDLSQLGGD

(single underline: HNH domain; double underline: RuvC domain)

The dCas9 of the present disclosure encompasses completely inactive Cas9 or partially inactive Cas9. For example, the dCas9 may have one of the two nuclease domain inactivated, while the other nuclease domain remains active. Such a partially active Cas9 may also be referred to as a Cas9 nickase, due to its ability to cleave one strand of the targeted DNA sequence. The Cas9 nickase suitable for use in accordance with the present disclosure has an active HNH domain and an inactive RuvC domain and is able to cleave only the strand of the target DNA that is bound by the sgRNA. The Cas9 nickase of the present disclosure may comprise mutations that inactivate the RuvC domain, e.g., a D10A mutation. It is to be understood that any

mutation that inactivates the RuvC domain may be included in a Cas9 nickase, e.g., insertion, deletion, or single or multiple amino acid substitution in the RuvC domain. In a Cas9 nickase described herein, while the RuvC domain is inactivated, the HNH domain remains activate. Thus, while the Cas9 nickase may comprise mutations other than those that inactivate the RuvC domain (e.g., D10A), those mutations do not affect the activity of the HNH domain. In a non-limiting Cas9 nickase example, the histidine at position 840 remains unchanged. The sequence of an exemplary Cas9 nickase suitable for the present disclosure is provided below.

Cas9 Nickase (D10A) (SEQ ID NO: 3)
 MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTRHSIKKNLIGALLEDSGETAEATRLKRTAR
 RRYTRKRNICYLQEIFSNEAKVDDSFHRLLEESFLVEEDKKHERHPFGNIIDVAYHEKPYTIYHRL
 KKLVDSTDKDALKRLLTQYLAHMIFGRKHFLLIEGDLNDPNSDVKLFIQLVQTYNQLEENPINASGVDA
 KAILSARLSKSRLRLENLIAQQLPEKKNGLFGNNLIALSGLTPNFKSNPDLAEDAKLQLSKDITYDDLDN
 LLAQIGDQYADLFLAAKNLSDAILLSDILRVRNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
 YKEIFFDQSKNGYAGYIDGQASQEEFYKPIKPLIKEMDGTETEELLVKLNRDILLRKQRTFDNGSIIPHQIHL
 GELHAILRRQEDFYPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKG
 SAQSFIERMNTFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAPLSGEQKKAIVDLLFK
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT
 LFEDREMIEERLKTYAHLFDDKVMQLKRRRTGWRGLSRKLINGIRDQSGKTILDFLKSDGFANRN
FMQLIHDDSLTPKEDIQKAQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDELVKVMGRHKPENIVI
EMARENQTTOKGQKNSRERMKRIEEGIKELGSQILKEHPVENTOLQNEKLYLYLQNGRDMYVQDQEL
DINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQOLLNAKLITQRK
FDNLTKAERGGLELDKAGFIKRQLVETRQITKHVAQILDLSRMNTKYDENDKLIREVKVITLKSCLVSD
FRKDFQFYKVREINNYHHAAHDAYLNAAVVTALIKKPKLESEFVYGDYKVDVRKMIAKSEQEIGKA
TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGDRFATVRKVLSPQVNIVKKTEVO
TGGFSKESILPKRNSDKLIARKKDWDPKYGGFDSPPTVAYSVLVVAKEKGSKKLKSVKELLGITIME
RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKMLASAGELQKGNEALPSKYVNFLYLA
SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENI
IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIGHQSITGLYETRIDLSQLGGD
 (single underline: HNH domain; double underline: RuvC domain)

S. aureus Cas9 Nickase (D10A)

(SEQ ID NO: 5)

MKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRERRHRIQRVKK
 LLFDYNNLTDHSELSGINPYEARVKGLSQKLSEEEFSAAHLHLAKRRGVHNVNVEEEDTGNELSTKEQI
 SRNSKALEEKYVAELQLERLKKDGEVRGSINRFTKTSYVKEAKQQLLKVKQKAYHLDQSFIDTYIDLLE
 TRRTYYEGPGEGSPFGWKDIKEYEMLMGHTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENE
 KLEYYEKFQIIENVFQKQKKPTLQIAKEILVNEEDIKGYVRTSTGKPEFTNLKVYHDIKITARKEIIEN
 AELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGHNLSLKAINLILDELWHTNDNQIA

-continued

```

IFNRLKLVPKKVQLSQQKEIPTTLVDDFILSPVVKRSFIQSIVINAIKKYGLPNDIIIELAREKNSKDAQK
MINEMQKRNRQTNERIEEEIRTGKENAKYLIEKIKLHDMQEGKCLYSLEAPILEDLLNNPFNYEVDHIIIP
RSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSLSSDKISYETFKKHILNLAKGKRISKTKKEYLLEER
DINRFPSVQKDFINRNLVDTRYATRGLMNLRLSRYFRVNNLDVKVSINGGFTSFLRRKWFKKERNKGY
KHHAEADALIIANADIFTKEWKKLDKAKKVMENQMFEEKQAESMPITEQEYKEFIFTPHQIKHICDFK
DYKQSHRVDKKPNEELINDTLYSTRKDKGNTLIVNNNLNGLYDKDNDKLKLINKSPEKLLMYHDP
QTYQKQLKLIMEQYGDEKPLYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNNAHLDITDDYPNSR
NKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKLKKISNQAEFIASFYNN
DLIKINGELYRIVIGVNNDLNLRIEVNMIDITYREYLENMNDKRPPIIKTIASKTQSICKYSTDILGNLYE
VSKKKHPQIITKG

```

It is appreciated that when the term “dCas9” or “nuclease-inactive Cas9” is used herein, it refers to Cas9 variants that are inactive in both HNH and RuvC domains as well as Cas9 nickases. For example, the dCas9 used in the present disclosure may include the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 3. In some embodiments, the dCas9 may comprise other mutations that inactivate RuvC or HNH domain. Additional suitable mutations that inactivate Cas9 will be apparent to those of skill in the art based on this disclosure and knowledge in the field, and are within the scope of this disclosure. Such additional exemplary suitable nuclease-inactive Cas9 domains include, but are not limited to, D839A and/or N863A (See, e.g., Prashant et al., *Nature Biotechnology*. 2013; 31(9): 833-838, the entire contents of which are incorporated herein by reference), or K603R (See, e.g., Chavez et al., *Nature Methods* 12, 326-328, 2015, the entire contents of which is incorporated herein by reference). The term Cas9, dCas9, or Cas9 variant also encompasses Cas9, dCas9, or Cas9 variant from any organism. Also appreciated is that dCas9, Cas9 nickase, or other appropriate Cas9 variants from any organisms may be used in accordance with the present disclosure.

A “deaminase” refers to an enzyme that catalyzes the removal of an amine group from a molecule, or deamination. In some embodiments, the deaminase is a cytidine deaminase, catalyzing the hydrolytic deamination of cytidine or deoxycytidine to uridine or deoxyuridine, respectively. In some embodiments, the deaminase is a cytosine deaminase, catalyzing the hydrolytic deamination of cytosine to uracil (e.g., in RNA) or thymine (e.g., in DNA). In some embodiments, the deaminase is a naturally-occurring deaminase from an organism, such as a human, chimpanzee, gorilla, monkey, cow, dog, rat, or mouse. In some embodiments, the deaminase is a variant of a naturally-occurring deaminase from an organism, that does not occur in nature. For example, in some embodiments, the deaminase or deaminase domain is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75% at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring deaminase from an organism.

A “cytosine deaminase” refers to an enzyme that catalyzes the chemical reaction “cytosine+H₂O→uracil+NH₃” or “5-methyl-cytosine+H₂O→thymine+NH₃.” As it may be apparent from the reaction formula, such chemical reactions result in a C to U/T nucleobase change. In the context of a gene, such nucleotide change, or mutation, may in turn lead to an amino acid residue change in the protein, which may affect the protein function, e.g., loss-of-function or gain-of-function.

One exemplary suitable class of cytosine deaminases is the apolipoprotein B mRNA-editing complex (APOBEC) family of cytosine deaminases encompassing eleven proteins that serve to initiate mutagenesis in a controlled and beneficial manner. The apolipoprotein B editing complex 3 (APOBEC3) enzyme provides protection to human cells

against a certain HIV-1 strain via the deamination of cytosines in reverse-transcribed viral ssDNA. These cytosine deaminases all require a Zn²⁺-coordinating motif (His-X-Glu-X₂₃₋₂₆-Pro-Cys-X₂₄-Cys; SEQ ID NO: 820) and bound water molecule for catalytic activity. The Glu residue acts to activate the water molecule to a zinc hydroxide for nucleophilic attack in the deamination reaction. Each family member preferentially deaminates at its own particular “hotspot”, ranging from WRC (W is A or T, R is A or G) for hAID, to TTC for hAPOBEC3F. A recent crystal structure of the catalytic domain of APOBEC3G revealed a secondary structure comprised of a five-stranded 3-sheet core flanked by six α-helices, which is believed to be conserved across the entire family. The active center loops have been shown to be responsible for both ssDNA binding and in determining “hotspot” identity. Overexpression of these enzymes has been linked to genomic instability and cancer, thus highlighting the importance of sequence-specific targeting. Another suitable cytosine deaminase is the activation-induced cytidine deaminase (AID), which is responsible for the maturation of antibodies by converting cytosines in ssDNA to uracils in a transcription-dependent, strand-biased fashion.

Herein, a “nucleobase editor” refers to a protein that edits a nucleotide base. “Edit” refers to the conversion of one nucleotide base to another. For example, the nucleobase may target C bases in a nucleic acid sequence and convert the C to T base. In some embodiments, the C to T editing is carried out by a deaminase, e.g., a cytosine deaminase. Other types of base conversions are also contemplated. In some embodiments, the nucleobase editor comprises a DNA binding domain that directs it to a target sequence.

As such, a base editor may be a cytosine deaminase-dCas9 fusion protein. In some embodiments, the base editor may be a deaminase-dCas9-UGI fusion protein. In some embodiments, the base editor may be a APOBEC1-dCas9-UGI fusion protein. In some embodiments, the base editor may be APOBEC1-Cas9 nickase-UGI fusion protein. In some embodiments, the base editor may be APOBEC1-dCpf1-UGI fusion protein. In some embodiments, the base editor may be APOBEC1-dNgAgo-UGI fusion protein. In some embodiments, the base editor may be a pmCDA1-Cas9 nickase-UGI fusion protein. In some embodiments, the base editor may be a human APOBEC3G-Cas9 nickase UGI fusion protein. In some embodiments, the base editor may comprise a second UGI domain. Non-limiting exemplary sequences of the nucleobase editors described herein are provided in Example 1, SEQ ID NOs: 293-302, 1071, and 1084. Such nucleobase editors and methods of using them for genome editing have been described in the art, e.g., in U.S. Pat. No. 9,068,179, US Patent Application Publications US 2015/0166980, US 2015/0166981, US 2015/0166982, US20150166984, and US20150165054, and US Provisional Applications, U.S. Ser. No. 62/245,828, filed Oct. 23, 2015; 62/279,346, filed Jan. 15, 2016; 62/311,763, filed Mar. 22, 2016; 62/322,178, filed Apr. 13, 2016, 62/357,352, filed Jun.

30, 2016, U.S. Pat. No. 62,370,700, filed Aug. 3, 2016; 62/398,490, filed Sep. 22, 2016; 62/408,686, filed Oct. 14, 2016; PCT Application PCT/US2016/058344, filed Oct. 22, 2016, U.S. patent application Ser. No. 15/311,852, filed Oct. 22, 2016; Komor et al. (2017) Improved Base Excision Repair Inhibition and Bacteriophage Mu Gam Protein Yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv.*, 3: eaao4774; and in Komor et al., *Nature*, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage, 533, 420-424 (2016), the entire contents of each of which are incorporated herein by reference.

The term “target site” or “target sequence” refers to a sequence within a nucleic acid molecule (e.g., a DNA molecule) that is deaminated by the fusion protein provided herein. In some embodiments, the target sequence is a polynucleotide (e.g., a DNA), wherein the polynucleotide comprises a coding strand and a complementary strand. The meaning of a “coding strand” and “complementary strand” is the common meaning of the terms in the art. In some embodiments, the target sequence is a sequence in the genome of a mammal. In some embodiments, the target sequence is a sequence in the genome of a human. The term “target codon” refers to the amino acid codon that is edited by the base editor and converted to a different codon via deamination of C base. In some embodiments, the target codon is edited in the coding strand. In some embodiments, the target codon is edited in the complementary strand.

The term “linker,” as used herein, refers to a chemical group or a molecule linking two molecules or moieties, e.g., two domains of a fusion protein, such as, for example, a nuclease-inactive Cas9 domain and a nucleic acid editing domain (e.g., a deaminase domain). Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (e.g., a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 2-100 amino acids in length, for example, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

The term “mutation,” as used herein, refers to a substitution of a residue within a sequence, e.g., a nucleic acid or amino acid sequence, with another residue, or a deletion or insertion of one or more residues within a sequence. Mutations are typically described herein by identifying the original residue followed by the position of the residue within the sequence and by the identity of the newly substituted residue. Various methods for making the amino acid substitutions (mutations) provided herein are well known in the art, and are provided by, for example, Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)).

The terms “nucleic acid” and “nucleic acid molecule,” as used herein, refer to a compound comprising a nucleobase and an acidic moiety, e.g., a nucleoside, a nucleotide, or a polymer of nucleotides. Typically, polymeric nucleic acids, e.g., nucleic acid molecules comprising three or more nucleotides are linear molecules, in which adjacent nucleotides are linked to each other via a phosphodiester linkage. In some embodiments, “nucleic acid” refers to individual nucleic acid residues (e.g. nucleotides and/or nucleosides).

In some embodiments, “nucleic acid” refers to an oligonucleotide chain comprising three or more individual nucleotide residues. As used herein, the terms “oligonucleotide” and “polynucleotide” can be used interchangeably to refer to a polymer of nucleotides (e.g., a string of at least three nucleotides). In some embodiments, “nucleic acid” encompasses RNA as well as single and/or double-stranded DNA. Nucleic acids may be naturally occurring, for example, in the context of a genome, a transcript, an mRNA, tRNA, rRNA, siRNA, snRNA, a plasmid, cosmid, chromosome, chromatid, or other naturally occurring nucleic acid molecule. On the other hand, a nucleic acid molecule may be a non-naturally occurring molecule, e.g., a recombinant DNA or RNA, an artificial chromosome, an engineered genome, or fragment thereof, or a synthetic DNA, RNA, DNA/RNA hybrid, or including non-naturally occurring nucleotides or nucleosides. Furthermore, the terms “nucleic acid,” “DNA,” “RNA,” and/or similar terms include nucleic acid analogs, e.g., analogs having other than a phosphodiester backbone. Nucleic acids can be purified from natural sources, produced using recombinant expression systems and optionally purified, chemically synthesized, etc. Where appropriate, e.g., in the case of chemically synthesized molecules, nucleic acids can comprise nucleoside analogs such as analogs having chemically modified bases or sugars, and backbone modifications. A nucleic acid sequence is presented in the 5' to 3' direction unless otherwise indicated. In some embodiments, a nucleic acid is or comprises natural nucleosides (e.g. adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine); nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, 0(6)-methylguanine, and 2-thiocytidine); chemically modified bases; biologically modified bases (e.g., methylated bases); intercalated bases; modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose); and/or modified phosphate groups (e.g., phosphorothioates and 5'-N-phosphoramidite linkages).

The terms “protein,” “peptide,” and “polypeptide” are used interchangeably herein, and refer to a polymer of amino acid residues linked together by peptide (amide) bonds. The terms refer to a protein, peptide, or polypeptide of any size, structure, or function. Typically, a protein, peptide, or polypeptide will be at least three amino acids long. A protein, peptide, or polypeptide may refer to an individual protein or a collection of proteins. One or more of the amino acids in a protein, peptide, or polypeptide may be modified, for example, by the addition of a chemical entity such as a carbohydrate group, a hydroxyl group, a phosphate group, a farnesyl group, an isofarnesyl group, a fatty acid group, a linker for conjugation, functionalization, or other modification, etc. A protein, peptide, or polypeptide may also be a single molecule or may be a multi-molecular complex. A protein, peptide, or polypeptide may be just a fragment of a naturally occurring protein or peptide. A protein, peptide, or polypeptide may be naturally occurring, recombinant, or synthetic, or any combination thereof. The term “fusion protein” as used herein refers to a hybrid polypeptide which comprises protein domains from at least two different proteins. One protein may be located at the amino-terminal (N-terminal) portion of the fusion protein or at the carboxy-terminal (C-terminal) protein thus forming an “amino-ter-

minal fusion protein" or a "carboxy-terminal fusion protein," respectively. A protein may comprise different domains, for example, a nucleic acid binding domain (e.g., the gRNA binding domain of Cas9 that directs the binding of the protein to a target site) and a nucleic acid cleavage domain or a catalytic domain of a nucleic-acid editing protein. In some embodiments, a protein comprises a proteinaceous part, e.g., an amino acid sequence constituting a nucleic acid binding domain, and an organic compound, e.g., a compound that can act as a nucleic acid cleavage agent. In some embodiments, a protein is in a complex with, or is in association with, a nucleic acid, e.g., RNA. Any of the proteins provided herein may be produced by any method known in the art. For example, the proteins provided herein may be produced via recombinant protein expression and purification, which is especially suited for fusion proteins comprising a peptide linker. Methods for recombinant protein expression and purification are well known, and include those described by Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), the entire contents of which are incorporated herein by reference.

The term "subject," as used herein, refers to an individual organism, for example, an individual mammal. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal. In some embodiments, the subject is a non-human primate. In some embodiments, the subject is a rodent. In some embodiments, the subject is a sheep, a goat, a cattle, a cat, or a dog. In some embodiments, the subject is a vertebrate, an amphibian, a reptile, a fish, an insect, a fly, or a nematode. In some embodiments, the subject is a research animal. In some embodiments, the subject is genetically engineered, e.g., a genetically engineered non-human subject. The subject may be of either sex and at any stage of development.

The term "recombinant" as used herein in the context of proteins or nucleic acids refers to proteins or nucleic acids that do not occur in nature, but are the product of human engineering. For example, in some embodiments, a recombinant protein or nucleic acid molecule comprises an amino acid or nucleotide sequence that comprises at least one, at least two, at least three, at least four, at least five, at least six, or at least seven mutations as compared to any naturally occurring sequence.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS OF THE INVENTION

The immune system is critical in preventing the outgrowth of cancers, and "immunosurveillance" exists to provide immunological resistance against cancer development (e.g., as described in Old et al., *Annu Rev Med* 1964; 15: 167-186; Burnet et al., *Prog Exp Tumor Res* 1970; 13: 1-27; and Graziano et al., *Cancer Treat Res* 2005; 123: 89-111, each of which is incorporated herein by reference).

Despite the presence of immunosurveillance, cancers can develop in apparently immunocompetent animals and humans, due to the ability of cancer cells to evade immunosurveillance (e.g., as described in Hanahan et al., *Cell* 2000; 100: 57-70 and Zitvogel et al., *Nat Rev Immunol* 2006; 6: 715-727, each of which is incorporated herein by reference). The evasion of cancer cells from immunosurveillance occurs via various well-characterized mechanisms, including induction of T-cell tolerance by autochthonous tumors (e.g., as described in Willimsky et al., *Nature* 2005; 437: 141-146, incorporated herein by reference), cancer immu-

noediting (e.g., as described in Dunn et al., *Nat Immunol* 2002; 3: 991-998, incorporated herein by reference), and development of an immune suppressive cancer microenvironment (e.g., as described in Zou et al., *Nat Rev Cancer* 2005; 5: 263-274, incorporated herein by reference). Therapeutic cancer vaccines or adoptive immunotherapy are being developed and tested as potential approaches to strengthen the immune responses after tumor arise in order to slow their progression and prevent their recurrence. Immunotherapeutic approaches, e.g., cancer vaccines have been described but are only partially successful (e.g., as described in Finn et al., *Nat Rev Immunol* 2003; 3: 630-641, incorporated herein by reference).

Described herein are systems, methods, compositions, and kits for producing immunogenic peptides derived from tumor specific antigens (e.g., heteroclitic epitopes or cryptic epitopes) that may be used as cancer vaccines in vivo or ex vivo. Targeted mutations are introduced into tumor-specific antigens using a gene editing agent, e.g., a nucleobase editor comprising a programmable DNA binding domain (e.g., a catalytically-inactive Cas9 or Cas9 nuclease) fused to a cytosine deaminase, to generate altered versions of peptides arising from the tumor-specific antigens (heteroclitic epitopes) or peptides arising from normally untranslated regions of the tumor-specific antigen genes (cryptic peptides). The heteroclitic peptides or cryptic peptides may be generated in vivo in a subject (e.g., a subject who has cancer) and presented to the adaptive immune system via the MHC class I or MHC class II pathway, which in turn induces a strong adaptive immune response, e.g., T cell response and B cell response. Such an adaptive immune response is antigen specific and is effective in reducing tumor growth and preventing metastasis.

The advantage of the cancer vaccines of the present disclosure is that the vaccine (e.g., antigenic peptides derived from tumor-specific antigens) is generated from the genome and the proteome of the malignant cells in vivo and is highly personalized. The cancer vaccines described herein are also highly cancer-specific and do not induce unwanted immune response against "self," since the immunogenic epitopes are derived from tumor-specific antigens. Further, the adaptive immune response induced by the cancer vaccine described herein confer "memory" to the immune system, promoting the immune system to efficiently recognize "neoepitopes" generated due to the highly mutagenic nature of the cancer genome, thus preventing metastasis and facilitate remission. To enhance the efficacy of the cancer vaccines described herein, combination therapies using an immune checkpoint inhibitor in conjunction with the cancer vaccine is also contemplated, aiming to enhance the tumor antigen specific immune response.

The methods of producing endogenous cancer vaccines in vivo are enabled by the targeted nucleobase editing technology described herein. Such base editing technology is described in the art, e.g., in U.S. Pat. No. 9,068,179, US Patent Application Publications US 2015/0166980, US 2015/0166981, US 2015/0166982, US20150166984, and US20150165054, and US Provisional Applications, U.S. Ser. No. 62/245,828, filed Oct. 23, 2015; 62/279,346, filed Jan. 15, 2016; 62/311,763, filed Mar. 22, 2016; 62/322,178, filed Apr. 13, 2016, 62/357,352, filed Jun. 30, 2016, U.S. Pat. No. 62,370,700, filed Aug. 3, 2016; 62/398,490, filed Sep. 22, 2016; 62/408,686, filed Oct. 14, 2016; PCT Application PCT/US2016/058344, filed Oct. 22, 2016, U.S. patent application Ser. No. 15/311,852, filed Oct. 22, 2016; and in Komor et al., *Nature*, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage,

533, 420-424 (2016), the entire contents of each of which are incorporated herein by reference.

Immunogenic Peptides or Epitopes

Some aspects of the present disclosure provide immunogenic peptides or epitopes derived from tumor-specific antigens and how these peptides or epitopes elicit tumor-specific immune response. A large number of proteins that specifically express in tumor cells or are upregulated in tumor cells have been identified (Hassane et al., Holland-Frei Cancer Medicine, 6th edition). The known tumor specific antigens are classified into different classes.

a) Cancer-testis antigens: The first TAAs ever identified that can be recognized by T cells belong to this class, which was originally called cancer-testis (CT) antigens because of the expression of its members in histologically different human tumors and, among normal tissues, only in spermatocytes/spermatogonia of testis and, occasionally, in placenta. Since the cells of testis do not express class I and II HLA molecules, these antigens cannot be recognized by T cells in normal tissues and can therefore be considered as immunologically tumor-specific. Well-known examples of CT antigens are the MAGE family members or NY-ESO-1.

b) Differentiation antigens: These TAAs are shared between tumors and normal tissue from which the tumor arose; most are found in melanomas and normal melanocytes. Many of these melanocyte lineage-related proteins are involved in the biosynthesis of melanin and are therefore not tumor specific but nevertheless are widely used for cancer immunotherapy. Examples include, but are not limited to, tyrosinase and Melan-A/MART-1 for melanoma, and PSA for prostate cancer.

c) Overexpressed TAAs: Genes encoding widely expressed TSAs have been detected in histologically different types of tumors as well as in many normal tissues, generally with lower expression levels. It is possible that many of the epitopes processed and potentially presented by normal tissues are below the threshold level for T-cell recognition, while their overexpression in tumor cells can trigger an anticancer response by breaking previously established tolerance. Examples of this class of TAAs are Her-2/neu, Survivin, Telomerase and WT1.

d) Tumor specific antigens: These unique TAAs arise from mutations of normal genes (such as β -catenin, CDK4, etc.). Some of these molecular changes are associated with neoplastic transformation and/or progression. Tumor specific antigens are generally able to induce strong immune responses without bearing the risk for autoimmune reactions against normal tissues. On the other hand, these TAAs are in most cases only relevant to the exact tumor on which they were identified and are usually not shared between many individual tumors.

e) TAAs arising from abnormal post-translational modifications: Such TSAs may arise from proteins which are neither specific nor overexpressed in tumors but nevertheless become tumor associated by posttranslational processes primarily active in tumors. Examples for this class arise from altered glycosylation patterns leading to novel epitopes in tumors (e.g., MUC1).

f) Oncoviral proteins: These TSAs are viral proteins that may play a critical role in the oncogenic process, and because they are foreign (not of human origin), they can evoke a T-cell response. Examples of such proteins are the human papilloma type 16 virus proteins, E6 and E7, which are expressed in cervical carcinoma.

TAAs are a starting point for the development of a tumor vaccine. The methods for identifying and characterizing the TAAs are based on the use of cytotoxic T lymphocytes (CTL)

that can be isolated from patients or healthy subjects, or they are based on the generation of differential transcription profiles or differential peptide expression patterns between tumors and normal tissues.

In some embodiments, the tumor-specific antigen is expressed in a broad range of different types of cancers. In some embodiments, the tumor-specific antigen is expressed only in one or a few types of cancers. The anti-cancer immune response described herein is antigen-specific. As such, an immune response induced by a tumor-specific antigen is specific to cancer types where the said antigen is expressed. Non-limiting, exemplary tumor-specific antigens that may be edited to generate immunogenic epitopes are provided in Tables 1-3. It is appreciated that the examples are for illustration purpose only, and the methods described herein may be applied to any tumor-specific antigen.

In some embodiments, the immunogenic peptide or epitope is a portion of the tumor-specific antigen. For example, the immunogenic peptide or epitope may be a portion of the tumor-specific antigen that is 5-40 amino acids long. In some embodiments, the immunogenic peptide or epitope is 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 amino acids long. In some embodiments, the immunogenic peptide or epitope comprises modifications, e.g., amino acid substitutions (also termed "heteroclitic epitopes"), as compared to the native sequence in the tumor specific antigen. In some embodiments, the immunogenic peptide or epitope comprises more than one amino acid substitutions (e.g., 2, 3, 4, 5, or more) compared to the native sequence of the tumor-specific antigen it is derived from. In some embodiments, a heteroclitic peptide or epitope may be at least 60%, at least 70%, at least 80%, at least 90%, at least 98%, or at least 99% identical to the native sequence that it is derived from. In some embodiments, a heteroclitic peptide or epitope is 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identical to the native sequence that it is derived from.

In some embodiments, a heteroclitic peptide or epitope is more immunogenic than a peptide of its native sequence. For example, a heteroclitic epitope may be at least 30% more immunogenic (i.e., induces a stronger immune response) than its corresponding native peptide. In some embodiments, a heteroclitic epitope may be at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 9-fold, at least 10-fold, at least 20-fold, at least 30-fold, at least 40-fold, at least 50-fold, at least 60-fold, at least 70-fold, at least 80-fold, at least 90-fold, at least 100-fold, or more immunogenic than its corresponding native peptide.

In some embodiments, the immunogenic peptide or epitope is a cryptic peptide or epitope, e.g., generated from translation of a non-coding region of the tumor specific antigen gene or translation of a different reading frame of a coding region of the tumor specific antigen. A cryptic peptide or epitope may be more immunogenic (i.e., induces a stronger immune response) than any native peptide derived from the tumor associated antigen. For example, a cryptic peptide or epitope may be at least 30% more immunogenic than any native peptide derived from the tumor associated antigen. In some embodiments, a cryptic peptide or epitope may be at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at

least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 9-fold, at least 10-fold, at least 20-fold, at least 30-fold, at least 40-fold, at least 50-fold, at least 60-fold, at least 70-fold, at least 80-fold, at least 90-fold, at least 100-fold, or more immunogenic than any native peptide derived from the tumor associated antigen. One skilled in the art is familiar

with how to assess the immune response induced by an antigen, e.g., measuring antibody titers.

Tumor specific antigens from which antigenic epitopes (e.g., heteroclitic epitopes and cryptic epitopes) may be derived are provided in Tables 1-3. (HLA—human leukocyte antigen type)

TABLE 1

Tumor associated antigen - Differentiation						
Gene/ protein	Tumor	HLA	HLA frequency (%)	Peptide	SEQ ID NO	Position
CEA	gut carcinoma	A2	44	YLSGANLNL	340	605-613
		A2	44	IMIGVLVGV	341	691-699
		A2	44	GVLVGVALI	342	694-702
		A2	44	LLTFWNPPT	343	24-33
		A3	22	HLFGYSWYK	344	61-69
		A24	20	QYSWFVNNGTF	345	268-277
		A24	20	TYACFVSNL	346	652-660
		DR3	21	AYVCGIQNSVSANRS	347	568-582
		DR4	24	DTGFYTLHVIKSSDLVNNEA	348	116-140
				TCQFRV		
		DR4	24	YSWRINGIPQQHTQV	349	625-639
		DR7	25	TYYRPGVNLSLSC	350	425-437
		DR7	25	EIIYPNALSNI	351	99-111
gp100/ Pmel17	melanoma	DR9	3	YACFVSNLATGRNNNS	352	653-667
		DR11	25	LWWVNNQSLPVSP	353	177-189, 355-367
		DR13	19	LWWVNNQSLPVSP	354	177-189, 355-367
		DR14	6	LWWVNNQSLPVSP	355	177-189, 355-367
		DR14	6	EIIYPNASLLIQN	356	99-111
		DR14	6	NSIVKSIITVSAAG	357	666-678
		A2	44	KTWGQYWQV	358	154-162
		A2	44	(A) MLGTHTM	359	177(8)-186
		A2	44	ITDQVPFSV	360	209-217
		A2	44	YLEPGPVTA	361	280-288
		A2	44	LLDGTTATLRL	362	457-466
		A2	44	VLYRYGSFSV	363	476-485
		A2	44	SLADTNSLAV	364	570-579
		A2	44	RLMKQDFSV	365	619-627
mammagl obin-A	breast cancer	A2	44	RLPRIFCSC	366	639-647
		A3	22	LIYRRRLMK	367	614-622
		A3	22	ALLAVGATK	368	17-25
		A3	22	IALNFPGSQK	369	86-95
		A3	22	RSYVPLAHR	370	195-202
						and 191 or 192
		A3	22	ALNFPGSQK	371	87-95
		A11	13	ALNFPGSQK	372	87-95
		A24	20	VYFFLPDHL	373	intron 4
		A32	8	RTKQLYPEW	374	40-42 and 47-52
		A68	8	HTMEVTVYHR	375	182-191
		B7	17	SSPGCQPPA	376	529-537
		B35	20	VPLDCVLYRY	377	471-480
Melan-A/ MART-1	melanoma	B35	20	LPHSSSHWL	378	630-638
		Cw8	—	SNDGPTLII	379	71-78
		DQ6	63	GRAMLGHTHMEVTY	380	175-189
		DR4	24	WNRQLYPEWTEAQRLD	381	44-59
		DR7	25	TTEWVETTARELPIPEPE	382	420-437
		DR7	25	TGRAMLGHTHMEVTYH	383	174-190
		DR53	49	GRAMLGHTHMEVTY	384	175-189
		A3	22	PLLENVISK	385	23-31
		A2	44	(E) AAGIGILTV	386	26(27)-35
		A2	44	ILTVILGV	387	32-40
		B35	20	EAAGIGILTV	388	26-35
		B45	2	AAEAAGIGIL(T)	389	24-33 (34)
		Cw7	41	RNGYRALMDKS	390	51-61
		DP5	3	YTTAEEAAGIGILTVILGV	391	21-50
		DQ6	63	LLLIGCWYCRR	392	25-36
				EEAAIGILTVI		

TABLE 1-continued

Tumor associated antigen - Differentiation						
Gene/ protein	Tumor	HLA	HLA frequency (%)	Peptide	SEQ ID NO	Position
NY-BR-1	breast cancer	A2	DR1	AAGIGILTVILGVL	393	27-40
			DR1	APPAYEKLpSABQf	394	100-111
			DR3	EEAAGIGILTVI	395	25-36
			DR4	RNGYRALMDKSLHVGTQC	396	51-73
			ALTRR			
			DR11	MPREDAHFIYGYPKKGHGH S	397	1-20
ERBB2	breast cancer	A2	DR52	KNCEPVPNAPPAYEKLSE	398	91-110
OA1	melanoma	A24	20	LYSACFWWL	401	126-134
PAP	prostate cancer	A2	44	FLFLLFFWL	402	18-26
			44	TLMSAMTNL	403	112-120
			44	ALDVYNGLL	404	299-307
PSA	prostate carcinoma	A2	44	FLTPKKLQCV	405	165-174
			44	VISNDVCAQV	406	178-187
RAB38/ NY-MEL-1	melanoma	A2	44	VLHWDPETV	407	50-58
TRP-1/ 9B75	melanoma	A31	5	MSLQRQFLR	408	alt. ORF
			24	ISPNSVFSQWRVVCDLED YD	409	277-297
		DR15	20	SLPYWNFATG	410	245-254
			21	SQWRVVCDSLEDYDT	411	284-298
TRP-2	melanoma	A2	44	SVYDFVVWL	412	180-188
		A2	44	TLDSQVMSL	413	360-368
		A31	5	LLGPGRPYR	414	197-205
		A33	5	LLGPGRPYR	415	197-205
		Cw8	—	ANDPIFVVL	416	387-395
		DR3	21	QCTEVRADTRPWSGP	417	60-74
		DR15	20	ALPYWNFATG	418	241-250
tyrosinase	melanoma	A1	26	KCDICTDEY	419	243-251
		A1	26	SSDYVIPIGTY	420	146-156
		A2	44	MLLAVLYCL	421	1-9
		A2	44	CLLWSFQTSA	422	8-17
		A2	44	YMDGTMQSV	423	369-377
		A24	20	AFLPWHRLF	424	206-214
		A24	20	IYMDGTADFSF	425	368-373
						and
		A26	8	QCSGNFMGF	426	90-98
		B35	20	TPRLPSSADVEF	427	309-320
		B35	20	LPSSADVEF	428	312-320
		B38	5	LHHAFVDEIF	429	388-397
		B44	21	SEIWRDIDFD	430	192-200
		DR4	24	QNILLSNAPLGPOFP	431	56-70
		DR4	24	SYLQDSDPDSFQD	432	450-462
		DR15	20	FLLHHAFVDSIFEQWLQRH RP	433	386-406
WT1	testis, ovary, bone marrow, spleen	A1 A24 DP5 DP5 DR4	26 20 3 3 24	TSEKRPFMCA CMTWNQMNL LSHLQMHSRKH KRYFKLQLSHLQMHSRKH KRYFKLQLSHLQMHSRKH	434 435 436 437 438	317-327 235-243 337-347 332-347 332-347
CD33	leukemia	A2	44	AIISGDSPV	439	65-73

TABLE 2

Tumor specific antigen - Tumor specific					
Gene/protein	HLA	HLA frequency (%)	Peptide	SEQ ID NO	Position
BAGE-1	Cw16	7	AARAVFLAL	440	1-10
D393-CD20n	DR4	24	KPLFRRMSSL <u>ELVIA</u>	441	28-42
Cyclin-A1	A2	44	FLDRFLSCM	442	227-235
	A2	44	SLIAAAAFCLA	443	341-351
GAGE-1, 2, 8	Cw6	18	YRPRPRRY	444	9-16
GAGE-3, 4, 5, 6, 7	A29	6	YYWPRPRRY	445	10-18
GnTVf	A2	44	VLPDVFIRC(V)	446	intron
HERV-K-MEL	A2	44	MLAVISCAV	447	1-9
KK-LC-1	B15	13	RQKRILVNL	448	76-84
KM-HN-1	A24	20	NYNNFYRFL	449	196-204
	A24	20	EYSKECLKEF	450	499-508
	A24	20	EYLSLSDKI	451	770-778
LAGE-1	A2	44	MLMAQEALAF	452	alt. ORF (1-11)
	A2	44	SLLMWITQC	453	157-165
	A31	5	LAAQERRVPR	454	alt. ORF (18-27)
	A68	8	ELVRRILSR	455	103-111
	B7	17	APRGVRMVA	456	alt. ORF (46-54)
	DP4	75	SLLMWITQCFLPVF	457	157-170
	DR3	21	QGAMLAQERRQRPRAAEVPR	458	alt. ORF (14-33)
	DR4	24	AADHRQLQLSISSCLQQL	459	139-156
	DR11	25	CLSRRPWKRWSAGSCPMPHL	460	alt. ORFT (81-102)
	DR12	5	CLSRRPWKRWSAGSCPMPHL	461	alt. ORFT (81-102)
	DR13	19	ILSRDAAPLPRPG	462	108-120
	DR15	20	AGATGGRGPRGAGA	463	37-50
LY6K	A24	20	RYCNLEGPP	464	119-128
	DP5	3	KWTEPYCVIAAVKIFPRFFMVAKQ	465	61-84
	DR15	20	KCKKIRYCNLLEGPPINSSVF	466	114-133
MAGE-A1	A1	26	EADPTGHSY	467	161-169
	A2	44	KVLEYVIKV	468	278-286
	A3	22	SLFRAVITK	469	96-104
	A68	8	EVDGREGHSA	470	222-231
	B7	17	RVRFFFPSL	471	289-298
	B35	20	EADPTGHSY	472	161-169
	B37	3	REPVTKAEML	473	120-129
	B44	21	KEADPTGHSY	474	160-169
	B53	2	DPARYEFLW	475	258-266
	B57	8	ITKKVADLVGF	476	102-112
	Cw2	10	SAPPTTINF	477	62-70
	Cw3	17	SAYGEPRKL	478	230-238
	Cw7	41	RVRFFFPSL	479	289-298
	Cw16	7	SAYGEPRKL	480	230-238
	DP4	75	TSCILESLFRAVITK	481	90-104
	DP4	75	PRALAETSYVKLEY	482	268-282
	DR13	19	FLLLKYRAREPVTKAE	483	112-127
	DR15	20	EYVIKVSARVRF	484	281-292
MAGE-A2	A2	44	YLQLVFGIEV	485	157-166
	A24	20	EYQLVFGI	486	156-164
	B37	3	REPVTKAEML	487	127-136
	Cw7	41	EGDCAPEEK	488	212-220
	DR13	19	LLKYRAREPVTKAE	489	121-134
MAGE-A3	A1	26	EVDPIGHLY	490	168-176
	A2	44	FLWGPRALVd	491	271-279
	A2	44	KVAELVHFL	492	112-120
	A24	20	TFPDLESEF	493	97-105
	A24	20	VAELVHFLL	494	113-121

TABLE 2-continued

Tumor specific antigen - Tumor specific					
Gene/protein	HLA	HLA frequency (%)	Peptide	SEQ ID NO	Position
	B18	6	MEVDPIGHLY	495	167-176
	B35	20	EVDPIGHLY	496	168-176
	B37	3	REPVTKAEML	497	127-136
	B40	6	AELVHFLLL	498	114-122
	B44	21	MEVDPIGHLY	499	167-176
	B52	5	WQYFFPVIF	500	143-151
	Cw7	41	EGDCAPEEK	501	212-220
	DP4	75	KKLLTQHFPQENYLEY	502	243-258
	DP4	75	RKVAELVHFLLLKYR	503	111-125
	DQ6	63	KKLLTQHFPQENYLEY	504	243-258
	DR1	18	ACYEFLWGPRLVETS	505	267-282
	DR4	24	RKVAELVHFLLLKYR	506	111-125
	DR4	24	VIFSKASSSQL	507	149-160
	DR7	25	VIRSKASSSQL	508	149-160
	DR7	25	VPGIELMEVDPIGHL	509	161-175
	DR11	25	GDNQIMPKAGLLIIV	510	191-205
	DR11	25	TSYVKVLHHMVKISG	511	281-295
	DR13	19	RKVAELVHFLLLKYRA	512	111-126
	DR13	19	FLLLKYRAREPVTKAE	513	119-134
MAGE-A4	A1	26	EVDPASNTY	514	169-177
	A2	44	GVYDGREHTV	515	230-239
	A24	20	NYKRCFPVI	516	143-151
	B37	3	SESLKMIF	517	156-163
MAGE-A6	A34	1	MVKISGGPR	518	290-298
	B35	20	EVDPIGHVY	519	168-176
	B37	3	REPVTKAEML	520	127-136
	Cw7	41	EGDCAPEEK	521	212-220
	Cw16	7	ISGGPRISY	522	293-301
	DR13	19	LLKYRAREPVTKAE	523	121-134
MAGE-A9	A2	44	ALSVVMGVYV	524	223-231
MAGE-A10	A2	44	GLYDGMEHLI	525	254-262
	B53	2	DPARYEFLW	526	290-298
MAGE-A12 m	A2g	44	FLWGPRALV	527	271-279
	Cw7	41	VRIGHLYIL	528	170-178
	Cw7	41	EGDCAPEEK	529	212-220
	DP4	75	REPFTKAEMLGSVR	530	127-141
	DR13	19	AELVHFLLLKYRAR	531	114-127
MAGE-C1	A2	44	ILFGISLREV	532	959-968
	A2	44	KVVEFLAML	533	1083-1091
	DQ6	63	SSALLSIFQSSPE	534	137-149
	DQ6	63	SFSYTLSSL	535	450-458
	DR15	20	VSSFFFSYTL	536	779-787
MAGE-C2	A2	44	LLFGLALIEV	537	191-200
	A2	44	ALKDVEERV	538	336-344
	B44	21	SESIKKKVL	539	307-315
	B57	8	ASSTLYLVF	540	42-50
	DR15	20	SSTLYLVFSPSSFST	541	43-57
mucink			PDTTRPAPGSTAPPAHGVTSAA	542	
NA88-A	B13	6	QQQHFLQKV	543	
	A2	44	SLLMWITQC	544	157-165
NY-ESO-1/ LAGE-2	A2	44	MLMMAQEALAFL	545	alt. ORF (1-11)
	A24	20	YLAMPFATPME	546	91-101
	A31	5	ASGPGGGAPR	547	53-62
	A31	5	LAAQERRVPR	548	alt. ORF (18-27)
	A68	8	TVSGNILTIR	549	127-136
	B7	17	APRGPHGGAASGL	550	60-72
	B35	20	MPFPATPMEEAL	551	94-104
	B49		KEFTVSGNILTII	552	124-135
	B51	12	MPFPATPMEEA	553	94-102
	B52	5	FATPMEEAL	554	96-104
	C12	12	FATPMEEALAR	555	96-106
	Cw3	17	LAMPFATPM	556	92-100

TABLE 2-continued

Tumor specific antigen - Tumor specific					
Gene/protein	HLA	HLA frequency (%)	Peptide	SEQ	
				ID NO	Position
Cw6	18	ARGPESRLL		557	80-88
DP4	75	SLLMWITQCFLPVF		558	157-170
DP4	75	LLEFYLAMPFATPMEEAELARRSLAQ		559	87-111
DR1	18	LLEFYLAMPFATPMEEAELARRSLAQ		560	87-111
DR1	18	EFYLYLAMPFATPM		561	89-100
DR1	18	PGVLLKEFTVSGNILTIRLTAADHR		562	119-143
DR2	25	RLLFYLAMPFA		563	86-97
DR3	21	QGAMLAQERRQERPRRAAEVPR		564	alt. ORF (14-33)
DR4	24	PFATPMEEAELARR		565	95-107
DR4	24	PGVLLKEFTVSGNILTIRLT		566	119-138
DR4	24	VLLKEFTVSG		567	121-130
DR4	24	AADHRQLQLSISSCLQQL		568	139-156
DR4	24	LLEFYLAMPFATPMEEAELARRSLAQ		569	87-111
DR52b	25	LKEFTVSGNILTIRL		570	123-137
DR7	25	PGVLLKEFTVSGNILTIRLTAADHR		571	119-143
DR7	25	LLEFYLAMPFATPMEEAELARRSLAQ		572	87-111
DR8	4	KEFTVSGNILT		573	124-134
DR9	3	LLEFYLAMPFATPM		574	87-100
DR15	20	AGATGGRGPRGAGA		575	37-50
SAGE	A24	20	LYATVIHDI	576	715-723
Sp17	A1	26	ILDSSSEDK	577	103-111
SSX-2	A2	44	KASEKIFYV	578	41-49
	DP1	14	EKIQKAFDDIAKYFSK	579	19-34
	DR1	18	FGRLQGISPKI	580	101-111
	DR3	21	WEKMKASEKIFYVYMKRK	581	37-54
	DR4	24	KIFYVYMKRKYEAMT	582	45-59
	DR11	25	KIFYVYMKRKYEAM	583	45-58
SSX-4	DP10	2	INKTSGPKRGKHAWTHRLRE	584	151-170
	DR3	21	YFSKKKEWEKMKSSSEKIVVY	585	31-50
	DR8	4	MKLNYEVMTKLGFKVTLPPF	586	51-70
	DR8	4	KHAWTHRLRERKQLVVYEEI	587	161-180
	DR11	25	LGFKVTLPFFMRSKRAADFH	588	61-80
	DR15	20	KSSEKIVVYVMKLNYEVMTK	589	41-60
	DR52	41	KHAWTHRLRERKQLVVYEEI	590	161-180
survivin/ BIRC5	A2	44	ELTLGEFLKL	591	96-106
	A24	20	AYACNTSTL	592	intron B2
TAG-1	A2	44	SLGWLFLLL	593	78-86
	B8	14	LSRLSNRLL	594	42-50
TAG-2	B8	14	LSRLSNRLL	595	42-50
TRAG-3	DR1	18	CEFHACWPAFTVLGE	596	34-48
	DR4	24	CEFHACWPAFTVLGE	597	34-48
	DR7	25	CEFHACWPAFTVLGE	598	34-48
TRP2-INT2g	A68	8	EVISCKLIKR	599	intron 2
XAGE-1b/ GAGED2a	A2	44	RQKKIRIQL	600	21-29
	DR4	24	HLGSRQKKIRIQLRSQ	601	17-32
	DR9	3	CATWKVICKSCISQTPG	602	33-49
BCR-ABL (b3a2)	A2	44	<u>SSKALQRPV</u>	603	926-934
	B8	14	<u>GFKQSSKAL</u>	604	922-930
	DR4	24	<u>ATGFKQSSKALQRPVAS</u>	605	920-936
	DR9	3	<u>ATGFKQSSKALQRPVAS</u>	606	920-936

TABLE 3

Overexpressed tumor specific antigen						
Gene/Protein	Normal tissue expression	HLAa	HLA	Peptide	SEQ ID NO	Position
			Freq (%)			
adipophilin	adipocytes, macrophages	A2	44	SVASTITGV	607	129-137
AIM-2	ubiquitous (low level)	A1	26	RSDSGQQARY	608	intron
ALDH1A1	mucosa, keratinocytes	A2	44	LLYKLADLI	609	88-96
BCLX (L)	ubiquitous (low level)	A2	44	YLNDHLEPWI	610	173-182
BING-4	ubiquitous (low level)	A2	44	CQWGRWLQQL	611	ORF2
CALCA	thyroid	A2	44	VLLQAGSLHA	612	16-25
CD45	proliferating cells, testis, multiple tissues (low level)	A24	20	KFLDALISL	613	556-564
CD274	multiple tissues (lung, heart, dendritic cells, etc.) and induced by IFN-γ	A2	44	LLNAFTVTW	614	15-23
CPSF	ubiquitous (low level)	A2	44	KVHPVIWSL	615	250-258
		A2	44	LMLQNALTTM	616	1360-1369
cyclin D1	ubiquitous (low level)	A2	44	LLGATCMFV	617	101-109
		DR4	24	NPPSMVAAGSVVAAV	618	198-212
DKK1	testis, prostate, mesenchymal stem cells	A2	44	ALGGHPLLGV	619	20-29
ENAH (hMen a)	breast, prostate stroma and epithelium of colon-rectum, pancreas, endometrium	A2	44	TMNGSKSPV	620	502-510
EpCAM	epithelial cells	A24	20	RYQLDPKFI	621	173-181
EphA3	many	DR11	25	DVTFNIIICKKCG	622	356-367
EZH2	ubiquitous (low level)	A2	44	FMVEDETVL	623	120-128
		A2	44	FINDEIFVEL	624	165-174
		A24	20	KYDCFLHPF	625	291-299
		A24	20	KYVGIEREM	626	735-743
FGF5	brain, kidney	A3	22	NTYASPRFKF	627	172-176 and 217-220
glypican-3	placental and multiple tissues	A2	44	FVGEFFTDV	628	144-152
		A24	20	EYILSLEEL	629	298-306
G250/MN/CAIX	stomach, liver, pancreas	A2	44	HLSTAFARV	630	254-262
HER-2/neu	ubiquitous (low level)	A2	44	KIFGSLAFL	631	369-377
		A2	44	IISAVVGIL	632	654-662
		A2	44	ALCRWGLLL	633	5-13
		A2	44	ILHNGAYSL	634	435-443
		A2	44	RLLQETELV	635	689-697
		A2	44	VVLGVVFGI	636	665-673
		A2	44	YMIMVKCWM	637	952-961
		A2	44	HLYQQCQVV	638	48-56
		A2	44	YLVPQQGFFC	639	1023-1032
		A2	44	PLQPQBLQV	640	391-399
		A2	44	TLEEITGYL	641	402-410
		A2	44	ALIHMHNTHL	642	466-474
		A2	44	PLTSIIISAV	643	650-658
		A3	22	VLRENTSPK	644	754-762
		A24	20	TYLPPTNASL	645	63-71
HLA-DOB	B lymphocytes, monocytes, blood cells, adrenals, . . .	A2	44	FLLGLIFLL	646	232-240

TABLE 3-continued

Overexpressed tumor specific antigen						
Gene/Protein	Normal tissue expression	HLAa	HLA Freq (%)	Peptide	SEQ ID NO	Position
Hepsin	kidney, liver, skin,	A2	44	SLLSGDWVL	647	191-199
		A2	44	GLQLGVQAV	648	229-237
		A2	44	PLTEYYIQPV	649	268-276
IDO1	lymph nodes, placenta, and many cell types in the course of inflammatory response	A2	44	ALLEIASCL	650	199-207
IGF2B3	ubiquitous (low level)	A2	44	NLSSAEVVV	651	515-523
		A3	44	RLLVPTQFV	652	199-207
IL13Ralpha2		A2	44	WLPFGFIL	653	345-353
Intestinal carboxyl esterase	liver, intestine, kidney	B7	17	SPRWWP	654	alt. ORF
alpha-foetoprotein	liver	A2	44	GVALQTMKQ	655	542-550
		A2	44	FMNKFYIEI	656	158-166
		DR13	19	QLAVSVILRV	657	364-373
Kallikrein 4	prostate and ovarian cancer cancer	A2	44	FLGYLILGV	658	11-19
		DP4	75	SVSESDTIRSIAS	659	125-139
		DR4	24	LLANGRMPVTLQCVN	660	155-169
		DR7	25	RMPTVLCVNVS	661	160-174
KIF20A	ubiquitous (low level)	A2	44	LLSDDDV	662	12-20
		A2	44	AQPDTAPLPV	663	284-293
		A2	44	CIAEQYHTV	664	809-817
Lengsin	eye lens and low level in multiple tissues	A2	44	FLPEFGISSA	665	270-279
M-CSF	liver, kidney	B35	20	LPAVVGLSPGEQEY	666	alt. ORF
MCSP	endothelial cells, chondrocytes, smooth muscle cells	DR11	25	VGQDVSVLFRVTGALQ	667	693-708
mdm-2	ubiquitous (brain, muscle, lung)	A2	44	VLFYLGQY	668	53-60
Meloe	ubiquitous (low level)	A2	44	TLNDECWPA	669	36-44
		DQ2	41	ERISSTLNDECWPA	670	31-44
		DQ6	63	FGRLQGISPKI	671	32-44
		DR1	18	TSREQFLPSEGAA	672	11-23
		DR11	25	CPPWHPSERISSTL	673	24-37
Midkine	ubiquitous (low level)	A2	44	ALLALTSAV	674	13-21
		A2	44	AQCQETIRV	675	114-122
		DR4	24	LTLLALLALTSAVAK	676	9-23
MMP-2	ubiquitous	A2	44	GLPPDVQRVH	677	560-568
MMP-7	ubiquitous (low level)	A3	22	SLFPNSPKWTSK	678	96-107
MUC1	glandular epithelia	A2	44	STAPPVHN	679	950-958
		A2	44	LLLLTVLTV	680	12-20
		DR3	21	PGSTAPPAGV	681	repeated region
MUC5AC	surface mucosal cells, respiratory tract, and stomach epithelia	A24	20	TCQPTCRSL	682	716-724
p53	ubiquitous (low level)	A2	44	LLGRNSFEV	683	264-272
		A2	44	RMPEAAPV	684	65-73
		B46	0.1	SQKTYQGSY	685	99-107
		DP5	3	PGTRVRAMAIYKQ	686	153-165
		DR14	6	HLIRVEGNLRVE	687	193-204

TABLE 3-continued

Overexpressed tumor specific antigen						
Gene/Protein	Normal tissue expression	HLAa	HLA Freq (%)	Peptide	SEQ ID NO	Position
PAX5	hemopoietic system	A2	44	TLPGYPPHV	688	311-319
PBF	ovary, pancreas, spleen, liver	B55	4	CTACRWKKACQR	689	499-510
PRAME	testis, ovary, endometrium, adrenals	A2	44	VLDGLDVLL	690	100-108
		A2	44	SLYSFPEPEA	691	142-151
		A2	44	ALYVDSLFFL	692	300-309
		A2	44	SLLQHLIGL	693	425-433
		A24	20	LYVDSLFFLC	694	301-309
PSMA	prostate, CNS, liver	A24	20	NYARTEDFF	695	178-186
RAGE-1	retina	A2	44	LKLSGGVRL	696	352-360
		A2	44	PLPPARNNGLG	697	32-40
		B7	17	SPSSNRIRNT	698	11-20
RGS5	heart, skeletal muscle, pericytes	A2	44	LAALPHSCL	699	5-13
RhoC	ubiquitous (low level)	A3	22	RAGLQVRKNK	701	176-185
		A2	44	ALWPWILLMA(T)	702	11-19(20)
RNF43		A24	20	NSQPVWLCL	703	721-729
		B7	17	LPRWPPPQL	704	antisense
secernin 1	ubiquitous	A2	44	KMDAEHPEL	705	196-204
SOX10	ubiquitous (low level)	A2	44	AWISKPPGV	706	332-340
		A2	44	SAWISKPPGV	707	331-340
STEAP1	prostate	A2	44	MIAVFPLIV	708	292-300
		A2	44	HQQYFYKIPILVINK	709	102-116
survivin	ubiquitous	A2	44	ELTLGEFLKL	710	95-104
		DR1	18	TLGEFLKLDRERAKN	711	97-111
Telomerase	testis, thymus, bone marrow, lymph nodes	A2	44	ILAKFLHWLE	712	540-548
		A2	44	RLVDDFLLV	713	865-873
		DR7	25	RPGLLGASVLGLDDI	714	672-686
		DR11	25	LTDLQPYMRQFVAHL	715	766-780
		hTERT (572)		WLFFYRKSV(R)	716	572-581
TPBG	multiple tissues (esophagus, bladder, etc.)	A2	44	RLARLALV	717	17-25
VEGF	ubiquitous (low level)	B27	7	SRFGGAVVR	718	-i
WT1	testis, ovary, bone marrow, spleen	A1	26	TSEKRPPMCAY	719	317-327
		A24	20	CMTWNQMNL	720	235-243
		DP5	3	LSHLQMHSRKH	721	337-347
		DP5	3	KRYFKLSHLQMHSRKH	722	332-347
		DR4	24	KRYFKLSHLQMHSRKH	723	332-347

The identification of heteroclitic peptides are described in the art. For example, in previous studies (Selby, et al., *J. Immunol.*, 162(2):669 (1999), Skipper, et al., *J. Exp. Med.* 183:527 (1996), the entire contents of each of which are incorporated herein by reference), heteroclitic epitopes were fortuitously identified by eluting naturally occurring mutant peptides from melanoma cells, or by systematically screening a large number of epitopes consisting of substitutions at almost every position in the epitope (Zaremba, et al., *Cancer Research*, 57:4570 (1997), Loftus, et al., *Cancer Research* 58:2433 (1998), Blake, et al., *J. Exp. Med.* 18:121 (1996), the entire contents of each of which are incorporated herein by reference). Alternatively, heteroclitic epitopes were identified by screening random combinatorial peptide libraries which also has required the arduous synthesis and screening of large numbers of peptides (Pinilla, et al., *Current Opinion in Immunology* 11:193-202 (1999), the entire contents of each of which are incorporated herein by reference). Genetic approaches, such as screening of DNA expression libraries, have provided another method for generating CTL epitopes and analogs (Boon, et al., *Annu. Rev. Immunol.* 12:337-65 (1994), Gavin, et al., *Eur. J. Immunol.* 24(9):2124-33 (1994), the entire contents of each of which are incorporated herein by reference).

Generating Cancer Vaccine in Tumor Cells

Some aspects of the present disclosure provide systems, compositions, and methods of editing genes encoding tumor specific antigens *in vivo* (e.g., in tumor cells in a subject) or *ex vivo* (e.g., in isolated tumor cells) to introduce mutations in the genes encoding tumor specific antigens. In some embodiments, such mutations lead to the production of heteroclitic peptides that are more immunogenic than native peptides of the tumor-specific antigen. In some embodiments, such mutations lead to the translation of a non-coding region of the tumor specific antigen, which results in cryptic peptides that are more immunogenic than any native peptides from the tumor specific antigen.

The gene editing methods described herein, rely on nucleobase editors as described in, in U.S. Pat. No. 9,068,179, US Patent Application Publications US 2015/0166980, US 2015/0166981, US 2015/0166982, US20150166984, and US20150165054, and US Provisional Applications, U.S. Ser. No. 62/245,828, filed Oct. 23, 2015; 62/279,346, filed Jan. 15, 2016; 62/311,763, filed Mar. 22, 2016; 62/322,178, filed Apr. 13, 2016, 62/357,352, filed Jun. 30, 2016, U.S. Pat. No. 62,370,700, filed Aug. 3, 2016; 62/398,490, filed Sep. 22, 2016; 62/408,686, filed Oct. 14, 2016; PCT Application PCT/US2016/058344, filed Oct. 22, 2016, U.S. patent application Ser. No. 15/311,852, filed Oct. 22, 2016; and in Komor et al., *Nature*, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage, 533, 420-424 (2016), the entire contents of each of which are incorporated herein by reference.

The nucleobase editors are highly efficient at precisely editing a target base in any of the tumor associated antigen genes described herein, and a DNA double stand break is not necessary for the gene editing, thus reducing genome instability and preventing possible oncogenic modifications that may be caused by other genome editing methods. The nucleobase editors described herein may be programmed to target and modify a single base. In some embodiments, the target base is a cytosine (C) base and may be converted to a thymine (T) base via deamination by the nucleobase editor.

To edit the polynucleotide encoding a tumor associated antigen, the polynucleotide is contacted with a nucleobase editor as described herein. In some embodiments, the tumor-associated antigen encoding polynucleotide is contacted

with a nucleobase editor and a guide nucleotide sequence, wherein the guide nucleotide sequence targets the nucleobase editor the target base (e.g., a C base) in the tumor-associated antigen encoding polynucleotide.

In some embodiments, the tumor-associated antigen encoding polynucleotide is the tumor-associated antigen gene locus in the genomic DNA of a cell (e.g., a tumor cell). In some embodiments, the tumor cell is a cultured cell. In some embodiments, the tumor cell is *in vivo*. In some embodiments, the tumor cell is *ex vivo*. In some embodiments, the tumor cell is from a mammal. In some embodiments, the mammal is a human. In some embodiments, the mammal is a rodent. In some embodiments, the rodent is a mouse. In some embodiments, the rodent is a rat.

In some embodiments, the tumor-associated antigen encoding polynucleotide may be a DNA molecule comprising a coding strand and a complementary strand, e.g., the tumor-associated antigen gene locus in the genome of a tumor cell. In some embodiments, the tumor-associated antigen encoding polynucleotide may also include coding regions (e.g., exons) and non-coding regions (e.g., introns or splicing sites). In some embodiments, the target base (e.g., a C base) is located in the coding region (e.g., an exon) of the tumor-associated antigen encoding polynucleotide. As such, the conversion of a base in the coding region may result in an amino acid change in the tumor-associated antigen protein sequence, i.e., a mutation. Tumor associated antigens comprising the desired mutation(s), once degraded (e.g., via any of the protein degradation pathways, such as degradation by the proteasome) results in immunogenic heteroclitic epitopes.

In some embodiments, the target base is located in a non-coding region of the tumor-associated antigen gene, e.g., in an intron or a splice site. In some embodiments, a target base is located in a splice site, and the editing of such target base causes alternative splicing of the tumor-associated antigen mRNA. In some embodiments, the alternative splicing leads to translation of a non-coding region of the tumor-associated antigen gene, generating cryptic epitopes. The immunogenic epitopes (e.g., heteroclitic epitopes or cryptic epitopes) may be presented by the tumor cell, or a professional antigen presenting cell, and be recognized by the immune system, thus eliciting a tumor-specific immune response (e.g., T-cell response or B-cell response).

To edit a tumor-associated antigen gene, the tumor-associated antigen gene (a polynucleotide molecule) may be contacted with the nucleobase editor, wherein the nucleobase editor binds to its target sequence and edits the desired base. For example, the nucleobase editor may be expressed in a cell where editing is desired (e.g., a tumor), allowing editing of the tumor-associated antigen gene by the nucleobase editor. In some embodiments, the binding of the nucleobase editor to its target sequence in the tumor-associated antigen gene is mediated by a guide nucleotide sequence, e.g., a nucleotide molecule comprising a nucleotide sequence that is complementary to one of the strands of the target sequence in the tumor-associated antigen gene. Thus, by designing the guide nucleotide sequence, the nucleobase editor may be programmed to edit any target base in any tumor associated antigen gene. In some embodiments, the guide nucleotide sequence is co-expressed with the nucleobase editor in a tumor cell where editing is desired. In some embodiments, a nucleobase editor/gRNA complex is delivered to the cell where editing is desired (e.g., a tumor cell).

Provided herein are non-limiting, exemplary heteroclitic epitopes and cryptic epitopes that may be produced via base editing and strategies for making them.

Codon Change

Using the nucleobase editors described herein, several amino acid codons may be converted to a different codon via deamination of a target base within the codon. For example, in some embodiments, a cytosine (C) base is converted to a thymine (T) base via deamination by a nucleobase editor comprising a cytosine deaminase domain (e.g., APOBEC1 or AID). It is worth noting that during a C to T change via deamination (e.g., by a cytosine deaminase such as APOBEC1 or AID), the cytosine is first converted to a uridine (U), leading to a G:U mismatch. The G:U mismatch is then converted by DNA repair and replication pathways to T:A pair, thus introducing the thymine at the position of the original cytosine. As such, deamination of a C base results in a C-G base pair being replaced by a T-A base pair.

As is familiar to one skilled in the art, conversion of a base in an amino acid codon may lead to a change of the amino acid the codon encodes. Cytosine deaminases are capable of converting a cytosine (C) base to a thymine (T) base via deamination. Thus, it is envisioned that, for amino acid codons containing a C base, the C base may be directly converted to T. For example, codon for leucine (CTC) may be changed to a TTC (phenylalanine) codon via the deamination of the first C on the coding strand. For amino acid codons that contain a guanine (G) base, a C base is present on the complementary strand; and the G base may be converted to an adenine (A) via the deamination of the C on the complementary strand. For example, an ATG (Met/M) codon may be converted to a ATA (Ile/I) codon via the deamination of the third C on the complementary strand. In some embodiments, two C to T changes are required to convert a codon to a different codon. Non-limiting examples of possible mutations that may be made in any tumor associated antigens by the nucleobase editors of the present disclosure are summarized in Table 4.

TABLE 4

Exemplary Codon Changes via Base Editing		
Target codon	Base-editing reaction (s)	Edited codon
CTT (Leu/L)	1st base C to T on coding strand	TTT (Phe/F)
CTC (Leu/L)	1st base C to T on coding strand	TTC (Phe/F)
ATG (Met/M)	3rd base C to T on complementary strand	ATA (Ile/I)
GTT (Val/V)	1st base C to T on complementary strand	ATT (Ile/I)
GTA (Val/V)	1st base C to T on complementary strand	ATA (Ile/I)
GTC (Val/V)	1st base C to T on complementary strand	ATC (Ile/I)
GTG (Val/V)	1st base C to T on complementary strand	ATG (Met/M)
TCT (Ser/S)	2nd base C to T on coding strand	TTT (Phe/F)
TCC (Ser/S)	2nd base C to T on coding strand	TTC (Phe/F)
TCA (Ser/S)	2nd base C to T on coding strand	TTA (Leu/L)
TCG (Ser/S)	2nd base C to T on coding strand	TTG (Leu/L)
AGT (Ser/S)	2nd base C to T on complementary strand	AAT (Asp/N)
AGC (Ser/S)	2nd base C to T on complementary strand	AAC (Asp/N)
CCT (Pro/P)	1st base C to T on coding strand	TCT (Ser/S)
CCC (Pro/P)	1st base C to T on coding strand	TCC (Ser/S)
CCA (Pro/P)	1st base C to T on coding strand	TCA (Ser/S)
CCG (Pro/P)	1st base C to T on coding strand	TCG (Ser/S)
CCT (Pro/P)	2nd base C to T on coding strand	CTT (Leu/L)
CCC (Pro/P)	2nd base C to T on coding strand	CTC (Leu/L)
CCA (Pro/P)	2nd base C to T on coding strand	CTA (Leu/L)
CCG (Pro/P)	2nd base C to T on coding strand	CTG (Leu/L)
ACT (Thr/T)	2nd base C to T on coding strand	ATT (Leu/L)
ACC (Thr/T)	2nd base C to T on coding strand	ATC (Leu/L)
ACA (Thr/T)	2nd base C to T on coding strand	ATA (Leu/L)
ACG (Thr/T)	2nd base C to T on coding strand	ATG (Met/M)
GCT (Ala/A)	2nd base C to T on coding strand	GTT (Val/V)
GCC (Ala/A)	2nd base C to T on coding strand	GTC (Val/V)

TABLE 4-continued

Exemplary Codon Changes via Base Editing		
Target codon	Base-editing reaction (s)	Edited codon
GCA (Ala/A)	2nd base C to T on coding strand	GTA (Val/V)
GCG (Ala/A)	2nd base C to T on coding strand	GTG (Val/V)
GCT (Ala/A)	1st base C to T on complementary strand	ACT (Thr/T)
GCC (Ala/A)	1st base C to T on complementary strand	ACC (Thr/T)
GCA (Ala/A)	1st base C to T on complementary strand	ACA (Thr/T)
GCG (Ala/A)	1st base C to T on complementary strand	ACG (Thr/T)
CAT (His/H)	1st base C to T on complementary strand	TAT (Tyr/Y)
CAC (His/H)	1st base C to T on complementary strand	TAC (Tyr/Y)
GAT (Asp/D)	1st base C to T on complementary strand	AAT (Asp/N)
GAC (Asp/D)	1st base C to T on complementary strand	AAC (Asp/N)
GAA (Glu/E)	1st base C to T on complementary strand	AAA (Lys/K)
GAG (Glu/E)	1st base C to T on complementary strand	AAG (Lys/K)
TGT (Cys/C)	2nd base C to T on complementary strand	TAT (Tyr/Y)
TGC (Cys/C)	2nd base C to T on complementary strand	TAC (Tyr/Y)
CGT (Arg/R)	1st base C to T on coding strand	TGT (Cys/C)
CGC (Arg/R)	1st base C to T on coding strand	TGC (Cys/C)
AGA (Arg/R)	2nd base C to T on complementary strand	AAA (Lys/K)
AGG (Arg/R)	2nd base C to T on complementary strand	AAG (Lys/K)
CGG (Arg/R)	2nd base C to T on complementary strand	CAG (Gln/Q)
CGG (Arg/R)	1st base C to T on coding strand	TGG (Trp/W)
GGT (Gly/G)	2nd base C to T on complementary strand	GAT (Asp/D)
GGC (Gly/G)	2nd base C to T on complementary strand	GAC (Asp/D)
GGA (Gly/G)	2nd base C to T on complementary strand	GAA (Glu/E)
GGG (Gly/G)	2nd base C to T on complementary strand	GAG (Glu/E)
GGT (Gly/G)	1st base C to T on complementary strand	AGT (Ser/S)
GGC (Gly/G)	1st base C to T on complementary strand	AGC (Ser/S)
GGA (Gly/G)	1st base C to T on complementary strand	AGA (Arg/R)
GGG (Gly/G)	1st base C to T on complementary strand	AGG (Arg/R)

Such amino acid substitutions introduced via base editing generate heteroclitic epitopes. Non-limiting examples of heteroclitic epitopes that may be generated from tumor associated antigens by nucleobase editors are summarized in Table 5.

In some embodiments, to bind to its target sequence and edit the desired base, the nucleobase editor depends on its guide nucleotide sequence (e.g., a guide RNA). In some embodiments, the guide nucleotide sequence is a gRNA sequence. A gRNA typically comprises a tracrRNA framework allowing for Cas9 binding, and a guide sequence, which confers sequence specificity to fusion proteins disclosed herein. In some embodiments, the guide RNA comprises the structure 5'-[guide sequence]-guuuuagcua-gaaaua-
gaggcaccgagucggcguuu uuuu-3' (SEQ ID NO: 336), wherein the guide sequence comprises a sequence that is complementary to the target sequence. Other suitable tracrRNA framework sequences are provided in Table 11. The guide sequence is typically about 20 nucleotides long. In certain embodiments, the guide sequence may be 15-25 nucleotides long. In some embodiments, the guide sequence is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides long. Such suitable guide RNA sequences typically comprise guide sequences that are complementary to a nucleic sequence within 50 (e.g., within 50, 45, 40, 35, 30, 35, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, or 10) nucleotides upstream or downstream of the target nucleotide to be edited.

Guide sequences that may be used to target the nucleobase editor to its target sequence to induce specific mutations in tumor associated antigen genes are provided in Table 5. It is to be understood that the mutations and guide sequences presented herein are for illustration purpose only and are not meant to be limiting.

TABLE 5

Heteroclitic epitopes						
Antigen Name/Epitope amino acid position (Exemplary condition)	Heteroclitic epitope (mutation), Genomic target region	Programmable guide-RNA sequence	SEQ ID NO	gRNA (PAM) (C-edited)	Genome Editor type ^a	
gp100/209-217 (melanoma)	<u>I</u> IDQVPFSV (T2101) (SEQ ID NO: 786) ttcacccat <u>a</u> Ct ggtaagggttta ggaaggggca (SEQ ID NO: 831)	GCCUUACACCAUUAUCUGGUAA CACAUUACUGGUAGGGUU ACCAUUACUGGUAGGGUUU CCAUUACUGGUAGGGUUUA AUUACUGGUAGGGUUUAGG UUACUGGUAGGGUUUAGGA UACUGGUAGGGUUUAGGAAG ACUGGUAGGGUUUAGGAAG CAGCCUUACACCAUUAUCUGGU UUACUGGUAGGGUUUAGGA	724 870 871 872 873 	(GGG) (TAG) (AGG) (GGAA) (AAG) (AGG) (GGG) (GGG) (AAGGT) (AGGG)	20 (C14) 20 (C9) 20 (C8) 20 (C7) 20 (C5) 20 (C4) 20 (C3) 20 (C2) 20 (C16) 20 (C5)	SpBE3 SpBE3 SpBE3 VQR-SpBE3 SpBE3 SpBE3 SpBE3 SpBE3 SaBE3 St3BE3
gp100/280-288 (melanoma)	YLEPGPVTV (A288V) (SEQ ID NO: 818) agt <u>c</u> actg <u>C</u> cc agg <u>t</u> gt <u>c</u> c <u>t</u> g cagg <u>t</u> gc <u>c</u> att cc (SEQ ID NO: 832)	UCACUGCCCAGGGGUCCUG CACUGCCCAGGGGUCCUGC	725 889	(CAG) (AGG)	20 (C7) 20 (C6)	SpBE3 SpBE3
gp100/154-162 (melanoma)	<u>K</u> IWGQYW QV (T1551) (SEQ ID NO: 787) tat <u>t</u> ct <u>g</u> ga <u>q</u> a <u>C</u> t <u>g</u> gg <u>g</u> tg agg <u>g</u> act <u>c</u> c <u>t</u> t ct (SEQ ID NO: 833)	AUGUCUGGAAGACCUGGGU UGUCUGGAAGACCUGGGUG GUCUGGAAGACCUGGGUGA UCUGGAAGACCUGGGUGAG	726 890 891 892	(GAG) (AGG) (GGG) (GGAC)	20 (C13) 20 (C12) 20 (C11) 20 (C10)	SpBE3 SpBE3 SpBE3 VQR-SpBE3
MART-1/26-35 (melanoma)	EVAGIGILT V (A27V) (SEQ ID NO: 819) tgg <u>t</u> cc <u>ag</u> ggc <u>cg</u> C <u>agg</u> ate gg <u>c</u> at <u>c</u> c <u>t</u> g <u>at</u> c gtgg <u>t</u> c <u>t</u> gg ga (SEQ ID NO: 834)	GGCCGCAGGGAUCCGCAUCC GCAGGGAUCCGCAUCCUGAU UGGUCCAGGGCGCAGGGAU AGGGCCGCAGGGAUCCGCAU	727 893 894 895	(TGAT) (CGTG) (CGG) (CCTGAT)	20 (C6) 20 (C2) 20 (C14) 20 (C8)	VQR-SpBE3 VQR-SpBE3 SpBE3 KKH-SaBE3
NY-ESO-1/157- 165 (multiple tumors, e.g., melanoma and breast cancer)	SLLMWITQ Y (C165Y) (SEQ ID NO: 788) ctt <u>c</u> c <u>t</u> gtt <u>g</u> at gtgg <u>t</u> ac <u>ac</u> gc agt <u>G</u> ctt <u>t</u> tc <u>g</u> c cc (SEQ ID NO: 835)	ACUGCGUGAUCCACAUCAAC CACUGCGUGAUCCACAUCAA	728 897	(AGG) (CAG)	20 (C-1) 20 (C1)	SpBE3 SpBE3
TYR/369-377 (melanoma)	YMNG <u>I</u> MSQ V (T3731) (SEQ ID NO: 789) tat <u>g</u> a <u>t</u> gg <u>aa</u> <u>C</u> aat <u>t</u> cc <u>c</u> ag gt <u>ac</u> agg <u>g</u> at <u>c</u> tg <u>cc</u> (SEQ ID NO: 836)	AUGGAACAAUGUCCAGGU UGGAACAAUGUCCAGGUAC GGAACAAUGUCCAGGUACA GAACAAUGUCCAGGUACAG UGGAACAAUGUCCAGGUAC	729 898 899 900 901	(CAG) (AGG) (GGG) (GGAT) (AGGGAT)	20 (C7) 20 (C6) 20 (C5) 20 (C4) 20 (C6)	SpBE3 SpBE3 SpBE3 VQR-SpBE3 SaBE3

TABLE 5-continued

Heteroclitic epitopes						
Antigen Name/Epitope amino acid position (Exemplary condition)	Heteroclitic epitope (mutation), Genomic target region	Programmable guide-RNA sequence	SEQ ID NO	(PAM)	gRNA size (C-edited)	Genome Editor type ^a
TyRP-1/240- 251 (melanoma)	DAEKYDIC TDEY (C244Y) (SEQ ID NO: 790) tgggactggcg ggatgcagaa aagtGtgacat ttgcacagatg agtacatggga (SEQ ID NO: 837)	CACACUUUUCUGCAUCCGC GUCACACUUUUCUGCAUCCC	730 902	(CAG) (GCCAGT)	20 (C3) 20 (C5)	SpBE3 KKH-SaBE3
Survivin/95-102 (multiple tumors, e.g., melanoma, breast cancer, and leukemia)	ELILGEFLK L (T971) (SEQ ID NO: 791) tttgaagaatta a <u>C</u> ccttggta attttgaactg gac (SEQ ID NO: 838)	AUUAACCCUUGGUGAAUUUU CCUUGGUGAAUUUUUGAAC	731 903	(TGAA) (TGG)	20 (C6) 20 (C-1)	VQR-SpBE3 SpBE3
HER2/657-666 (breast cancer)	AMVGILLVV V (V658M) (SEQ ID NO: 793) gttccgcggccc agccctgtac gtccatcatctct gcg <u>G</u> tgggtgg cattctgtggtc (SEQ ID NO: 839)	CGCAGAGAUGAUGGACGUCA ACCGCAGAGAUGAUGGACGU CCAACCACCGCAGAGAUGAU GCCAACCCACCGCAGAGA AAUGCCAACCACCGCAGAGA AGAAUGCCAACCACCGCAGA CCCGAGAGAUGAUGGACGUC	732 904 905 906 907 908 909	(GAG) (CAG) (GGAC) (TGG) (TGAT) (GATGAT) (AGAG)	20 (C-1) 20 (C2) 20 (C8) 20 (C9) 20 (C12) 20 (C14) 20 (C1)	VQR-SpBE3 SpBE3 VQR-SpBE3 SpBE3 VQR-SpBE3 KKH-SaBE3 EQR-SpBE3
HER2/911-920 (breast cancer)	IIWELMTFG A (T9121) (SEQ ID NO: 794) aggtgtgaCtg tgtggagctg atgactttggg gCcaaacctta cgatggatcc cageccggga gatccct (SEQ ID NO: 840)	AGGUGUGACUGUGGGAGC UGUGACUGUGUGGGAGCU UUAGGUGUGACUGUGUGGA GUGUGGGAGCUGAUGACUU aggtgtgaCtg tgtggagctg atgactttggg gCcaaacctta cgatggatcc cageccggga gatccct (SEQ ID NO: 840)	733 910 911 912	(TGAT) (TGAC) (GCTGAT) (TGGGG)	20 (C9) 20 (C6) 20 (C11) 20 (C-2)	VQR-SpBE3 VQR-SpBE3 KKH-SaBE3 St3BE3
HER2/911-920 (breast cancer)	ITWELMTF GV (A920V) (SEQ ID NO: 795) gacttttgggg <u>C</u> caaaccttac gatgggatccc agccccgg (SEQ ID NO: 841)	ACUUUUGGGGCCAACCUUA UUUGGGGCCAACCUUACGA UUGGGGCCAACCUUACGAU UUUGGGGCCAACCUUACGA gacttttgggg <u>C</u> caaaccttac gatgggatccc agccccgg (SEQ ID NO: 841)	734 913 914 915	(CGAT) (TGG) (GGG) (TGGGAT)	20 (C11) 20 (C8) 20 (C7) 20 (C8)	VQR-SpBE3 SpBE3 SpBE3 SaBE3
hTERT/540- 549 (breast cancer)	I LAKFLHWL I (M5491) (SEQ ID NO: 792) cgtgaggagat cctggccaagtt cctgcactggct gat <u>G</u> agtgtgt acgtcgctcgag ctg (SEQ ID NO: 842)	CAUCAGGCCAGUGCAGGAACU CACACUCAUCAGCCAGUGCA ACACACUCAUCAGCCAGUGC UACACACUCAUCAGCCAGUG GACGUACACACUCAUCAGCC GACGACGUACACACUCAUCA	735 916 917 918 919 920	(TGG) (GGAA) (AGG) (CAG) (AGTG) (GCCAGT)	20 (C1) 20 (C7) 20 (C8) 20 (C9) 20 (C12) 20 (C11)	SpBE3 VQR-SpBE3 SpBE3 SpBE3 VQR-SpBE3 KKH-SaBE3

TABLE 5-continued

Heteroclitic epitopes							
Antigen Name/Epitope amino acid position (Exemplary condition)	Heteroclitic epitope (mutation), Genomic target region	Programmable guide-RNA sequence	SEQ ID NO	(PAM)	gRNA size (C-edited)	Genome Editor type ^a	
hTERT/572- 580 (breast cancer)	WLFFYRKS V (R572W) (SEQ ID NO: 716) ccaaagcctat cttttctgtatgc Cggcttcattg gtcacacttcgg ttcca (SEQ ID NO: 843)	UUCUGAUGCUCGGCUCUUUC (SEQ ID NO: 716)	736	(TGG)	20 (C11)	SpBE3	
SSX2/41-49 (multiple tumors)	KVSEKIFYV (A42V) (SEQ ID NO: 797) ggaaaagatg aaagCctegg agaaaatcttct atgttatatatga agagaaagtat gaggctatgac t (SEQ ID NO: 844)	GAAAAGAUGAAAGCCUCGGA GCCUCGGAGAAAUCUUCUA (SEQ ID NO: 797)	737 921	(GAAAAT) (TGTG)	20 (C14) 20 (C2)	KKH-SaBE3 VQR-SpBE3	
WT1/235-243 (Leukemia)	YMTWNQM NL (C235Y) (SEQ ID NO: 798) attataccaaa tgacatcccag cttgaatGcatg acctggaa (SEQ ID NO: 845)	CAGGUCAUGCAUUCAGCUG CCAGGUCAUGCAUUCAGCU UCCAGGUCAUGCAUUCAGC GCAUUCAAGCUGGAUGUCA UCCAGGUCAUGCAUUCAGC (SEQ ID NO: 845)	738 922 923 924 925	(GGAT) (GGG) (TGG) (TTTGGT) (TCGGAT)	20 (C10) 20 (C11) 20 (C12) 20 (C2) 20 (C12)	VQR-SpBE3 SpBE3 SpBE3 KKH-SaBE3 SaBE3	
WT1/235-243 (Leukemia)	CITWNQMN L (M236I) (SEQ ID NO: 799) attataccaaa tgacatcccag cttgaatGcatG acctggaatca gat (SEQ ID NO: 846)	UGAUUCCAGGUCAUGCAUUC UCCAGGUCAUGCAUUCAGC CAGGUCAUGCAUUCAGCUG CCAGGUCAUGCAUUCAGCU UCCAGGUCAUGCAUUCAGC (SEQ ID NO: 846)	739 926 927 928 929	(AAG) (TGGGAT) (GGAT) (GGG) (TGG)	20 (C12) 20 (C8) 20 (C6) 20 (C7) 20 (C8)	SpBE3 SaBE3 VQR-SpBE3 SpBE3 SpBE3	
CD33/65-73 (Leukemia)	VIISGDSPV (A65V) (SEQ ID NO: 796) ctggttccggg aaggagCcat tatattccagg actctccagt (SEQ ID NO: 847)	GGGAAGGAGCCAUUAUACC GGAAAGGAGCCAUAUAUCCA GAAGGAGCCAUUAUACCAG GCCAUUAUACCGGGACUC ctggttccggg aaggagCcat tatattccagg actctccagt (SEQ ID NO: 847)	740 930 931 932	(AGG) (GGG) (GGAC) (TCCAGT)	20 (C10) 20 (C9) 20 (C8) 20 (C2)	SpBE3 SpBE3 VQR-SpBE3 KKH-SaBE3	
EpCAM/184- 192 (multiple tumors)	ILYENNVI (T1921) (SEQ ID NO: 800) atgagaataat gttataca <u>tattg</u> atctgggtcaa attttctc (SEQ ID NO: 848)	UAAUGUUUAUCACAUUUGAUC UUAUCACAUUUGAUCUGGUU AAUAAUGUUUAUCACUAUUGA atgagaataat gttataca <u>tattg</u> atctgggtcaa attttctc (SEQ ID NO: 848)	741 933 934	(TGG) (CAAAAT) (TCTGGT)	20 (C12) 20 (C7) 20 (C14)	SpBE3 KKH-SaBE3 KKH-SaBE3	

TABLE 5-continued

Heteroclitic epitopes						
Antigen Name/Epitope amino acid position (Exemplary condition)	Heteroclitic epitope (mutation), Genomic target region	Programmable guide-RNA sequence	SEQ ID NO	(PAM)	gRNA size (C-edited)	Genome Editor type ^a
CEA-CAM/24- 31 (multiple tumors, e.g., colorectal cancer, lung cancer, breast cancer)	LLTFWNPPI — (T3141) (SEQ ID NO: 801) taaccttctggaa accgcggcaC caactgccaagc tcactattgaatc caggccgt (SEQ ID NO: 849)	ACCACUGCCAAGCUCACAUU GGAACCCGCCAACACUGCC ACCACUGCCAAGCUCACAUU	742 935 936	(TGA) (AAG) (TGAA)	20 (C2) 20 (C13) 20 (C2)	SpBE3 SpBE3 VQR-SpBE3
CEA-CAM/310- 318 (multiple tumors, e.g., colorectal cancer, lung cancer, breast cancer)	RITVTTITV — (T3111) (SEQ ID NO: 802) gacactggct caataggaaCc acagtccacgac gtacacgtct atggtaaagtgg atccacgaa (SEQ ID NO: 850)	CCUAAUAGGACCACAGUCA CAAUAGGACCACAGUCACGA GACCACAGUCACAGACGUCA CUAAUAGGACCACAGUAC AGGACCACAGUCACGACGAU	743 937 938 939 940	(CGAC) (CGAT) (CAG) (GACGAT) (CACAGT)	20 (C12) 20 (C9) 20 (C3) 20 (C11) 20 (C5)	VQR-SpBE3 VQR-SpBE3 SpBE3 KKH-SaBE3 KKH-SaBE3
CEA-CAM/687- 695 (multiple tumors, e.g., colorectal cancer, lung cancer, breast cancer)	AVVGIMIGV — (T688V) (SEQ ID NO: 803) tctcttggtctct cagctggggc aCtgtcggt catgatggagt gctgggtgggt t (SEQ ID NO: 851)	UGGGGCCACUGUCGGCAUCA GCCACUGUCGGCAUCAUGAU CACACUGGCAUCAUGAUUG ACUGUCGGCAUCAUGAUU GCUGGGGACACUGUCGGCAU GCCACUGUCGGCAUCAUGAU GCCACUGUCGGCAUCAUGAU GCCACUGUCGGCAUCAUGAU	744 941 942 943 944 945 946 947	(TGAT) (TGG) (GGAG) (GAG) (AGTG) (CATGAT) (TGGAGT) (TGGAG)	20 (C9) 20 (C5) 20 (C4) 20 (C3) 20 (C2) 20 (C11) 20 (C5) 20 (C5)	VQR-SpBE3 SpBE3 EQR-SpBE3 SpBE3 VQR-SpBE3 KKH-SaBE3 SaBE3 St3BE3
CEA-CAM/691- 699 (multiple tumors, e.g., colorectal cancer, lung cancer, breast cancer)	IMIGMLVGV — (V695M) (SEQ ID NO: 804) tctcaqctggg gccactgtcg catcatgttgg aGtgtcggtt gggtt (SEQ ID NO: 852)	UCCAAUCAUGAUGCCGACAG ACUCCAAUCAUGAUGCAG CACUCCAAUCAUGAUGCCGA CAGCACUCCAAUCAUGAUGC CCCCAACAGCACUCCAAUCA AGCACUCCAAUCAUGAUGCC ACCCCAACCAGCACUCCAAU	745 948 949 950 951 952 953	(TGG) (AGTG) (CAG) (CGAC) (TGAT) (GACAGT) (CATGAT)	20 (C-1) 20 (C2) 20 (C3) 20 (C6) 20 (C12) 20 (C5) 20 (C14)	SpBE3 VQR-SpBE3 SpBE3 VQR-SpBE3 VQR-SpBE3 KKH-SaBE3 KKH-SaBE3
MAGEA3/112- 120 (multiple tumors)	KVAELVYF L (H118Y) (SEQ ID NO: 805) aggtgtggccga gtttgttCattttc tgctctcaagt atcgagccag ggagccggtc ac (SEQ ID NO: 853)	GUUGGUUUCAUUUUUCUGCUCC UGGUUCAUUUUUCUGCUCC CACUCCAAUCAUGAUGCCGA CAGCACUCCAAUCAUGAUGC CCCCAACAGCACUCCAAUCA AGCACUCCAAUCAUGAUGCC ACCCCAACCAGCACUCCAAU	746 954	(TCAAGT) (AAG)	20 (C8) 20 (C6)	KKH-SaBE3 SpBE3
MAGEA3/181- 190 (also in MAGE A1 (multiple tumors)	YLGLSYDG LL (C181Y) (SEQ ID NO: 806) ctatgtccttgtc acctGccttagg tcttccttatgtat (SEQ ID NO: 854)	AGGUGACAAGGACAUAGGAG GGCAGGUGACAAGGACAUAG AGGCAGGUGACAAGGACAU UAGGCAGGUGACAAGGACAU CUAGGCAGGUGACAAGGAC UAGGCAGGUGACAAGGACAA AGAGACCUAGGCAAGGUGACA UAGGCAGGUGACAAGGACAU	747 955 956 957 958 959 960 961 962	(TGG) (AGTG) (GAG) (GGAG) (AGG) (TAG) (GGAC) (AGG) (AGGAGT)	20 (C-1) 20 (C2) 20 (C3) 20 (C4) 20 (C5) 20 (C6) 20 (C11) 20 (C13) 20 (C5)	SpBE3 VQR-SpBE3 SpBE3 EQR-SpBE3 SpBE3 SpBE3 VQR-SpBE3 SpBE3 SaBE3 SpBE3

TABLE 5-continued

Heteroclitic epitopes						
Antigen Name/Epitope amino acid position (Exemplary condition)	Heteroclitic epitope (mutation), Genomic target region	Programmable guide-RNA sequence	SEQ ID NO	gRNA size (C-edited)	Genome Editor type ^a	
MAGEA3/181-190 (also in MAGE A2/A12 (multiple tumors)	YLGLSYDG LL (C181Y) (SEQ ID NO: 806) tacatcccttgtca cctGcctggc ctctcttacgtat (SEQ ID NO: 855)	AGGUGACAAGGAUGUACAAG GCAGGUACAAGGAUGUACA GGCAGGUACAAGGAUGUAC GAGGCCAGGCAGGUACAA AGAGGCCAGGCAGGUACAA CAGGCAGGUACAAGGAUGU GAGAGGCCAGGCAGGUAC	748 963 964 965 966 967 968	(TGG) (AGTG) (AAG) (GGAT) (AGG) (ACAAGT) (AAGGAT)	20 (C-1) 20 (C2) 20 (C3) 20 (C11) 20 (C13) 20 (C5) 20 (C14)	SpBE3 VQR-SpBE3 SpBE3 VQR-SpBE3 SpBE3 KKH-SaBE3 SaBE3
MAGEA3/181-190 (also in MAGE A3 (multiple tumors)	YLGLSYDG LL (C181Y) (SEQ ID NO: 806) gtacatctttgcc acctGcctggg ccttccttacgaa (SEQ ID NO: 856)	AGGUGGAAAAGAUUACAAG GCAGGUUGCAAGAUUACA GGCAGGUUGCAAGAUUAC GAGGCCAGGCAGGUUGCAA AGAGGCCAGGCAGGUUGCA CAGGCAGGUUGCAAGAUU GAGAGGCCAGGCAGGUUGC	749 969 970 971 972 973 974	(TGG) (AGTG) (AAG) (AGAT) (AAG) (ACAAGT) (AAAGAT)	20 (C-1) 20 (C2) 20 (C3) 20 (C11) 20 (C13) 20 (C6) 20 (C12)	SpBE3 VQR-SpBE3 SpBE3 VQR-SpBE3 SpBE3 KKH-SaBE3 KKH-SaBE3
MAGEA3/181-190 (also in MAGE A4 (multiple tumors)	YLGLSYDG LL (C181Y) (SEQ ID NO: 806) tacaccccttgtc acctGcctggg ccttccttatgtat (SEQ ID NO: 857)	AGGCAGGUACAAGGGUGUA CAGGCAGGUACAAGGGUG CCAGGCAGGUACAAGGGUG AGGCCAGGCAGGUACAAG AAGGCCAGGCAGGUACAA AAAGGCCAGGCAGGUACAA GAAAGGCCAGGCAGGUAC CCCAGGCAGGUACAAGGU CAGGCAGGUACAAGGGUG AAAGGCCAGGCAGGUACAA	750 975 976 977 978 979 980 981 982 983	(GGTG) (AGG) (TAG) (GGTG) (GGG) (AGG) (AAG) (GTAGGT) (AGGTG) (GGGTG)	20 (C4) 20 (C5) 20 (C6) 20 (C9) 20 (C10) 20 (C11) 20 (C12) 20 (C7) 20 (C5) 20 (C11)	VQR-SpBE3 SpBE3 SpBE3 VQR-SpBE3 SpBE3 SpBE3 SpBE3 KKH-SaBE3 St3BE3 St3BE3
MUC-1/92-101 (multiple tumors)	AIWGQDVT SV (T931) (SEQ ID NO: 807) tcacgtgcac ctggggacag gatgtcaccc gtccccactc caggcca (SEQ ID NO: 858)	UCAGCUGCCACCUGGGGACA CUGGGGACAGGUAGUCACCU ACCUGGGACAGGUAGUCAC	751 984 985	(GGAT) (CGG) (CTCGGT)	20 (C11) 20 (C-1) 20 (C2)	VQR-SpBE3 SpBE3 KKH-SaBE3

^aGenome-editor types abbreviations: SpBE3 = APOBEC1-SpCas9n-UGI; VQR-SpBE3 = APOBEC1-VQR-SpCas9n-UGI; EQR-SpBE3 = APOBEC1-EQR-SpCas9n-UGI; VRER-SpBE3 = APOBEC1-VRER-SpCas9n-UGI; SaBE3 = APOBEC1-SaCas9n-UGI; KKH-SaBE3 = APOBEC1-KKH-SaCas9n-UGI; St3BE3 = APOBEC1-St3Cas9n-UGI; St1BE3 = APOBEC1-St1Cas9n-UGI. Guide sequences (the portion of the guide RNA that targets the nucleobase editor to the target sequence) are provided. The guide sequences may be used with any tracrRNA framework sequences known in the art to generate the full guide RNA sequence.

Cryptic Epitopes

Some aspects of the present disclosure provide strategies for generating cryptic epitopes in tumor cells. In some embodiments, such strategies involve alterations of splicing sites in a tumor associated antigen gene. Altered splicing site may lead to altered splicing of an mRNA that encodes a tumor associated antigen. One outcome of altered splicing is the translation of an otherwise non-coding region of the gene, leading to otherwise "hidden peptides," i.e., cryptic epitopes. The splicing site typically comprises an intron donor site, a Lariat branch point, and an intron acceptor site. The mechanism of splicing are familiar to those skilled in the art. The intron donor site has a consensus sequence of GGGTRAGT, and the C bases paired with the G bases in the intron donor site consensus sequence may be targeted by a nucleobase editor, thereby altering the intron donor site. The

50 Lariat branch point also has a consensus sequence, e.g., YTRAC, wherein Y is a pyrimidine, and R is a purine. The C base in the Lariat branch point consensus sequence may be targeted by the nucleobase editors described herein, leading to skipping of the following exon. The intron acceptor site has a consensus sequence ofYNCAGG, wherein Y is a pyrimidine, and N is any nucleotide. The C base of the consensus sequence of the intron acceptor site, and the C base paired with the G bases in the consensus sequence of the intron acceptor site may be targeted by the nucleobase editors described herein, thereby altering the intron acceptor site, in turn leading to skipping of an exon. 55 General strategies of altering the splicing sites are described in Table 6.

60

65

TABLE 6

Exemplary Alteration of Intron-Exon Junction via Base Editing				
Target site	Consensus Sequence	Base-editing reaction (s)	Edited sequence	Outcome
Intron donor	GGGTRAGT (example)	2 nd or 3 rd base C to T on complementary strand	GAGTRAGT (example)	Intron sequence is translated as exon, in frame premature STOP codon
Lariat branch point	YTRAC (example)	5 th base C to T on coding strand	YTRAT (example)	The following exon is skipped from the mature mRNA, which may affect the coding frame
Intron acceptor	Y(rich)NCAGG (example)	2 nd to last base C to T on complementary strand	Y(rich)NCAAG (example)	The exon is skipped from the mature mRNA, which may affect the coding frame
Start codon	ATG (Met/M)	3 rd base C to T on complementary strand	ATA (Ile/I)	The next ATG is used as start codon, which may affect the coding frame

Non-limiting, exemplary cryptic epitopes that may be produced using the base editing methods described herein are provided in Table 7.

TABLE 7

Cryptic Epitopes						
Antigen Name and Exemplary Condition(s)	Cryptic epitope(s) and Genomic target region	Programmable guide-RNA sequence	SEQ ID NO (PAM)	gRNA size (C-edited)	Genome Editor type ^a	
gp100 (PMEL gene) (melanoma)	VYFFLPDHL (SEQ ID NO: 808) AAGCTTTGTTT ATGTCTGGAAAG ACCTGGggtag ggactcccttcagcc tatcatccccac (SEQ ID NO: 859) intron 4, intron donor site	UCCCCUACCCCAAGGUUUCC GUCCCCUACCCCAAGGUUC AGAAGGGAGUCCCUCACCC UGUCUGGAAGACCUGGGUG GUCUGGAAGACCUGGGUGA AUGUCUGGAAGACCUGGGU GAGAACGGAGUCCCUCACCC	752 (AGAC) 986 (CAG) 987 (AGG) 988 (AGG) 989 (GGG) 990 (GAG) 991 (CAG)	20 (C9) 20 (C9/10) 20 (n.a.) 20 (n.a.) 20 (n.a.) 20 (n.a.) 20 (n.a.)	VQR-SpBE3 SpBE3 WT Cas9 WT Cas9 WT Cas9 WT Cas9 WT Cas9	
gp100 (PMEL gene) (melanoma)	VYFFLPDHL (SEQ ID NO: 808) ccaaaaaaacttcagG CCAATACTGGC AAGTTCT (SEQ ID NO: 860) intron 4, intron acceptor site	CCUGAAGUUUUUUGGAAUGAA UUGGCCUGAAGUUUUUUGGAA AGUAUUGGCCUGAAGUUUU CAGUAUUGGCCUGAAGUUUU CAGUAUUGGCCUGAAGUUUU CAGUAUUGGCCUGAAGUUUU CUUUUCAUCCCCAAACUUC AACUUGCCAGUAUUGGCCUG GCUUUUCAUCCCCAAACUU	753 (AAG) 992 (TGAA) 993 (GGAA) 994 (TGG) 995 (TGAAAT) 996 (AGG) 997 (AAG) 998 (CAG)	20 (C1/2) 20 (C4/5) 20 (C9/10) 20 (C10/11) 20 (C10/11) 20 (n.a.) 20 (n.a.) 20 (n.a.)	SpBE3 VQR-SpBE3 VQR-SpBE3 SpBE3 SaBE3 WT Cas9 WT Cas9 WT Cas9	
TYRP1 (melanoma)	MSLQRQFLR (SEQ ID NO: 809) gcacttttattcaaggc agaatggatgtctcta a (SEQ ID NO: 861) ORF1, target start codon	CUCAUUCUGCUUGAAAAG ACUCAUUCUGCUUGAAA GCACUCAUUCUGCUUGAAA UUAGGAGCACUCAUUCUG GCACUCAUUCUGCUUGAAA UUAGGAGCACUCAUUCUG CACUCAUUCUGCUUGAAA ACUCUUUUUCAAGCAGAAU	754 (AGTG) 999 (GAG) 1000 (AAG) 1001 (TGAA) 1002 (AAGAGT) 1003 (TGAAAT) 1004 (AGAG) 1005 (GAG)	20 (C3) 20 (C4) 20 (C6) 20 (C11) 20 (C6) 20 (C11) 20 (C5) 20 (n.a.)	VQR-SpBE3 SpBE3 SpBE3 VQR-SpBE3 SaBE3 KKH-SaBE3 EQR-SpBE3 WT Cas9	
MGAT5 (melanoma)	VLPDVFIRCV (SEQ ID NO: 810) tcatacgtctgtgggt tttctgtcttacagTT GGGTTC (SEQ ID NO: 862) target intron2-	UGUAAGACAGAAAACCACAC CUGUAAGACAGAAAACCACA CAAGUCCAACAAACACUGU UCUGUCUUACAGUUGUUUGU	755 (AGCG) 1006 (CAG) 1007 (AAG) 1008 (TGG)	20 (C-1) 20 (C1) 20 (n.a.) 20 (n.a.)	VRER-SpBE3 SpBE3 WT Cas9 WT Cas9	

TABLE 7-continued

Cryptic Epitopes						
Antigen Name and Exemplary Condition(s)	Cryptic epitope(s) and Genomic target region	Programmable guide-RNA sequence	SEQ ID NO	gRNA size (C-edited)	Genome Editor type ^a	
	exon3 junction, last base of intron 2					
LAGE-1 (multiple tumors, e.g., melanoma and breast cancer)	MLMAQEALAFL (SEQ ID NO: 811) LAAQERRQVR (SEQ ID NO: 812) APRGVRMVA (SEQ ID NO: 813) QGAMLAQQERR VPRAAEVPR (SEQ ID NO: 814) CLSRRPWKRS WSAGSCPMP HL (SEQ ID NO: 815) tctctgagagccgggc agaggctccggagcc at <u>G</u> caggccgaagg c (SEQ ID NO: 863) target OFR1 start site	CUUCGGCCUGCAUGGCUCGG AUUGCUCCGGAGCCUCUGCC GCCUUUCGGCCUGCAUGGCUC CCUUCGGCCUGCAUGGCUC CAGAGGCUCCCGGAGCAGC GCAGAGGCUCCCGGAGCAGC GCCUUUCGGCCUGCAUGGCUC	756 (GAG) 1009 (CGG) 1010 (CGG) 1011 (GGAG) 1012 (AGG) 1013 (CAG) 1014 (CGG)	20 (C11) 20 (C -1) 20 (C13) 20 (C12) 20 (n.a.) 20 (n.a.) 20 (n.a.)	SpBE3 SpBE3 SpBE3 EQR-SpBE3 WT Cas9 WT Cas9 WT Cas9	
TRP-2 (melanoma)	EVISCKLIKR (SEQ ID NO: 816) TATTCTGTTAG AGATACATTATT <u>A</u> gggggtttttcc (SEQ ID NO: 864) target intron 2 donor site	CCUAUAUAUGUAUCUUAAC ACCUUAUAUAUGUAUCUUA ACCUUAUAUAUGUAUCUUA CACCUAUAUAUGUAUCUUA UUAGAGAUACAUUAUAGGU GUUAGAGAUACAUUAUAGG	757 (AGAA) 1015 (CAG) 1016 (CAGAAT) 1017 (ACAGAAT) 1018 (GGG) 1019 (TCG)	20 (C1/2) 20 (C2/3) 20 (C2/3) 20 (C3/4) 20 (n.a.) 20 (n.a.)	VQR-SpBE3 SpBE3 SaBE3 StIBE3 WT Cas9 WT Cas9	
TRP-2 (melanoma)	EVISCKLIKR (SEQ ID NO: 816) tatgtttccaaattgtttc <u>a</u> gGACCAAGGAC GCCCT (SEQ ID NO: 865) target intron 2 acceptor site	UCCUGGUCCUGAAACAUUG GUCCUGGUCCUGAAACAAU CGUCCUGGUCCUGAAACAAU UUUCCCAAUUGUUUCAGGAC UUUCCCAAUUGUUUCAGGAC	758 (GGAA) 1020 (GGG) 1021 (TGG) 1022 (CAG) 1023 (AGG)	20 (C8/9) 20 (C9/10) 20 (C10/11) 20 (n.a.) 20 (n.a.)	VQR-SpBE3 SpBE3 SpBE3 WT Cas9 WT Cas9	
BIRC5/ Survivin (multiple tumors, e.g., melanoma, breast cancer, and leukemia)	AYACNTSTL (SEQ ID NO: 817) agtggactgcccgttta atccctt <u>C</u> auctgcctt tcgcgtgt (SEQ ID NO: 866) target intron 2 spliceosome branch site	AGCUGAAGGGAUAAAAGCGG GGCAGCUGAAGGGAUAAAAG AAGGCAGCUGAAGGGAUAAA AAAGGCAGCUGAAGGGAUAAA GCAGCUGAAGGGAUAAAAGC GACUGCCGUUUUAUCCCCUU CUGCCGUUUUAUCCCCUCA	759 (CAG) 1024 (CGG) 1025 (AGCG) 1026 (AAG) 1027 (GGCAGT) 1028 (CAG) 1029 (GCT)	20 (C3) 20 (C6) 20 (C8) 20 (C9) 20 (C5) 20 (n.a.) 20 (n.a.)	SpBE3 SpBE3 VRER-SpBE3 SpBE3 KKH-SaBE3 WT Cas9 WT Cas9	
BIRC5/ Survivin (multiple tumors, e.g., melanoma, breast cancer, and leukemia)	AYACNTSTL (SEQ ID NO: 817) tccgctgttggatttt ct <u>a</u> gAGGAAAC ATAA (SEQ ID NO: 867) target intron 2 acceptor site	UCUAGAAAAUAAAACAAC CUCUAGAAAAUAAAACAAC GAACAUAAAAGCAUUCGUC	760 (AGCG) 1030 (CAG) 1031 (CGG)	20 (C2) 20 (C3) 20 (n.a.)	VRER-SpBE3 SpBE3 WT Cas9	

TABLE 7-continued

Cryptic Epitopes					
Antigen Name and Exemplary Condition(s)	Cryptic epitope(s) and Genomic target region	Programmable guide-RNA sequence	SEQ ID NO (PAM)	gRNA size (C-edited)	Genome Editor type ^a
Bcr-Abl-OOF (Leukemia)	SSKALQRPV (SEQ ID NO: 603) GFKQSSKAL (SEQ ID NO: 604) ATGFKQSSKAL QRPVAS (SEQ ID NO: 605) ATGFKQSSKAL QRPVAS (SEQ ID NO: 606) tccccctttcttc gAACCCCTCA GC (SEQ ID NO: 868) target intron 1 acceptor site	UGGAAGAGAAAGGGGGAAAC CUGGAAGAGAAAGGGGGAA GCUUCUGGAAGAGAAAGGG GGCUUCUGGAAGAGAAAGGG AGGGCUUCUGGAAGAGAAAG AAGGGCUUCUGGAAGAGAAA GAAGGGCUUCUGGAAGAGAA UGAAGGGCUUCUGGAAGAGAA UCUGGAAGAGAAAGGGGGAA AGGGCUUCUGGAAGAGAAAG AAGGGCUUCUGGAAGAGAAA GAAGGGCUUCUGGAAGAGAA UUCCCCCUUUCUCCUUCAG CUGUUCCCCCUUUCUUC gc (SEQ ID NO: 868) target intron 1 acceptor site	761 (AGAA) 1032 (CAG) 1033 (GGAA) 1034 (GGG) 1035 (GGG) 1036 (GGG) 1037 (GGG) 1038 (AGG) 1039 (AAG) 1040 (ACAGAAA) 1041 (GGGG) 1042 (GGGG) 1043 (AGGG) 1044 (AAG) 1045 (CAG)	20 (C-1) 20 (C1) 20 (C5) 20 (C6) 20 (C7) 20 (C8) 20 (C9) 20 (C10) 20 (C11) 20 (C2) 20 (C8) 20 (C9) 20 (C10) 20 (n.a.) 20 (n.a.)	VQR-SpBE3 SpBE3 VQR-SpBE3 SpBE3 SpBE3 SpBE3 SpBE3 SpBE3 SpBE3 St1BE3 St3BE3 St3BE3 WT Cas9 WT Cas9
Bcr-Abl-OOF (Leukemia)	SSKALQRPV (SEQ ID NO: 603) GFKQSSKAL (SEQ ID NO: 604) ATGFKQSSKAL QRPVAS (SEQ ID NO: 605) ATGFKQSSKAL QRPVAS (SEQ ID NO: 606) tccccctttcttc AAAAGCTCCGG GTCT (SEQ ID NO: 869) target intron 2 acceptor site	UGAGAAGAAAAGGAACCAA GCUUUUACCUAGAGAAAA AGCUUUUACCUAGAGAAAA GAGCUUUUACCUAGAGAGA UUACCUAGAGAGAAAGGA CCGGAGCUUUUACCUAGAG CCCAGAGCUUUUACCUAGAG ACCGGAGCUUUUACCUAGA CACCGAGAGAAAGGAACC GAUUUGGUUCCUUUCUUC UUCCUUUCCUUCAGGUGAA UGAUUUGGUUCCUUUCUUC tcctttttcttc gtcg gtct (SEQ ID NO: 869) target intron 2 acceptor site	762 (CAG) 1046 (GGAA) 1047 (AGG) 1048 (AAG) 1049 (CCAAT) 1050 (AGA) 1051 (AAG) 1052 (GAAGAA) 1053 (AAATC) 1054 (AGG) 1055 (AAG) 1056 (CAG)	20 (C-1) 20 (C9/10) 20 (C10/11) 20 (C11/12) 20 (C5/6) 20 (C7/8) 20 (C6/7) 20 (C8/9) 20 (C6/7) 20 (n.a.) 20 (n.a.) 20 (n.a.)	SpBE3 VQR-SpBE3 SpBE3 SpBE3 KKH-SaBE3 SpBE3 KKH-SaBE3 St3BE3 WT Cas9 WT Cas9 WT Cas9

^aGenome-editor types abbreviations: SpBE3 = APOBEC1-SpCas9n-UGI; VQR-SpBE3 = APOBEC1-VQR-SpCas9n-UGI; EQR-SpBE3 = APOBEC1-EQR-SpCas9n-UGI; VRER-SpBE3 = APOBEC1-VRER-SpCas9n-UGI; SaBE3 = APOBEC1-SaCas9n-UGI; KKH-SaBE3 = APOBEC1-KKH-SaCas9n-UGI; St3BE3 = APOBEC1-St3Cas9n-UGI; St1BE3 = APOBEC1-St1Cas9n-UGI. Guide sequences (the portion of the guide RNA that targets the nucleobase editor to the target sequence) are provided. The guide sequences may be used with any tracrRNA framework sequences known in the art to generate the full guide RNA sequence.

In some embodiments, the nucleobase editor may be used to introduce a premature stop codon (a stop codon that occurs upstream of the normal stop codon) into a tumor specific antigen gene (e.g., TAA, TAG, and TGA). In some embodiments, introduction of a premature stop codon destabilizes the tumor specific antigen. In some embodiments, destabilization of the tumor specific antigen leads to enhanced presentation of immunogenic epitopes (e.g., heteroclitic epitopes or cryptic epitopes).

Premature stop codons are introduced by changing one or more bases in a target codon that encodes a target residue. For example, nucleobase editors including a cytosine deaminase domain are capable of converting a cytosine (C) base to a thymine (T) base via deamination. Thus, it is envisioned that, for amino acid codons containing a C base, the C base may be converted to T. For example, a CAG (Gln/Q) codon may be changed to a TAG (amber) codon via the deamination of the first C on the coding strand. For sense codons that contain a guanine (G) base, a C base is present on the complementary strand; and the G base may be converted to an adenosine (A) via the deamination of the C on the complementary strand. For example, a TGG (Trp/W) codon may be converted to a TAG (amber) codon via the deami-

nation of the second C on the complementary strand. In some embodiments, two C to T changes are required to convert a codon to a nonsense codon. For example, a CGG (R) codon is converted to a TAG (amber) codon via the deamination of the first C on the coding strand and the deamination of the second C on the complementary strand.

In some embodiments, the target residue is located in a flexible loop region of the tumor specific antigen. In some embodiments, tandem premature stop codons are introduced. Non-limiting examples of codons that may be changed to stop codons via base editing are provided in Table 8.

TABLE 8

Conversion to Stop Codon		
Target codon	Base-editing process	Edited codon
CAG (Gln/Q)	1 st base C to T on coding strand	TAG (amber)
<u>T</u> GG (Trp/W)	2 nd base C to T on complementary strand	<u>T</u> AG (amber)
CGA (Arg/R)	1 st base C to T on coding strand	TGA (opal)
<u>C</u> AA (Gln/Q)	1 st base C to T on coding strand	<u>T</u> AA (ochre)
T <u>G</u> G (Trp/W)	3 rd base C to T on complementary strand	T <u>G</u> A (opal)

TABLE 8-continued

Conversion to Stop Codon		
Target codon	Base-editing process	Edited codon
<u>CGG</u> (Arg/R)	1 st base C to T on coding strand and 2 nd base C to T on complementary strand	<u>TAG</u> (amber)
<u>CGA</u> (Arg/R)	1 st base C to T on coding strand and 2 nd base C to T on complementary strand	<u>TAA</u> (ochre)

^{*}single underline: changes on the coding strand

double underline: changes on the complementary strand

In some embodiments, cryptic epitopes are generated by shifting the coding frame of a tumor specific antigen gene.

In some embodiments, the coding frame is shifted by changing a start codon (ATG) to a sense codon that cannot be used as a start codon. As such, translation will start at the next start codon in the coding region, and the coding frame may be shifted. In some embodiments, a normal sense codon may be edited to generate a start codon to allow translation to start at the newly generated start codon, which may also lead to shifting of the coding frame. Alterations of start codons and the resulting shift in the coding frame generate peptides that would not otherwise be generated from the tumor specific antigen gene (i.e., cryptic epitopes). Non-limiting, exemplary start codon alterations that may be achieved by the nucleobase editors described herein are provided in Table 9.

TABLE 9

Alteration of Start Codons via Base Editing		
Target codon	Base-editing process	Edited codon
Cognate Start codon <u>ATG</u> (Met/M)	3 rd base C to T on complementary strand	ATA (Ile/I, next ATG is used as start codon)
ACG (Thr/T)	2 nd base C to T on coding strand	ATG (Met/M)
<u>G</u> TG (Val/V)	1 st base C to T on complementary strand	<u>A</u> TG (Met/M)
<u>GCG</u> (Ala/A)	1 st base C to T on complementary strand and 2 nd base C to T on coding strand	<u>A</u> TG (Met/M)

The tumor associated antigens listed in Table 5 and Table 6 and their respective gene sequences and protein sequences are known in the art. The amino acid sequence of the listed tumor specific antigens are listed in Table 10.

TABLE 10

Amino acid sequences of human tumor specific antigens		
Name of tumor associated antigen	Amino acid sequence	SEQ ID NO
Melanocyte protein PMEL (gp100)	MDLVKRCLLHLAVIALLAVGATKVPQRNQDWLGVSRQLRTKAWNRNQLYPEW TEAQLRDCWRGGQVSLKVNDGPTLIGANASFSIALNFPGSQKVLPDGQVIW VNNTIINGSQVWGGQPVYQETDDACIFPDGGPCPSGSWSQRSFVWVKWT GQYWQVLGGPVSGLSIGTGRAMLGHTMEVTYHRGRSRSYVPLAHSSAFT ITDQVPSVSQSLRAALDGNKHFLRNQPLTFALQLHDPGSGLAEADLSYT DFGDSSGTLSRALVVTHTYLEPGPVTAQVVLQAAIPLTSCGSSPVPGTTDG HRPTAEAPNTTAGQVPTTEVGTTGGQAPTAEPSTGTTSVQVPITEVISTAPV QMPTAESTGMTPKEPVPVSEVMGTTLAEMSTPEATGMMTPEAEVSIIVLSGTTA QVTTTEWVETTARELPIPEPEGPDASSIMSTESITGSGLGPLDGTTATRLRK RQVPLDCVLYRGFSVTLDIVQGIESAELQAVPSPGEGDAFELTVSCQGGL PKEACMEISSLPGCOPAQLCQPVLPSPACQLVLHQILKGGSCTYCLNVSLA DTNSLAVVSTQLIMPQGEAQLGQVPLIVGILLVLMAVVLASLIYRRRLMKD FSVPQLPHSSSHWLRLPRIFCSCPICGENSPILLSQVQ	763
Melanoma antigen recognized by T-cells 1 (MLANA/M ART-1)	MPREDAHFIYGYPKKGHGHSYTTAEAAAGIGILTIVLGVLILLGCWYCRRRN GYRALMDKSLHVGTQCALTRCPQEGFDHRDSKVSLQEKNCEPVVPNAPPAY EKLSAEQSPPPYSP	764
5, 6-dihydroxyindole-2-carboxylic acid oxidase (TYRP1)	MSAPKLSSLGCIFFPLLLFQQARAQFPQCATVEALRSGMCPDLSPVSGPG TDRCGSSSGRGRCEAVTADSRPHSPQYPHDGRDRREVPLRFFNRTCHCNGN FSGHNCGTCRPGWGRGAAACDQRVLIVRRNLLDLSKEEKNHFVRALDMAKRTTH PLFVIATRSEELIGPDGNTPQFENISIYNYFWTHYYSVKKTFLGVGQESF GEVDFSHEGPFLTWHRYHLLRLEKDMQEMLOQEPSFLPYWNFATGKNCVDI CTDDLMGSRSNFDSTLISPNSVFSQWRVVCDSLEDYDTLGLCNSTEDGPIR RNPAGNVARPMVQRLPEPODVAQCLEVGLFDTPPFYNSNSTNSFRNTVEGYSD PTGKYDPAVRSLHNLAHLFLNGLTGGQTHLSPNDPPIFVLLHTFTDAVEDEWLR RYNADISTPLENAPIGHNRQYNMVPFWPPVNTTEMFVTAPDNLGYTYEIOW PSREFSVPEIIIAIAVVGALLLVALIFGTASYLIRARRSMDEANQPLLTDQYQ CYAEEYEKLQNPNQSVV	765

TABLE 10-continued

Amino acid sequences of human tumor specific antigens		
Name of tumor associated antigen	Amino acid sequence	SEQ ID NO
Alpha-1, 6-mannosyl-glycoprotein 6-beta-N-acetylglucosaminyltransferase A (MGAT5)	MALFTPWLSSQKLGFFLVTFGFIWGMMLLHFTIQQRTOPESSMLREQILDLSKRYIKALAEENRNVVDGPyAGVMTAYDLKKTAVLLDNILQRIGKLESKV DNLLVVNGTGTNSTNSTTAVPSLVALEKINVADINQAQEKCVLPPMDGYPHCEGKIKWKMKDWRSDPCYADYGVDGSTCSFFIYLSEVENWCPhLPWRAKNPYE EADHNSLAEIRTDNFNLYSMMKKHEEFRWMLRIRRMAWIAQIKSLAEKQ NLEKRKRKKVVLHGLLTKESGFKIAETAFSGGPLGELVQWSIDLTSLYLLG HDIRISASLAELKEIMKKVVGNRSGPCPTVGDRIVELIYIDIVGLAQFKKTLG PSWVHQCMRLVLDLSFGTEPEFHANYAQSKGHKTPWGKWNLNPQOFYTMFP HTPDNNFLGFVVEQHLNNSDIHHINEIKRNQQLSVYGVVDSFWKNNKKIYLDI IHTYMEVHATVGGSSTKNIPSYVKNHGILSGRDLQFLLLRETKLTVGLGFPYE GPARPLEIAANGCAFLNPKPNPKSSKNTDFIIGKPTLRELTSQHPYAEVFIG RPHVWTVDLNQSEEVEDAVKAILNQKIEPYMPYEFTECEGMLQRIANAFIEKQD FCHGQVMWPPLSALQVKAEPGQSKCQVCQESQLICEPSFFQHNLKDQMLK YKTCQSSELAKDILVPSFDPKNKHCVFQGDLLLFSCAGAHPHRQRVCPCRDFIKGQVALCKDCL	766
Cancer/testis antigen 1 (CTAG1; LAGE2/NY-ESO-1)	MQAEGRRGTGGSTGDADGPGGPIPDGPGGNAGGPGEAGATGGRGPRGAGAAR ASGPGGGAPRGPHGAASGLNGCCRCGARGPESRLLFYLAMPFATPMEAEL ARRSLAQDAPPVPGVLLKEFTVSGNILTIRLTAADHRQLQLSISSCLQQL SLLMWITQCFLPVFLAQPPSGQRR	767
CTL-recognized antigen on melanoma (CAMEL)	MLMAQEALAFMLAQGAMLAAQERRVPRAAEVPGAQGQQGPRGRREEAPRGVRM AVPLLRRMEEAPAGPGGRTAACFSCTSRLSRPWRKSWSAGSCPMPHLSP DQGRF	768
L-dopachrome tautomerase (DCT)	MSPLWWGFLLSCLGCKILPGAGQQFPRVCMTVDLNVKECCPRLGAESANVC GSQQGQRCQCTEVRADTRPWSPGYILRNQDDRELPWRKFFTRCKCTGNFAGY NCDCDKFPGWTGPNCERKPPVIRQNIHSLSPOEREQFLGALDALKRVHPDY VTTQHWLGLLGPNGTQPQFANCVSVDFFVWLHYYSVRDTLLGPGRPYRAID FSHQGPAPFTWHRYHLLCLERDLQRLIGNESFALPYWNFATGRNECDVCTDQ LFGAARPDDPTLISRNSRFSSWETVCDSDLDDYHNLVTLCMGTYEGLLRRNQM GRNSMKLPTLKDIRDCLSLSQKFDDNPFFQNNTFSFRNALEGEDKADGTLDSQ VMSLHNLVHSFLNGTMALPHSAANDPIFVVLHSFTDAIFDEWMKRFNPPADA WPQELAPIGHNRMYNMVPFFPPVTNEELFLTSQDQLGYSYADLPVSVEETPG WPTTLVVMGTLVALVGLFVLLAFLQYRRLRKGYTPLMETHLSSKRYTEEA	769
Tyrosinase (TYR)	MLLAVLYCLLWSFQTSAQHFPRACVSSKNLMEKECCPPWSGDRSPCGQLSGRGSCQNILLSNAPLGPQPFPTGVDDRESWPSVFYNRTQCQCSGNFMGFCGNCK FGFWGPNCTERLLVRRNIFDLSAPEGKDKFFAYLTLLAKHTISSDVVIPIGTY GQMKNGSTPMENDINIYDLFVWMMHYVSMALLGGEIWRDIDFAHEAPAFL PWHLFLLRWBEQEIQLTGDEFNTIPYWDWRDAEKCDICTDEYMGQQHTNP NLLSPASFFSSWQIVCSRLEEYNSHQSLCNGTPEGLRRNPGNHDKSRTPRL PSSADFECFLSLTQYESGSMKAANFSFRNTLEGFASPLTGIADASQSSMHN ALHIYMNGTMSQVQGSANDPIFIPLLHHAFVDSIFEQWLRRRHPHQEVYPEANA PIGHNRESYMFPIPLYRNGDFIISKDLGYDYSYLDSDPDSFQDYIKSYL EQASRIWSWLLGAAMVGAVLTALLAGLVSLLCRHKRKQLPEEKQPLLMEKED YHSLYQSHL	770
Baculoviral IAP repeat-containing protein 5 (BIRC5)	MGAPTLPPAWQPFLKDHRISTFKNWPFLLEGCACTPERMAEAGFIHCPTENEPLAQCFFCFKELEGWEPDDDPIEEHKKHSSGCAFLSVKKQFEELTGEFLKL DRERAKNKIAKETNNKKKEFEETAKKVRRAIEQLAAMD	771
Receptor tyrosine-protein kinase erbB-2 (ERBB2/HER 2)	MELAALCRWGLLLALLPPGAASTQVCTGTDMKLRLPASPETHLDMLRHLYQGCQVQGNLELTLYLPTNASLFLQDIQEVOQGYVILIAHNQVRQVPLQRLRIVRG TQLFEDNYALAVLDNGDPLNNTPVGTGASPGLRELQLRSLTEILKGGLIQRNPQLCYQDTIILWKDIFHKNQNLALTLDINTNRSRACHCPSPMCCKGSRCWGES SEDCQSLTRTVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLHENH SGICELCEKPAVLVNTDTFESMPNPEGRYTFGASCVTACPYNLLSTDVGSCTLVCPPLHNQEVTAEDGTCRCEKCSKPCARVCYGLGMELHREVRAVTSANIQEF AGCKKIFGSLAFLPSEGDGPASNTAPLQPEQLOVQFETLEEITGYLYISAWPDSLPDFSVFQNLQVIRGRILHNGAYSLTLQGLGIWLGLRSLRELGSGLALIHINTHLCFVHTPVWDQFLFRNPHQALLHTANRPEDECVGEGLACHQLCARGHCGWGPGBTQCVNCSQFLRGQECVEECRVLQGLPREYVNRHCLPCHPECQPNQGSVTCFGPEADQCVACAHYDPPFCVARCPSPGVKPDLSYMPIWKFPDEEGACQPCPINCTHSCVLDLDDKGCPABQRASPLTSII SAVVGILLVVLGVVFGILIKRROQKIRKYTMRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKVLGSYAFGTVYKGWI PDGENVKIPVAIKVLRRENTSPKANKEILDEAYVMAGVGSPYVSRLLGICLTSVQLVTQLMYGCCLLDHVRENRGRLGSQDLNWCMQIAKG	772

TABLE 10-continued

Amino acid sequences of human tumor specific antigens		
Name of tumor associated antigen	Amino acid sequence	SEQ ID NO
	MSYLEDVRLVHRLDLAARNVLVKSPNHSVITDFGLARLLDIDEDETEYHADGGKVPIKWMALESIILRRRFTHQSDVWSYGTVWELMTFGAKPYDGI PAREIPDLLEKGERLPOPPICTIDVYMIMVKCWMIDSECRPRFRELVSSEFSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSILLEDDDMGDLVDAEYLVPPQQGFCPDPAVGAGGMVHHHRHSSSTSRSGGDLTGLGPLEPSEEAAAPRSLAPSEGAGSVDGEDGLGMAAKGLQSPLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNQPDVRPQPPSPREGPLPAARPAGATLERPKTLSPGKNGVVKDVFAFGGAVENPEYLTPQGGAAPQPHPPPAFSPAIDNLYYWDQDP PERGSTFKGTPTAENPEYLGDDPV	
Myeloid cell surface antigen CD33	MPLLLLLPLLWAGALAMDPNFWLQVQESVTQEGLCVLVPCTFFHIPYYDKNSPVHGWFREGAIISRDSPVATNKLDQEVEQTQGRFRLLGDPNSRNCSLISDARRRDNGSYFFRMRERGSTKYSYKSPQLSVHVTDLTHRPKLLIPGTLLEPGHSKNLTCWSVSWACEQGTPPIFSWLSAAPTSLGPGRITHSSVLIITPRPQDHGTNLTCQVKFAGAVGVTTERTIQLNVTYVPQNPTTGIFPPGDGSGKQETRAGVHGAIGGAGVTALLALCLCLIFFIVKTHRKAARTAVGRNDTHPTGSASPKHQKKSLHGPETSSCSGAAPTVEMDEELHYASLNFHGMNPSKDTSTEYSEVRTQ	773
Telomerase reverse transcriptase (TERT)	MPRAPRCAVRSLLRSHREVLPLATFVRLGPQGWRLVQRGDPAFRALVAQCLCVCPWDAPPAAFSRQVSCLKELVARVLQRLCERGAKNVLAFGFALLDARGGGPPEAFTSVRSYLPNVTDALRGSGAWGLLRLRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGAATQARPPPHASGPRRRLGCERAWNHSVERAGVPLGLPAPCRGGSAARSLLPLPKRPRGAAPEPERTPVQGGSWAHPGRTGRPSDRGFVUVSPARPAEEATSLLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPWDTCPCPVVAETKHFLYSSGDKEQLRPSFLSSLRPSLTGARRIVETIPLGSRPWPMPGTPPRLPRLPQRYWQMRPLFLELLGNHAQCPCYGVLLKTHCPLRAAVTPAAGVCAKEKPQGSVAPEEEEDTDPRRLVQLLRQHS SPWQVYGFVRACLRLVPPGLWGSRHNRFLRNTKKFISLGKHAKLSQLELTWKMSVRDCAWLRRSPGVGCVPAEEHRLREEIILAKFLHWLMSVYVELLRSFFYVTETTFQKNRLFYRKSWSKLOSIGIRQHLKRVQLRELSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVMDYVVGARTPREFKERAERLTSRVKALFSQLNYERARRPGLLGASVGLDDDIHRAWRTFVLVRRAQDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPNQTYCVRYYAVVQKAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPRLDAVVIQSSSLNEASSGLEDFVFLRFMCHHAVRIRGKSYVQCOGIPQGSILSTLLCGMDENKLFAGIRRDLGLLRLVDDFLLLVTPLHTAKTFLRTLVRGVPEYCCVNULRKTUVNFPVDEALGGTAFVQMPAHGLFWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGRNMRRKLFGVLRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRHFACVQLPFLHQQWKNPTFFLRLVISDTASLCYSILKAKNACMSLGAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTVPLLGSLRATAQTLQSRKLPGTTLTAANPALPSDFKTIID	774
Protein SSX2	MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEEWEMKMASEKIFYVYMKRKYEAMTKLGFKATLPPFMCNKRAEDFQGNDLDNDPNRGNQVERPQMTFGRLQGISPKIMPKPAEEGNDSEEVPEASGPQNDGKELCPPGKPTTSEKIHERSGPKRGEHAWTHRLRERKQLVIYEIEISDPEEDE	775
Wilms tumor protein (WT1)	MGSVDRLNALLPAVPVSLGGGGCALPVSGAAQWAVPLDFAPPGASAYGSLGGPAPPAPPAPPAPPFSFIQKEPSWGAEPHEEQLSAFTVHSGQFTGTA GACRYGPFGPPPSQASSGQARMFPNAPYLPSCLESQPAIRNQGYSTVTEDGTPSYGHTPSHHAQFPNHSFKHEDPMGQQGSLGEEQQYSVPVVGCHPTDSCTGSQALLRRTPYSSDNLYQMTSQLECMTNQMNLGATLKGVAAGSSSSVVKRTEGQSNHSTGYEDNHTTPILCGAQYRIRHTHGVVERGIQDVRVPGVAPTLVR SASETSEKRPFMCAYPGCNKRYFKLSHLQMHSRKHTGEKPYQCDFKDERRFSRSQDQLKRHQRRTGKPFQCKTCQRKFSRSRSHLKTHTRTHTGKTEKPFSCRWPSCKKFARSDELVRHHNMQRNMTKLQLAL	776
Sperm-associated antigen 11A (SPAG11A)	MQRLLPSVTSLLLVALFPFGSSQARHVNHSAEALGELRERAPGQGINGFQLLRHAHKRDLPPRTPPYQVHISHQEARGPSFKICVGFLGPRWARGCSTGNEKYHLPYAAARDLQTFFLPFW	777
BCR/ABL fusion protein isoform X3	MVDPVGFAEAWAKAQFPDSEPPRMEMLRSVGDIEQELERCKASIRRLEQEVNQERFRMIYLQTLAKEKKSYDRQRWGFRRAAQAPDGASEPRASASRPQPAADGADPPPAEEPEARPDGEKGSPKGARPGTARRPGAAASGERDDRGPBPASVALRSNFERIKGHGQPGADAEPKPYVNVEFHHERGLVKVNDKEVSDRISSLGSQAMQMERKKSQHGAGSSVGDASRPPYGRSSSESSCGVDGDYEDAELNPRFLKDNLIDANGGSRPPWPPLPEYQPYQSIYVGGMMEGEGKGPLLRSQSTSEQEKRLTWPRRSYSPRSFEDCGGGYTPDCSSNENLTSSEEDFSSGQSSRVSPSPPTTYMERDKRSRSPQNSQSFDSSSPPTPQCHKRHRCPVVSEATIVGVRKTGQIWPNDGEGAFHGDAGSGFTPPGTYCAADRAEEQRHQDGLPYIDDSPSSPHLSSKGRGSRDAVLSGALESTKASELDLEKGLEMRKWVLSGILASETYLSHREALLLPMKPLKAAATTSQPVLTTSQQIETIFFKVPELYEIHKEFYDGLFPRVQOWSHQQRVGDLFQKLASQLGVYRVLGYNHNGEWCEAQTKNGQGWVPSNYITPVNS	778

TABLE 10-continued

TABLE 10-continued

Amino acid sequences of human tumor specific antigens		
Name of tumor associated antigen	Amino acid sequence	SEQ ID NO
precursor (EpCAM)	YDSKSLRTALQKEITTRYQLDPKFITSILYENNVITIDLVQNNSQKTQNDVD IADVAYYFEKDVKGESLFHSKMDLTVNGEQLLDPGQTLIYYVDEKAPEFS MQGLKAGVIAVIVVVVIAVVAGIVVLVISRKRMAYEAEIKEMGEMHREL NA	

In some embodiments, cancer vaccines containing immunogenic peptides from tumor specific antigens (e.g., heteroclitic epitopes and cryptic epitopes) are generated in vivo (e.g., in tumor cells in a subject) or ex vivo (e.g., in tumor cells isolated from a subject). In some embodiments, the tumor cells are treated with the nucleobase editors to generate the immunogenic peptides and are irradiated and administered to the subject as whole-cell cancer vaccines.

To edit the genes encoding the tumor associated antigens, the nucleobase editor and/or the guide nucleotide sequence is introduced into the cell (e.g., a tumor cell) where the editing occurs. In some embodiments, nucleic acid molecules (e.g., expression vectors) encoding the nucleobase editors and/or the guide nucleotide sequences are delivered into the cell, resulting in co-expression of nucleobase editors and/or the guide nucleotide sequences in the cell. The nucleic acid molecules encoding the nucleobase editors and/or the guide nucleotide sequences may be delivered into the cell using any known methods in the art, e.g., transfection (e.g., transfection mediated by cationic liposomes), transduction (e.g., via viral infection) and electroporation. In some embodiments, an isolated nucleobase editor/gRNA complex is delivered. Methods of delivering an isolated protein to a cell is familiar to those skilled in the art. For example, the isolated nucleobase editor in complex with a gRNA be associated with a supercharged, cell-penetrating protein or peptide, which facilitates its entry into a cell (e.g., as described in PCT Application Publication WO2010129023 and US Patent Application Publication US20150071906, incorporated herein by reference). In some embodiments, the isolated nucleobase editor in complex with a gRNA may be delivered by a cationic transfection reagent, e.g., the Lipofectamine CRISPRMAX Cas9 Transfection Reagent from ThermoFisher Scientific. In some embodiments, the nucleobase editor and the gRNA may be delivered separately. Other suitable delivery methods may also be used, e.g., AAV mediated gene transfer. Strategies for delivery a Cas9-based genome editing agent (e.g., the nucleobase editor described herein) using AAV have been described, e.g., in Zetsche et al., *Nature Biotechnology* 33, 139-142 (2015), incorporated herein by reference.

In some embodiments, once generated, the immunogenic peptide is displayed on the surface of the tumor cell via the MHC class I antigen presentation pathway. In some embodiments, the immunogenic peptide is displayed on the surface of an antigen presenting cell (APC) via the MHC class II antigen presentation pathway. In some embodiments, the APC is selected from the group consisting of: tumor cells, dendritic cells, mononuclear phagocytes, thymic epithelial cells, and B cells. In some embodiments, the immunogenic peptide elicits an adaptive immune response against the tumor-specific antigen where the peptide is derived from. In some embodiments, the immunogenic peptide elicits an adaptive immune response against the tumor. In some

embodiments, the adaptive immune response comprises promoting the maturation of dendritic cells, activation of CD4+ T lymphocytes, (T helper cells) activation of CD8+ T lymphocytes (cytotoxic T cells), activation and maturation of B lymphocytes, and/or production of tumor antigen-specific antibodies.

T helper cells (TH cells) assist other white blood cells in immunologic processes, including maturation of B cells into plasma cells and memory B cells, and activation of cytotoxic T cells and macrophages. These cells are also known as CD4+ T cells because they express the CD4 glycoprotein on their surfaces. Helper T cells become activated when they are presented with peptide antigens by MHC class II molecules, which are expressed on the surface of antigen-presenting cells (APCs). Once activated, they divide rapidly and secrete small proteins called cytokines that regulate or assist in the active immune response. These cells can differentiate into one of several subtypes, including TH1, TH2, TH3, TH17, TH9, or TFH, which secrete different cytokines to facilitate different types of immune responses. Signaling from the APC directs T cells into particular subtypes.

Cytotoxic T cells (e.g., TC cells, CTLs, T-killer cells, killer T cells) destroy virus-infected cells and tumor cells, and are also implicated in transplant rejection. These cells are also known as CD8+ T cells since they express the CD8 glycoprotein at their surfaces. These cells recognize their targets by binding to antigen associated with MHC class I molecules, which are present on the surface of all nucleated cells. Through IL-10, adenosine, and other molecules secreted by regulatory T cells, the CD8+ cells can be inactivated to an amergic state, which prevents autoimmune diseases.

Most cytotoxic T cells express T-cell receptors (TCRs) that can recognize a specific antigen. Antigens inside a cell are bound to class I MHC molecules, and brought to the surface of the cell by the class I MHC molecule, where they can be recognized by the T cell. If the TCR is specific for that antigen, it binds to the complex of the class I MHC molecule and the antigen, and the T cell destroys the cell, e.g., via inducing apoptosis. In order for the TCR to bind to the class I MHC molecule, the former must be accompanied by a glycoprotein called CD8, which binds to the constant portion of the class I MHC molecule. Therefore, these T cells are called CD8+ T cells.

Natural killer T cells (NKT cells—not to be confused with natural killer cells of the innate immune system) bridge the adaptive immune system with the innate immune system. Unlike conventional T cells that recognize peptide antigens presented by major histocompatibility complex (MHC) molecules, NKT cells recognize glycolipid antigen presented by a molecule called CD1d. Once activated, these cells can perform functions ascribed to both Th and Tc cells (i.e., cytokine production and release of cytolytic/cell killing

molecules). They are also able to recognize and eliminate some tumor cells and cells infected with microorganisms, e.g., bacteria or virus.

Memory T cells are a subset of antigen-specific T cells that persist long-term after an initial T cell response. They quickly expand to large numbers of effector T cells upon re-exposure to their cognate antigen, thus providing the immune system with “memory” against past antigens. The cancer vaccine described herein provides the immune system with “memory” against the tumor specific antigen, thereby eliciting strong immune response against newly emerged cancer cells or metastasized cancer cells.

Regulatory T cells (suppressor T cells) are crucial for the maintenance of immunological tolerance. Their major role is to shut down T cell-mediated immunity toward the end of an immune reaction and to suppress autoreactive T cells that escaped the process of negative selection in the thymus. Suppressor T cells along with Helper T cells can collectively be called Regulatory T cells due to their regulatory functions.

B cell activation occurs in the secondary lymphoid organs (SLOs), such as the spleen and lymph nodes. After B cells mature in the bone marrow, they migrate through the blood to SLOs, which receive a constant supply of antigen through circulating lymph. At the SLO, B cell activation begins when the B cell binds to an antigen via its BCR. The antigen can either be free-floating or presented by APCs such as macrophages or dendritic cells (DCs), and include proteins, glycoproteins, polysaccharides, whole virus particles, and whole bacterial cells. Some subtypes of B cell preferentially undergo T cell-dependent activation while other subtypes of cells preferentially undergo T cell-independent activation.

Antigens that activate B cells with the help of T-cell are known as T cell-dependent (TD) antigens and include foreign proteins. They are named as such because they are unable to induce a humoral response in organisms that lack T cells. B cell response to these antigens takes multiple days, though antibodies generated have a higher affinity and are more functionally versatile than those generated from T cell-independent activation.

Once a BCR binds a TD antigen, the antigen is taken up into the B cell through receptor-mediated endocytosis, degraded, and presented to T cells as peptide pieces in complex with MHC-II molecules on the cell membrane. T helper (TH) cells, typically follicular T helper (TFH) cells, that were activated with the same antigen recognize and bind these MHC-II-peptide complexes through their T cell receptor (TCR). Following TCR-MHC-II-peptide binding, T cells express the surface protein CD40L as well as cytokines such as IL-4 and IL-21. CD40L serves as a necessary co-stimulatory factor for B cell activation by binding the B cell surface receptor CD40, which promotes B cell proliferation, immunoglobulin class switching, and somatic hypermutation as well as sustains T cell growth and differentiation. T cell-derived cytokines bound by B cell cytokine receptors also promote B cell proliferation, immunoglobulin class switching, and somatic hypermutation as well as guide differentiation. After B cells receive these signals, they are considered activated.

Activated B cells participate in a two-step differentiation process that yields both short-lived plasmablasts for immediate protection and long-lived plasma cells and memory B cells for persistent protection. The first step, known as the extrafollicular response, occurs outside of lymphoid follicles but still in the SLO. During this step activated B cells proliferate, may undergo immunoglobulin class switching, and differentiate into plasmablasts that produce early, weak

antibodies mostly of class IgM. The second step consists of activated B cells entering a lymphoid follicle and forming a germinal center (GC), which is a specialized microenvironment where B cells undergo extensive proliferation, immunoglobulin class switching, and affinity maturation directed by somatic hypermutation. These processes are facilitated by TFH cells within the GC and generate both high-affinity memory B cells and long-lived plasma cells. Resultant plasma cells secrete large amounts of antibody and either stay within the SLO or, more preferentially, migrate to bone marrow.

Antigens that activate B cells without T cell help are known as T cell-independent (TI) antigens and include foreign polysaccharides and unmethylated CpG DNA. They are named as such because they are able to induce a humoral response in organisms that lack T cells. B cell response to these antigens is rapid, though antibodies generated tend to have lower affinity and are less functionally versatile than those generated from T cell-dependent activation.

As with TD antigens, B cells activated by TI antigens need additional signals to complete activation, but instead of receiving them from T cells, they are provided either by recognition and binding of a common microbial constituent to toll-like receptors (TLRs) or by extensive crosslinking of BCRs to repeated epitopes on a bacterial cell. B cells activated by TI antigens go on to proliferate outside of lymphoid follicles but still in SLOs (GCs do not form), possibly undergo immunoglobulin class switching, and differentiate into short-lived plasmablasts that produce early, weak antibodies mostly of class IgM, but also some populations of long-lived plasma cells.

Memory B cell activation begins with the detection and binding of their target antigen, which is shared by their parent B cell. Some memory B cells can be activated without T cell help, such as certain virus-specific memory B cells, but others need T cell help. Upon antigen binding, the memory B cell takes up the antigen through receptor-mediated endocytosis, degrades it, and presents it to T cells as peptide pieces in complex with MHC-II molecules on the cell membrane. Memory T helper (TH) cells, typically memory follicular T helper (TFH) cells, that were derived from T cells activated with the same antigen recognize and bind these MHC-II-peptide complexes through their TCR. Following TCR-MHC-II-peptide binding and the relay of other signals from the memory TFH cell, the memory B cell is activated and differentiates either into plasmablasts and plasma cells via an extrafollicular response or enter a germinal center reaction where they generate plasma cells and more memory B cells.

In some embodiments, the adaptive immune response results in the killing tumor cells, reducing tumor burden, reducing tumor size, and/or preventing metastasis. In some embodiments, the adaptive immune response is active against neo-epitopes associated with spontaneous somatic mutations. In some embodiments, the neo-epitope is specific to the lineage of tumor cells.

In some embodiments, the adaptive immune response elicited by the heteroclitic or cryptic epitopes is cross-reactive with the native tumor-specific antigen. In some embodiments, the adaptive immune response elicited by the heteroclitic or cryptic epitopes is cross-reactive with neoepitopes arising from spontaneous mutations occurring in the tumor specific antigen.

There are advantages associated with using heteroclitic epitopes in clinical applications. For example, heteroclitic epitopes have the ability to break/overcome tolerance by reversing a state of T cell anergy, activating non-tolerized

cross-reactive clones of T cells, or by mediating “immune deviation,” i.e., the type of CTL produced, such as Th1 or Th2. Recent studies indicate that heteroclitic epitopes are immunogenic (Zaremba, et al., *Cancer Research*, 57:4570 (1997); Rivoltini, et al., *Cancer Research*, 59:301 (1999); Selby, et al., *The Journal of Immunology* 162(2):669 (1999), the entire contents of each of which are incorporated herein by reference) in that they are capable of inducing CTLs that recognize endogenously processed epitopes. This is confirmed by studies in different immunological systems (Zugel, et al., *J. Immunol.*, 161:1705 (1998), Wang, et al., *J. Exp. Med.*, 190:983 (1999), Men, et al., *J. Immunol.*, 162: 3566, (1999), the entire contents of each of which are incorporated herein by reference). For example, studies by Zugel et al. have shown that T cell tolerance to an immunodominant T cell epitope in adult mice can be overcome by immunization with heteroclitic cross-reactive peptide analogs of that peptide.

In some embodiments, heteroclitic epitopes or cryptic epitopes modulate cytokine production from T cells (Pfeiffer, et al., *J. Exp. Med.*, 181:1569 (1995), Tao, et al., *J. Immunol.*, 158:4237 (1997), Salazar, et al., *Int. J. Cancer* 85(6):829-38 (2000), Nicholson, et al., *Int. Immunol.* 12(2): 205-13 (2000), the entire contents of each of which are incorporated herein by reference). The immune deviation induced by such analogs has implications in several disease states, where generation of a specific subset of Th cell responses correlate with tumor regression (Zitvogel, et al., *J. Exp. Med.*, 183:87 (1996), Celluzzi, et al., *J. Exp. Med.* 183:283 (1996), the entire contents of each of which are incorporated herein by reference) or affected the clinical outcome of autoimmune or infectious disease (Romagnani, et al., *Annu. Rev. Immunol.*, 12:227-57 (1994), the entire contents of which are incorporated herein by reference). Thus, immunization with heteroclitic epitopes offers the capacity to modulate cytokine production by induction of specific subsets of effector T cells, thereby altering the course of disease.

In some embodiments, heteroclitic epitopes offer an advantage in drug development since significantly smaller amounts of peptide are needed for treatment doses, due to their strong biological potency. This feature overcomes certain manufacturing and toxicity concerns. In this regard, it has been shown that a heteroclitic analog of a MART-1 peptide (Rivoltini, et al., *Cancer Research* 59:301 (1999), the entire contents of which are incorporated herein by reference), which generated antigen specific T cells in melanoma patients, was active at much lower concentrations than the native epitope. Similar results were reported by Schlom and colleagues (Zaremba, et al., *Cancer Research* 57:4570 (1997), the entire contents of which are incorporated herein by reference) regarding heteroclitic analog of the CEA derived CAP1 epitope.

Nucleobase Editors

The methods of generating immunogenic peptides or epitopes from tumor associated antigens as cancer vaccines described herein, are enabled by the use of the nucleobase editors. As described herein, a nucleobase editor is a fusion protein comprising: (i) a programmable DNA binding protein domain; and (ii) a deaminase domain. It is to be understood that any programmable DNA binding domain may be used in the based editors.

In some embodiments, the programmable DNA binding protein domain comprises the DNA binding domain of a zinc finger nuclease (ZFN) or a transcription activator-like effector domain (TALE). In some embodiments, the programmable DNA binding protein domain may be pro-

grammed by a guide nucleotide sequence, and is thus referred as a “guide nucleotide sequence-programmable DNA binding-protein domain.” In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a nuclease inactive Cas9, or dCas9. A dCas9, as used herein, encompasses a Cas9 that is completely inactive in its nuclease activity, or partially inactive in its nuclease activity (e.g., a Cas9 nickase). Thus, in some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a Cas9 nickase. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a nuclease inactive Cpf1. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a nuclease inactive Argonaute. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a nuclease inactive CasX or CasY, e.g., as described in Burstein et al., New CRISPR-Cas systems from uncultivated microbes, *Nature* 542, 237-241, 2017, incorporated herein by reference.

In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a dCas9 domain. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a Cas9 nickase. In some embodiments, the dCas9 domain comprises the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 3. In some embodiments, the dCas9 domain comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the Cas9 domains provided herein (e.g., SEQ ID NOs: 11-260), and comprises mutations corresponding to D10X (X is any amino acid except for D) and/or H840X (X is any amino acid except for H) in SEQ ID NO: 1. In some embodiments, the dCas9 domain comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the Cas9 domains provided herein (e.g., SEQ ID NOs: 11-260), and comprises mutations corresponding to D10A and/or H840A in SEQ ID NO: 1. In some embodiments, the Cas9 nickase comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the Cas9 domains provided herein (e.g., SEQ ID NOs: 11-260), and comprises mutations corresponding to D10X (X is any amino acid except for D) in SEQ ID NO: 1 and a histidine at a position correspond to position 840 in SEQ ID NO: 1. In some embodiments, the Cas9 nickase comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the Cas9 domains provided herein (e.g., SEQ ID NOs: 11-260), and comprises mutations corresponding to D10A in SEQ ID NO: 1 and a histidine at a position correspond to position 840 in SEQ ID NO: 1. In some embodiments, variants or homologues of dCas9 or Cas9 nickase (e.g., variants of SEQ ID NO: 2 or SEQ ID NO: 3, respectively) are provided which are at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 98%

identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% to SEQ ID NO: 2 or SEQ ID NO: 3, respectively, and comprises mutations corresponding to D10A and/or H840A in SEQ ID NO: 1. In some embodiments, variants of Cas9 (e.g., variants of SEQ ID NO: 2) are provided having amino acid sequences which are shorter, or longer than SEQ ID NO: 2, by about 5 amino acids, by about 10 amino acids, by about 15 amino acids, by about 20 amino acids, by about 25 amino acids, by about 30 amino acids, by about 40 amino acids, by about 50 amino acids, by about 75 amino acids, by about 100 amino acids, or more, provided that the dCas9 variants comprise mutations corresponding to D10A and/or H840A in SEQ ID NO: 1. In some embodiments, variants of Cas9 nickase (e.g., variants of SEQ ID NO: 3) are provided having amino acid sequences which are shorter, or longer than SEQ ID NO: 3, by about 5 amino acids, by about 10 amino acids, by about 15 amino acids, by about 20 amino acids, by about 25 amino acids, by about 30 amino acids, by about 40 amino acids, by about 50 amino acids, by about 75 amino acids, by about 100 amino acids, or more, provided that the dCas9 variants comprise mutations corresponding to D10A and comprises a histidine at a position corresponding to position 840 in SEQ ID NO: 1.

Additional suitable nuclease-inactive dCas9 domains will be apparent to those of skill in the art based on this disclosure and knowledge in the field, and are within the scope of this disclosure. Such additional exemplary suitable nuclease-inactive Cas9 domains include, but are not limited to, D10A/H840A, D10A/D839A/H840A, D10A/D839A/H840A/N863A mutant domains (See, e.g., Prashant et al., *Nature Biotechnology*. 2013; 31(9): 833-838, which are

incorporated herein by reference), or K603R (See, e.g., Chavez et al., *Nature Methods* 12, 326-328, 2015, which is incorporated herein by reference).

In some embodiments, the nucleobase editors utilized in the present invention comprise a Cas9 domain with decreased electrostatic interactions between the Cas9 domain and a sugar-phosphate backbone of a DNA, as compared to a wild-type Cas9 domain. In some embodiments, a Cas9 domain comprises one or more mutations that decreases the association between the Cas9 domain and a sugar-phosphate backbone of a DNA. In some embodiments, the nucleobase editors described herein comprises a dCas9 (e.g., with D10A and H840A mutations) or a Cas9 nickase (e.g., with D10A mutation), wherein the dCas9 or the Cas9 nickase further comprises one or more of a N497X, R661X, Q695X, and/or Q926X mutation of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 11-260, wherein X is any amino acid. In some embodiments, the nucleobase editors described herein comprises a dCas9 (e.g., with D10A and H840A mutations) or a Cas9 nickase (e.g., with D10A mutation), wherein the dCas9 or the Cas9 nickase further comprises one or more of a N497A, R661A, Q695A, and/or Q926A mutation of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 11-260. In some embodiments, the dCas9 domain (e.g., of any of the nucleobase editors provided herein) comprises the amino acid sequence as set forth in any one of SEQ ID NOs: 2-9. In some embodiments, the nucleobase editor comprises the amino acid sequence as set forth in any one of SEQ ID NOs: 293-302 and 321.

Cas9 variant with decreased electrostatic interactions between the Cas9 and DNA backbone

DKKYSIGLAIGNTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAATRLKRTARR
RYTRKRNRICYLQEIIFSNEMAKVDSFFHRLLEESFLVEEDKKHERHPFIFGNIVDEVAYHEKYPTIYHLRK
KLVDSTDKDADRILYILALAHMIPFRGHFLIEGDLNPNDNSVDFKLFIQLVQTYNQLEENPINASGVDSL
AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSGLTPNFKNSNFDLAEDAKLQLSKDTYDDDNLL
AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEMDGTEELLLVKLNREDLLRQKORTFDNGSIPHQIHLGE
LHAILRRQEDFYPFLKDDNREKIEKILTFPLPYYVGPLARGNSRFAMTRKSEETITPWNFEEVVDKGASA
QSFIERMTAFDKNLNEKVLPKHSLLYEFTVYNELTKVVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTFE
DRENIEERLKTYAHLFDDKVMKQLKRRRYTGWGALSRKLINGIRDQSKILDFLSGDFANRNFM
ALIHDDSLTFEDIQKAVSGQGDSLSHEHIANLAGSPAIKGGIQTQVVDELVKMGRHKPENVIEM
ARENOTQGQKNSRERMKRIEEGIKELGQILEPHVENTOLNEQLYLYLYQNGRDMYDVQELDIN
RLSDYDVHIVVPQFSLKDSDIDNKVLTRSDKNRGKDNVPEEEVVKKMNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKUGFIKQLVETRAIKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFR
KDFQFYKVREINNNYIHADHAYLNAVGTALIKYPKLEEEFVYGDYKVVDVRKMIAEQEIGKATA
KYFFYSNIMNNFFKTEITLANGEIRKRPLIETNGETEGTVWDKGRDFATVRKVLSMPQVNIVKTEVTG
GFSKESILPKRNSDKLIARKKDWDPKYGGFDSPTVASVLVVAKEGKSKLKSVKELLGITMERS
SFEKPIDPFLEAKGYKEVKDLIIKLPKYSLFEENGRKRMLASEGELQKGNELALPKYVNFLYLASH
YEKLKGSPEDENQKLFVEQKHYLDANDKVLSAYNKHRDKIREQAENIIH
LFLTLTNGAPAAFKYFDTIDRKETSTKEVLDTLFEDREMEEERLKTYAHLFDDKVMQLKRRRYTGW
ALSRKLINGIRDQGKTILDFLKSGFNMALIHDDSLTFKEDIQKAQSGQDSLHEHIANLAGS

(SEQ ID NO: 9, mutations relative to SEQ ID NO: 1 are bolded and underlined)

High fidelity nucleobase editor

(SEQ ID NO: 321)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNKHVENVNFIEKF
TTERYFCPNTRCSITWFLSWSPCGECRAITEFLSRYPHVTLIYIARLYHHADPRNRQGLRDLISSGVTI
QIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRYVLECILGLPPCNILRRKQPQLTFFTIALQ
SCHYQRLLPPHIWTGLKGSETPGTSEATPDKYIGSILAGINTSVGWAITDEYKVPSKKFKVLG
NTDRHSIKKNLIGALLFDSGETAATRLKRTARRYTRRKNRICYLQEIFNSNEAKVDDSFHRLLEESFL
VEEDKKHERHPIFGNIVDEVAEYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKPRGHFLIEGDLN
PDNSDVDKLFIQLVQTYNQLFEENPIASGVDAKAILSARLSRLENLIQLPGEKNGLFGNLIALS
LGLTPNFKNSNFDLAEAKLQLSKDTYDDLDNLLQAQIGDQYADLFLAKNLSDAILLSDILRVNTEITK
APLSAMIKRYDEHQDQLTLKALVRQQLPEKYEIFFDQSKNGYAGYIDGGAQEEFYKFIKPIEKM
DGTEELVKLNREDLLRKQRTFDNGSIPQHQIHLGEHAILRQEDFPYFLKDNREKIEKILTFRIPYVGP
LARGNSRFAMTRKSEEETITPWNFEVVDKGASQSFIERMTAFDKNLPNEKVLPKHSLLYEYFTVN
ELTKVYTEGMRKPAFLSGEQKKAIVDLFKTNRKVQLKEDYFKKIECFDSVEISGVEDRFNASL
ALSRKLINGIRDQGKTILDFLKSGFNMALIHDDSLTFKEDIQKAQSGQDSLHEHIANLAGS

- continued

```

PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH
PVENTQLQNEKLYLQLQNGRDMYVDQELDINRLSDYDVEDHIVPQSFLKDDSIDNKVLTRSDKNRGKS
DNVPSEEVKKMKNYWROLNNAKLITQRKFNDNLTKAERGGLSLEDLKAGFIKRQLVETRAITKHVAQIL
DSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAAHDAYLNAVVTALIKKYPK
LESEEFVYGDYKVYDVVRKMIAKSEQEIGKATAKYFFYSNIMMFVFTETILANGETRKRPLIETNGETGEIV
WDKGRDFATVRKVLSMPQVNIVKKTETQGGFSKESILPKRNSDKLIARKKDWPDKKYGGFDSPPTVA
YSVLVVAKVEKGKSKKLKSVKELLGITMERSSFEKNPIDFLEAKGYKEVKKDLLIKLPKYSLFLENGR
KRMLASAGELQKGNEALPSKVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR
VILADANLDKVLSAYNKHRSKPIREQAENIHFTLTNLGAPAAFKYFDTTIDRKRTSTKEVLDATLH
QSTTGLYETRIDLSQLGGD

```

In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a single effector of a microbial CRISPR-Cas system. Single effectors of microbial CRISPR-Cas systems include, without limitation, Cas9, Cpf1, C2c1, C2c2, and C2c3. Typically, microbial CRISPR-Cas systems are divided into Class 1 and Class 2 systems. Class 1 systems have multisubunit effector complexes, while Class 2 systems have a single protein effector. Cas9 and Cpf1 are Class 2 effectors. In addition to Cas9 and Cpf1, three distinct Class 2 CRISPR-Cas systems (C2c1, C2c2, and C2c3) have been described by Shmakov et al., "Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems", *Mol. Cell*, 2015 Nov. 5; 60(3): 385-397, the entire contents of which are herein incorporated by reference. Effectors of two of the systems, C2c1 and C2c3, contain RuvC-like endonuclease domains related to Cpf1. A third system, C2c2 contains an effector with two predicted HEPN RNase domains. Production of mature CRISPR RNA is tracrRNA-independent, unlike production of CRISPR RNA by C2c1. C2c1 depends on both CRISPR RNA and tracrRNA for DNA cleavage. Bacterial C2c2 has been shown to possess a unique RNase activity for CRISPR RNA maturation distinct from its RNA-activated single-stranded RNA degradation activity. These RNase functions are different from each other and from the CRISPR RNA-processing behavior of Cpf1. See, e.g., East-Seletsky, et al., "Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection", *Nature*, 2016 Oct. 13; 538(7624):270-273, the entire contents of which are hereby incorporated by reference. In vitro biochemical analysis of C2c2 in *Leptotrichia shahii* has shown that C2c2 is guided by a single CRISPR RNA and can be programmed to cleave ssRNA targets carrying complementary protospacers. Catalytic residues in the two conserved HEPN domains mediate cleavage. Mutations in the catalytic residues generate catalytically inactive RNA-binding proteins. See e.g., Abudayyeh et al., "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," *Science*, 2016 Aug. 5; 353(6299), the entire contents of which are hereby incorporated by reference.

The crystal structure of *Alicyclobacillus acidoterrestris* C2c1 (AacC2c1) has been reported in complex with a chimeric single-molecule guide RNA (sgRNA). See, e.g., Liu et al., "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism", *Mol. Cell*, 2017 Jan. 19; 65(2):310-322, incorporated herein by reference. The

crystal structure has also been reported for *Alicyclobacillus acidoterrestris* C2c1 bound to target DNAs as ternary complexes. See, e.g., Yang et al., "PAM-dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas endonuclease", *Cell*, 2016 Dec. 15; 167(7):1814-1828, the entire contents of which are hereby incorporated by reference. Catalytically competent conformations of AacC2c1, both with target and non-target DNA strands, have been captured independently positioned within a single RuvC catalytic pocket, with C2c1-mediated cleavage resulting in a staggered seven-nucleotide break of target DNA. Structural comparisons between C2c1 ternary complexes and previously identified Cas9 and Cpf1 counterparts demonstrate the diversity of mechanisms used by CRISPR-Cas9 systems.

In some embodiments, the nucleobase editors described herein comprise a C2c1, a C2c2, or a C2c3 protein. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a C2c1 protein. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a C2c2 protein. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a C2c3 protein. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring C2c1, C2c2, or C2c3 protein. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a naturally-occurring C2c1, C2c2, or C2c3 protein. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of SEQ ID NOs: 1057-1059. In some embodiments, the guide nucleotide sequence-programmable DNA binding protein comprises an amino acid sequence of any one SEQ ID NOs: 1057-1059. It should be appreciated that C2c1, C2c2, or C2c3 from other bacterial species may also be used in accordance with the present disclosure.

C2c1 (uniprot.org/uniprot/T0D7A2#) sp|T0D7A2|C2C1_ALIAG CRISPR-associated endonuclease C2c1 OS = *Alicyclobacillus acidoterrestris* (strain ATCC 49025/DSM 3922/CIP 106132/NCIMB 13137/GD3B) GN=c2c1 PE=1 SV=1 #)

```

C2c1 (uniprot.org/uniprot/T0D7A2#) sp|T0D7A2|C2C1_ALIAG CRISPR-
associated endonuclease C2c1 OS = Alicyclobacillus acidoterrestris
(strain ATCC 49025/DSM 3922/CIP 106132/MCIMB 13137/GD3B) GN = c2c1
PE = 1 SV = 1
MAVKS1KVKLRLLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSSLRQEONLYRRSPNGDGEQECDKTAE
(EQ ID NO: 1057)
MAVKS1KVKLRLLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSSLRQEONLYRRSPNGDGEQECDKTAE
ECKAELLERLRARQVENGHRGPGASDDELLQLARQLYELLVPQAIGAKGDAQQIARKFLSPЛАDKDAV

```

- continued

GGLGIAKAGNKPRWVRMREAGEPGWEEEKEKAETRKSADRTADVLRALADFGLKPLMRVYTDSEMS
 SVEWKPLRGQAVRTWDRDMFQQAIERMMWSWNSWNQRVGQEYAKLVEQKNRFEQKNFVGQEHLV
 HLVNQLQQDMKEASPGLESQEQTAYHTGRALRGSDKVFEKGK LAPDFDLYDAEIKNVQRRNT
 RRFGSHDLFAKLAEPPEYQALWREDASFTRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLG
 GNLHQYTFLFNEFGERRHAIRFHKLKVENGWAREVDDVTVPISMSEQLDNLLPRDPNEPIALYFRDYG
 AEQHFTGEFGGAKIQCRRDQLAHMHRRRGARDVYLNVSVRVQSQS EARGERPPYAAVFRLVGDNH
 RAFVHFDFKLSDYLAEPDDGKLGSEGGLSGLRVMSDLGLRTSASISVFRVARKDELKPNSKGVPFFF
 PIKGNDNLVAHVERSQQLKLPGETESKDLRAIREERQRTLRLRTQAYLRLLVRCGSEDVRRERSW
 AKLIEQPVDAANHMTPDWREAFENELQKLKSLHGICSDKEWMDAVYESVRRVWRHMGKQVRDWK
 DVRSGERPKIRGYAKDVGGNSIEQIEYLERQYKFLKSWSFFGKVSGQVIRAEKGSRAITLREHIDHAK
 EDRLKKLADRRIIMEALGYVYALDERGKGKWWAKYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSH
 RGVFQELINQAQVHDLLVGTMYAAFSSRFDARTGAPGIRCRRVPARCTQEHNPEPFPWWLNKFVVEHT
 LDACPLRADDLIPTGEGEIFVSPFSAEEGFHQIHADLNAAQNLQQLWSDFDISQIRLRCDWGEVDGE
 LVLIPLRTGKRTADSYSNKVFYTNTGVTYYERERGKKRRKVFAQEKLSEEAELLVEADEAREKSVVL
 MRDPSGIINRGNWTROKEFWSMVNQRIEGYLVQIIRSRVPLQDSACENTGDI

 C2c2 (uniprot.org/uniprot/P0DOC6) >sp|P0DOC6|C2C2_LEPSD CRISPR-
 associated endoribonuclease C2c2 OS = Leptotrichia shahii (strain DSM
 19757/CCUG 47503/CIP 107916/JCM 16776/LB37) GN = c2c2 PE = 1 SV = 1
 (SEQ ID NO: 1058)
 MGNLFGHKRWYEVRDCKDFKIKRKVKVCRNYDGNKYILNINENNKEKIDNNKFIRKYINYKNDNI
 LKEFTRKFHAGNLFKLKGKEGIIRIENNDDPLETEEVVLYIEAYGKSEKLKALGITKKKIIDEAIRQGITK
 DDKKIEIKRQENESEEIEIDIREDTNTKTLNDCSIIILRIIENDELETKKSIEIFKNINMSLYKIIIEKIIINETEK
 VFENRYYEEHLREKLLKDDKIDVILTNFMEIREKIKSNLEILGFVKFYLNVGDKKSKNKKMLVEKIL
 NINVDLTVEDIADFVIKELEFWNITKRIEKVKVNNEFLEKRNRNTYIKSYVLLDKHEKFKEIRENKKDK
 IVKFFVENIKNNSIKEKIEKILAECFKIDELIKKLEKELKGNCNTDEIFGIFKKHYKVNFDSSKKFSKKSDEEK
 ELYKIIYRYLKGRIEKILVNEQKVRLLKMEKIEIEKILNESILSEKILKRVQYTLHIMYLGKLHNDID
 MTTVNTDDFSRLHAKEELDLELTFFASTNMELNKIFSRENINNDENIDFFGGDREKNVYVLDKKILNSKI
 KIIRLDFFIDNKNNITNNFIRKFTKIGTNERNRILHAIKERDLQGTQDDYNKVINIIONLKISDEEVSKAL
 NLDVVFKDKNNIITKINDIKISEENNNDIYLPFSKVLPEILNLYRNNPKNEPFDTIFTEKIVLNALIYVN
 KELYKKLILEDDEENESKNIFLQELKKTGLNIDEIDENIIEENYYKNAQISASKGNNAIKKYQKKVIECY
 IGYLRKNYEELPDFSDPKMNQEIKKQIKDINDNKTHERITVKTSDDKTIVINDDPEYIISIFALLNSNAVIN
 KIRNRFFATSVWLNTSEYQNIIDILDEIMQLNLRNECITENWNLNLEEFIQKMKEIEKDFDDFKIQTKE
 IFNNYYEDIKNNILTEFKDDINGCDVLEKKLEKIVIFDDDETKEIDKKSNIHQDEQRKLSNINKDLKKV
 DQYIKDKDQEIKSKILCRIIFNSDFPLKKYKKEIDNLIEDMESENENKFQEIYYPKERKNELYIYKKNLFLNI
 GNPNFDFKIJGLISNDIKMADAKFLPNIDGKNIRKNKISEIDAILKNLNDKLNGSKEYKEKYIKKLKEND
 DFFAKNIQNKNYKSFEKDYNRVSEYKKIRDLVEFNYLNKIESYLDINWKLAQMARFERDMHYIVNGL
 RELGIKLSGYNTGISRAYPKRNGSDGFYTTAYYKFFDEESYKKFEKICYGFGIDLSENSEINKPENESIR
 NYISHFYIVRNPFADYSIAEQIDRVSNLLSSTRYNNSTYASVFEVFKKDVLNLDYDELKKFKLIGNNDI
 LERLMKPKKVSLELESYNSDYIKNLIIELLTKIENTNDTL

-continued

C2c3, translated from >CEPX01008730.1 marine metagenome genome assembly TARA_037_MES_0.1-0.22, contig TARA_037_MES_0.1-0.22_scaffold22115_1, whole genome shotgun sequence.

(SEQ ID NO: 1059)

```

MRSNYHGRNARQWRQKISGLARRTKEFVFTYKFPLETDAEIDPKAVQTYGIAEVGVGHGSЛИGLVC
AFHLSGFRLFSKAGEAMAFRNRSRYPTEFAEAKLSAIMGIQLPTLSPEGLDLIFQSPPRSRDGIAPVWSE
NEVRNRILYTNWTGRGPANKPDEHLLEIAGEIAKQVPPKFGGWDDLASDPDKALAAADKYFQSQQGDFP
SIAISLPAAIMLSPANSTVDFFEGDYIAIDPAETLLHQAVSRCAARLGRERPDLDQNKGPFVSSLQDALVS
SQNNGLSWLFGVGFQHWKEKSPKELIDEYKVPADQHGAVTQVKSFDLAIPLNPLFDTTTHYGEFRASVA
GKVRWSWANYWKRLLDLKSLATTEFTLPESISDPKAVSLEFSGLLVDPQQLKKVADSLPARLVSAEEAI
DRLMGVGIPTAADIAQVERVADEIGAFIGQVQQFNNQVKQKLENLQDADEEFLKGLKIELPSGDKEPP
AINRISGGAPAAAEIFELEEKLQRLLDARSEHFQTISEWAEENAVTLPIAAMVELERLRLAERGATGD
PEEYALRLLLQRIGRLANRVSPVSAGSIRELLKPVMEEREFPNLFFHNRLGSLYRSPYSTSRHQPFSIDVG
KAKAIDWIAGLDQISSDIEKALSGAGEALGDQLRDWINLAGFAISQRLRGLPDTVPNALAQVRCPPDDVR
IPPLLAMLLEEDIARDVCLKAFLNLYVSAINGCLFGALREGFIVRTFQRIGTDQIHYPVPKDAWEYPDR
LNTAKGPINAASWDIEKDGAIVKPVETVRNLSSTGFAGAGVSEYLVQAPHDWYTPLDLRDVAHLVT
GLPVEKNITKLKLTNRTAFRMVGASSFKTHLDSDLKIKLGDFIIIDQHYRQSVTYGGKVKISYEP
ERLQVEAAVPVVDTRDRTVPEPDTLFHDHVAILGERSVGFAVFDIKSCLRTEVVKPIHDNNGNPVVGT
VAVPSIRRLMKAVRSHRRRQPQVNQTYSTALQNYRENIGDVNCNRIDLMERYNAFPVLEFQIKN
FQAGAKQLEIVYGS

```

Cas9 recognizes a short motif (PAM motif) in the CRISPR repeat sequences in the target DNA sequence. A “PAM motif,” or “protospacer adjacent motif,” as used herein, refers a DNA sequence immediately following the DNA sequence targeted by the Cas9 nuclease in the CRISPR bacterial adaptive immune system. PAM is a component of the invading virus or plasmid, but is not a component of the bacterial CRISPR locus. Naturally, Cas9 will not successfully bind to or cleave the target DNA sequence if it is not followed by the PAM sequence. PAM is an essential targeting component (not found in the bacterial genome) which distinguishes bacterial self from non-self DNA, thereby preventing the CRISPR locus from being targeted and destroyed by nuclease.

Wild-type *Streptococcus pyogenes* Cas9 recognizes a canonical PAM sequence (5'-NGG-3'). Other Cas9 nucleases (e.g., Cas9 from *Streptococcus thermophilus*, *Staphylococcus aureus*, *Neisseria meningitidis*, or *Treponema denticola*) and Cas9 variants thereof have been described in the art to have different, or more relaxed PAM requirements. For example, in Kleinstiver et al., *Nature* 523, 481-485, 2015; Klenstiver et al., *Nature* 529, 490-495, 2016; Ran et al., *Nature*, April 9; 520(7546): 186-191, 2015; Kleinstiver et al., *Nat Biotechnol*, 33(12):1293-1298, 2015; Hou et al., *Proc Natl Acad Sci USA*, 110(39):15644-9, 2014; Prykhozhij et al., *PLoS One*, 10(3): e0119372, 2015; Zetsche et al., *Cell* 163, 759-771, 2015; Gao et al., *Nature Biotechnology*, doi:10.1038/nbt.3547, 2016; Want et al., *Nature* 461, 754-761, 2009; Chavez et al., doi: dx.doi.org/10.1101/058974; Fagerlund et al., *Genome Biol.* 2015; 16: 25, 2015; Zetsche et al., *Cell*, 163, 759-771, 2015; and Swarts et al., *Nat Struct Mol Biol*, 21(9):743-53, 2014, each of which is incorporated herein by reference.

Thus, the guide nucleotide sequence-programmable DNA-binding protein of the present disclosure may recognize a variety of PAM sequences including, without limita-

tion: NGG, NGAN, NGNG, NGAG, NGCG, NNGRRT, NGRRN, NNNRRT, NNNGATT, NNAGAAW, NAAAC, TTN, TTTN, and YTN, wherein Y is a pyrimidine, and N is any nucleobase. In some embodiments, the PAM is located 35 5' of the target base. In some embodiments, the PAM is located 3' of the target base.

One example of an RNA-programmable DNA-binding protein that has different PAM specificity is Clustered Regularly Interspaced Short Palindromic Repeats from *Prevotella* and *Francisella* 1 (Cpf1). Similar to Cas9, Cpf1 is also a class 2 CRISPR effector. It has been shown that Cpf1 mediates robust DNA interference with features distinct from Cas9. Cpf1 is a single RNA-guided endonuclease lacking tracrRNA, and it utilizes a T-rich protospacer-adjacent motif (TTN, TTTN, or YTN). Moreover, Cpf1 cleaves DNA via a staggered DNA double-stranded break. Out of 16 Cpf1-family proteins, two enzymes from *Acidaminococcus* and *Lachnospiraceae* are shown to have efficient 45 50 genome-editing activity in human cells.

Also useful in the present disclosure are nuclease-inactive Cpf1 (dCpf1) variants that may be used as a guide nucleotide sequence-programmable DNA-binding protein domain. The Cpf1 protein has a RuvC-like endonuclease domain that is similar to the RuvC domain of Cas9 but does not have a HNH endonuclease domain, and the N-terminal of Cpf1 does not have the alfa-helical recognition lobe of Cas9. It was shown in Zetsche et al., *Cell*, 163, 759-771, 2015 (which is incorporated herein by reference) that, the RuvC-like domain of Cpf1 is responsible for cleaving both DNA strands and inactivation of the RuvC-like domain inactivates Cpf1 nuclease activity. For example, mutations corresponding to D917A, E1006A, or D1255A in *Francisella novicida* Cpf1 (SEQ ID NO: 10) inactivates Cpf1 nuclease activity. In some embodiments, the dCpf1 of the present disclosure comprises mutations corresponding to D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/

95

D1255A, or D917A/E1006A/D1255A in SEQ ID NO: 10. It is to be understood that any mutations, e.g., substitution mutations, deletions, or insertions that inactivates the RuvC domain of Cpf1 may be used in accordance with the present disclosure.

Thus, in some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a nuclease inactive Cpf1 (dCpf1). In some embodiments, the dCpf1 comprises the amino acid sequence of any one SEQ ID NOs: 261-267. In some embodiments, the dCpf1 comprises an amino acid sequence that is at least 85%, at least 90%, at

96

least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at ease 99.5% identical to SEQ ID NO: 10, and comprises mutations corresponding to D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/D1255A, or D917A/E1006A/D1255A in SEQ ID NO: 10. Cpf1 from other bacterial species may also be used in accordance with the present disclosure.

Wild type *Francisella novicida* Cpf1 (SEQ ID NO: 10) (D917, E1006, and D1255 are bolded and underlined)

*Wild type Francisella novicida Cpf1 (SEQ ID NO: 10)
(D917, E1006, and D1255
are bolded and underlined)*
MSIYQEFVNKSYLSKTLRFELIPQGKTLENIKARG**LIL**DDEKRAKDYKKAK**QI**IDKYHQFFIEI**LSSVCIS**
EDLLQNYSDVYFKLKKSDDDNLQKDF**KSAKD**TIK**KQI**SEYIKD**SEKF**KNLFNQNLI**DAKKG**QESDL**LILW**
LKQSKDNGIE**LFKANS**DT**ID**I**DEALE**II**KSFKG**WT**TYF**KGFHENRKNV**YSSND**I**PTSII**Y**RIVDDNL**P**KFL**
NKAKYESLKDK**APAEAIN**YE**QIK**DL**AEE**LT**FDIDY**K**TE**V**NQRV**F**SL**DEV**EIA**N**FNNYLN**Q**S**GI**T**K**ENT**
IIGGKFV**NGENT**K**RKG**IN**EYIN**LY**SQ**Q**IND**KL**K**K**Y**K**M**S**V**LF**K**Q**I**LS**D**TE**S**K**F**VID**K**LE**DD**S**D**V**VTTM**
SFYEQIAAFK**TVEEKS**I**KETL**S**LLF**DD**L**K**Q**K**D**LS**K**Y**F**K**N**D**K**S**L**T**D**LS**QQ**VED**D**YS**V**IGT**A**V**LEY**I**T**
QIAPKNLDNPS**K**KE**Q**EL**I**AK**K**TE**K**AK**Y**LS**LET**I**K**LA**E**EF**N**K**H**R**D**ID**K**Q**C**R**F**E**E**I**L**AN**F**A**I**PM**I**F**D**E**I**A**Q**
NKDNLAQI**SI**K**YQ**N**Q**G**K**D**L**L**Q**ASA**E**DD**V**K**A**I**K**D**L**D**Q**T**N**LL**H**K**L**K**I****F**H**I****S**Q**SE**D**K**AN**I****L**D**K**DE**H**F**Y**
VF**E**E**C**Y**F**E**L**AN**I**V**P**LY**N**K**I**R**Y**I**T****Q**K**P**S**D**E**K**F**K**L**N**F**E**N**S**T**L**ANG**W**D**K**N**K**E**P**D**N**T**A**I**L****F**I**K**DD**K**YY**L****G**
MNKKNNK**I**F**DD**K**AI**K**EN**K**G**E**GY**K**K**I**V**Y**K**L**P**G**A**N**K**M**L**P**K**V**F**FS**A**K**I**K**F**Y**N**P**S**E**D**I**L**R**I**R**H**S**T**H**T****K**
SPQKGYEKF**E**F**N**I**E**D**CR**K**I**D**F**Y**K**Q**S**I**S**K**H**P**E**W**K**D**F**G**F**R**S**D**T**Q**R**Y**N**S**I**D**E**F**Y**R**E**V**N**Q**G**Y**K**L**T****F**E**I**S**E**
YIDSV**N**Q**G**K**L**Y**L****Q**I**Y**N**K**D**F**S**A**Y**S**K**G**R**P**N**L**H**T****L**Y**W**K**A**F**D**E**R**N**L**Q**D**V**V**Y**K**L**N**GE**A**E**L**F**Y**R**K**Q**S**I**P**
ITH**P**AK**E**A**I**AN**K**N**K**D**N**P**K**K**E**S**V**F**E**Y**D**L**I**K**D**K**R**F**T**E**D**K**F**F**H**C**P**I**T**I**N**F**K**S**G**A**N**K**E**D**E**I**N**L**L**K**E**K**A**
VHILS**IDR****G**R**H**L**A**Y**T**L**D**G**K**G**N**I**I**K**Q**D**T**F**N**I**I**G**N**D**R**M**K**T**N**Y**H**D**K**L**A**A**I**E**K**R**D**S**A**R**K**D**W**K**K**I**N****I****K**
MKEGYL**S**Q**V**H**E**I**A**K**L**V**I**E**Y**N**A**I**V****F****E**D**L**N**F**G**F**K**R**G**R**F**K**V**E**K**Q**V**Y**Q**K**L**E**K**M**L**E**K**L**N**Y**L**V**F**K**D**N**E**F**D**K**
TGGVL**R**A**Y**Q**L**T**A**P**F**E**T**F**K**K**M**G**K**Q**T**G**I**I**Y**Y**V**P**A**G**F**T**S**K**I**C**P**V**G**F**V**N**Q**L**Y**P**K**Y**E**S**V**S**K**S**Q**E**F**S**K**F**D**K**I**
NLDKGYFE**F**S**F**D**Y**K**N**F**G**D**K**A**A**K**G**K**W**T**I**A**F**G**S**R**L**I**N**F**R**N**S**D**K**H**N**W**D**T**R**V**Y**P**T**K**E**L**K**L**D****S****I****E**
GHGECI**KAA**I**C**G**E**DK**K**FF**A**KL**T**SV**L**N**T**I**L**Q**M**R**N**S**K**T**G**T**E**LD**L**I**S**P**V**AD**V**GN**N**FF**D**S**R**Q**A**P**K**N**M**P**Q**
DANGAYHIGLKGLMLLGRIKNNQEGKKLN**L**VI**K**NE**EY**FE**F**V**Q**N**R**NN

Francisella novicida Cpf1 D917A (SEQ ID NO: 261) (A917, E1006, and D1255 are bolded and underlined)
MSIYQEFVNKSYLSKTLRFELIPQGKTLENIKARG**LIL**DDEKRAKDYKKAK**QI**IDKYHQFFIEI**LSSVCIS**
EDLLQNYSDVYFKLKKSDDDNLQKDF**KSAKD**TIK**KQI**SEYIKD**SEKF**KNLFNQNLI**DAKKG**QESDL**LILW**
LKQSKDNGIE**LFKANS**DT**ID**I**DEALE**II**KSFKG**WT**TYF**KGFHENRKNV**YSSND**I**PTSII**Y**RIVDDNL**P**KFL**
NKAKYESLKDK**APAEAIN**YE**QIK**DL**AEE**LT**FDIDY**K**TE**V**NQRV**F**SL**DEV**EIA**N**FNNYLN**Q**S**GI**T**K**ENT**
IIGGKFV**NGENT**K**RKG**IN**EYIN**LY**SQ**Q**IND**KL**K**K**Y**K**M**S**V**LF**K**Q**I**LS**D**TE**S**K**F**VID**K**LE**DD**S**D**V**VTTM**
SFYEQIAAFK**TVEEKS**I**KETL**S**LLF**DD**L**K**Q**K**D**LS**K**Y**F**K**N**D**K**S**L**T**D**LS**QQ**VED**D**YS**V**IGT**A**V**LEY**I**T**
QIAPKNLDNPS**K**KE**Q**EL**I**AK**K**TE**K**AK**Y**LS**LET**I**K**LA**E**EF**N**K**H**R**D**ID**K**Q**C**R**F**E**E**I**L**AN**F**A**I**PM**I**F**D**E**I**A**Q**
NKDNLAQI**SI**K**YQ**N**Q**G**K**D**L**L**Q**ASA**E**DD**V**K**A**I**K**D**L**D**Q**T**N**LL**H**K**L**K**I****F**H**I****S**Q**SE**D**K**AN**I****L**D**K**DE**H**F**Y**
VF**E**E**C**Y**F**E**L**AN**I**V**P**LY**N**K**I**R**Y**I**T****Q**K**P**S**D**E**K**F**K**L**N**F**E**N**S**T**L**ANG**W**D**K**N**K**E**P**D**N**T**A**I**L****F**I**K**DD**K**YY**L****G**
MNKKNNK**I**F**DD**K**AI**K**EN**K**G**E**GY**K**K**I**V**Y**K**L**P**G**A**N**K**M**L**P**K**V**F**FS**A**K**I**K**F**Y**N**P**S**E**D**I**L**R**I**R**H**S**T**H**T****K**
SPQKGYEKF**E**F**N**I**E**D**CR**K**I**D**F**Y**K**Q**S**I**S**K**H**P**E**W**K**D**F**G**F**R**S**D**T**Q**R**Y**N**S**I**D**E**F**Y**R**E**V**N**Q**G**Y**K**L**T****F**E**I**S**E**
YIDSV**N**Q**G**K**L**Y**L****Q**I**Y**N**K**D**F**S**A**Y**S**K**G**R**P**N**L**H**T****L**Y**W**K**A**F**D**E**R**N**L**Q**D**V**V**Y**K**L**N**GE**A**E**L**F**Y**R**K**Q**S**I**P**

- continued

ITHPAKEAIANKNDNPKESVFEYDLIKDKRFTEDKFFFHCPIINFKSSGANKENDEINLLLKEAND
VHILSIARGERHLAYTLVDGKGNI **I**KQDTFNIIGNDRMKTNYHDKLAIAEKDRDSARKDWKKINNIKE
MKEGYLSQVVHIEAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDK
TGGVLRAYQLTAPFETPKKMKGQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICY
NLDKGYFEFSFDYKNFGDKAAGKWTIASFGSRLINFRNSDNHNDTREVYPTKELEKLLKDYSIEY
GHGECIKAACGESDKFFAKLTSVLNTILQMRNSKTGTTELDDYLISPADVNGNFFDSRQAPKNMPQDA
DANGAYHIGLKGLMLLGRIKNNQEGKKLNVIKNEEYFEFVQNRNN

Francisella novicida Cpf1 E1006A (SEQ ID NO: 262) (D917, A1006, and D1255 are bolded and underlined)
MSIYQEFVNKYSLKTRFELIPQGKTLENIKARGLILDEKRADYKKAQIIDKYHQFFIEEILSSVCIS

EDLLQNYSDVYF~~P~~KLKKSDDDNLQKDFSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILW
LKQSKDNGIELFKANSDITDIDEALEI **I**KSFKGWTTYFKGPHENRKNVSSNDIPTSIIYRIVDDNLPKFLE
NKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKENT
IIGGKFVNGENTKRKGINEYINLYSQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
SFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLDSQQVFDDYSVIGTAVLEYITQ
QIAPKNDNPSKKQELIAKKTEKAYLSETIKLAEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQ
NKDNLAQISIKYQNQGKDLLQASEDDVKAIKDLDQTNLLHKLKIFHISQEDKANILDKDEHFYL
VFEECYFELANIPVLYNKIRNYITQKPSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDKYYLGV
MNKKKNNKIFDDKAIKENKGEGYKKIVYKLPGANKMLPKVFFSAKIKFYNPSEILRIRNHSTHTKN
SPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREEVUQGYKLTFENISES
YIDSVVNQGKLYFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKK
ITHPAKEAIANKNDNPKESVFEYDLIKDKRFTEDKFFFHCPIINFKSSGANKENDEINLLLKEAND
VHILSIARGERHLAYTLVDGKGNI **I**KQDTFNIIGNDRMKTNYHDKLAIAEKDRDSARKDWKKINNIKE
MKEGYLSQVVHIEAKLVIEYNAIVVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDK
TGGVLRAYQLTAPPETPKKMKGQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICY
NLDKGYFEFSFDYKNFGDKAAGKWTIASFGSRLINFRNSDNHNDTREVYPTKELEKLLKDYSIEY
GHGECIKAACGESDKFFAKLTSVLNTILQMRNSKTGTTELDDYLISPADVNGNFFDSRQAPKNMPQDA
DANGAYHIGLKGLMLLGRIKNNQEGKKLNVIKNEEYFEFVQNRNN

Francisella novicida Cpf1 D1255A (SEQ ID NO: 263) (D917, E1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLKTRFELIPQGKTLENIKARGLILDEKRADYKKAQIIDKYHQFFIEEILSSVCIS

EDLLQNYSDVYF~~P~~KLKKSDDDNLQKDFSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILW
LKQSKDNGIELFKANSDITDIDEALEI **I**KSFKGWTTYFKGPHENRKNVSSNDIPTSIIYRIVDDNLPKFLE
NKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKENT
IIGGKFVNGENTKRKGINEYINLYSQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
SFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLDSQQVEDDYSVIGTAVLEYITQ
QIAPKNDNPSKKQELIAKKTEKAYLSETIKLAEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQ
NKDNLAQISIKYQNQGKDLLQASEDDVKAIKDLDQTNLLHKLKIFHISQEDKANILDKDEHFYL
VFEECYFELANIPVLYNKIRNYITQKPSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDKYYLGV
MNKKKNNKIFDDKAIKENKGEGYKKIVYKLPGANKMLPKVFFSAKIKFYNPSEILRIRNHSTHTKN
SPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREEVUQGYKLTFENISES
YIDSVVNQGKLYFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKK
ITHPAKEAIANKNDNPKESVFEYDLIKDKRFTEDKFFFHCPIINFKSSGANKENDEINLLLKEAND

- continued

VHILS IDRGERHLAYTLDGKGNIIKQDTPNIIGNDRMKTNYHDKLAIAIEKDRDSARKDWKKINNIKE
 MKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDK
 TGGVLRAYQLTAPFETFKKGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICY
 NLDKGYFEFSFDYKNFGDKAAKGWTIASFGSRLINFRNSDKHNWDREVYPTKELEKLLKDYSIEY
 GHGECIKAACGESDKFFAKLTSVLNTILQMRNSKTGTLEDYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGMLLGRKNNQEGKKLNVIKNEEYFEFVQNRNN

Francisella novicida Cpf1 D917A/E1006A (SEQ ID NO: 264) (A917, A1006, and D1255 are bolded and underlined)
 MSIYQEVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCIS
 EDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKQISEYIKDEKFNLFNQNLIDAKKGQESDLLW
 LKQSKDNGIELFKANSDITDIDEALEEIKSFKGTTYFKGHENRKNVSSNDIPTSIIYRIVDDNLPKFLE
 NKAKYESLKDKAPEAINYQIKDLAEELTFDIDYKTSEVNQRVFSLDEVEIANFNNYLNQSGITKENT
 IIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFQILSDTESKSFVIDKLEDDSVVTTMQ
 SFYEQIAAFKTVEEKSIKETLSLFDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQ
 QIAPKNLDNPSKKKEQELIAKTKEAKYLSETIKLAEEFNKHRDIDKQCRFEEILFAAIPMIFDEIAQ
 NKDNLAQISIKYQNGKDLLQQSAEDDVKAIKDLDQTNLLHKLKIFHISQSEDKAILDKDEHFYL
 VFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYLGV
 MNKKNNKIFDDKAIKENKGEGYKIVYKLLPGANKMLPVFFSAKIKFYNPSEDILRIRNHSTHTKNG
 SPQKGYEKFEFNIEDCRKFIDFYKQSIKHPEWKDFGFRFSDTQRYNSIDEFYREEVENQGYKLTFENISES
 YIDSVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAFDERNLQDVVYKLNGEAEFRKQSIPKK
 ITHPAEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKEDEINLLKEKAND
 VHILS IDRGERHLAYTLDGKGNIIKQDTPNIIGNDRMKTNYHDKLAIAIEKDRDSARKDWKKINNIKE
 MKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDK
 TGGVLRAYQLTAPFETFKKGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICY
 NLDKGYFEFSFDYKNFGDKAAKGWTIASFGSRLINFRNSDKHNWDREVYPTKELEKLLKDYSIEY
 GHGECIKAACGESDKFFAKLTSVLNTILQMRNSKTGTLEDYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGMLLGRKNNQEGKKLNVIKNEEYFEFVQNRNN

Francisella novicida Cpf1 D917A/D1255A (SEQ ID NO: 265) (A917, E1006, and A1255 are bolded and underlined)
 MSIYQEVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCIS
 EDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKQISEYIKDEKFNLFNQNLIDAKKGQESDLLW
 LKQSKDNGIELFKANSDITDIDEALEEIKSFKGTTYFKGHENRKNVSSNDIPTSIIYRIVDDNLPKFLE
 NKAKYESLKDKAPEAINYQIKDLAEELTFDIDYKTSEVNQRVFSLDEVEIANFNNYLNQSGITKENT
 IIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFQILSDTESKSFVIDKLEDDSVVTTMQ
 SFYEQIAAFKTVEEKSIKETLSLFDLKAQKLDLSKIYFKNDKSLTDLSQQVEDDYSVIGTAVLEYITQ
 QIAPKNLDNPSKKKEQELIAKTKEAKYLSETIKLAEEFNKHRDIDKQCRFEEILFAAIPMIFDEIAQ
 NKDNLAQISIKYQNGKDLLQQSAEDDVKAIKDLDQTNLLHKLKIFHISQSEDKAILDKDEHFYL
 VFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYLGV
 MNKKNNKIFDDKAIKENKGEGYKIVYKLLPGANKMLPVFFSAKIKFYNPSEDILRIRNHSTHTKNG
 SPQKGYEKFEFNIEDCRKFIDFYKQSIKHPEWKDFGFRFSDTQRYNSIDEFYREEVENQGYKLTFENISES
 YIDSVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAFDERNLQDVVYKLNGEAEFRKQSIPKK
 ITHPAEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKEDEINLLKEKAND

- continued

VHILSIRGERHLAYTLVDGKGNI IKQDTFNIIGNDRMKTNYHDKLAIAIEKDRDSARKDWKKINNIKE
 MKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDK
 TGGVLRAYQLTAPFETFKKMKGQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICY
 NLDKGYFEFSFDYKNFGDKAAKGWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLKDYSIEY
 GHGECIKAAICGESDKFFAKLTSVLNTILQMRNSKTGTTELDELYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGLMLLGRIKNQEGKKLNVIKNEEYFEFVQNRNN

Francisella novicida Cpf1 E1006A/D1255A (SEQ ID NO: 266) (D917, A1006, and A1255 are bolded and underlined)
 MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDDEKRAKDYKKAQI IDKYHQFFIEEILSSVCIS
 EDLLQNYSDVYF~~PKL~~KKSDDDNLQKDFSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILW
 LKQSKDNGIELPKANSDITDIDEALEI IKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLE
 NKAKYESLKDKPAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKENT
 IIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFQILSDTESKSFVIDKLEDDSDVTTMQ
 SFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDSVIGTAVLEYITQ
 QIAPKNLDNPSKKQEQELIAKKTEKAKYLSETIKLAEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQ
 NKDNLQISIKYQNQGKDLLQASAEDDVKAIKDLDQTNLLHKLKIFHISQEDKANILDKDEHFYL
 VFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGV
 MNKKNNKIFDDKAIKENGEGYKIVYKLLPGANKMLPKVFPSAKSIKFNPSEDILRIRNHSTHTKNG
 SPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREEVUQGYKLTFENISES
 YIDSVNQGKLFQIYNKDFSAYSKGRPNLHTLYWKAFDERNLQDVVKLNGEAEFYRKQSIPK
 ITHPAKEAIANKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSGANKENDEINLLKEKAND
 VHILSIRGERHLAYTLVDGKGNI IKQDTFNIIGNDRMKTNYHDKLAIAIEKDRDSARKDWKKINNIKE
 MKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDK
 TGGVLRAYQLTAPFETFKKMKGQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICY
 NLDKGYFEFSFDYKNFGDKAAKGWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLKDYSIEY
 GHGECIKAAICGESDKFFAKLTSVLNTILQMRNSKTGTTELDELYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGLMLLGRIKNQEGKKLNVIKNEEYFEFVQNRNN

Francisella novicida Cpf1 D917A/E1006A/D1255A (SEQ ID NO: 267) (A917, A1006, and A1255 are bolded and underlined)
 MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDDEKRAKDYKKAQI IDKYHQFFIEEILSSVCIS
 EDLLQNYSDVYF~~PKL~~KKSDDDNLQKDFSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILW
 LKQSKDNGIELPKANSDITDIDEALEI IKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLE
 NKAKYESLKDKPAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKENT
 IIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFQILSDTESKSFVIDKLEDDSDVTTMQ
 SFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDSVIGTAVLEYITQ
 QIAPKNLDNPSKKQEQELIAKKTEKAKYLSETIKLAEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQ
 NKDNLQISIKYQNQGKDLLQASAEDDVKAIKDLDQTNLLHKLKIFHISQEDKANILDKDEHFYL
 VFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGV
 MNKKNNKIFDDKAIKENGEGYKIVYKLLPGANKMLPKVFPSAKSIKFNPSEDILRIRNHSTHTKNG
 SPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREEVUQGYKLTFENISES
 YIDSVNQGKLFQIYNKDFSAYSKGRPNLHTLYWKAFDERNLQDVVKLNGEAEFYRKQSIPK
 ITHPAKEAIANKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSGANKENDEINLLKEKAND
 VHILSIRGERHLAYTLVDGKGNI IKQDTFNIIGNDRMKTNYHDKLAIAIEKDRDSARKDWKKINNIKE

- continued

```
MKEGYLSQVVHEIAKLVIEYNAIVVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDK
TGGVLRAYQLTAPFKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICY
NLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKHNWDREVYPTKELEKLKDYSIEY
GHGECIKAAICGESDKFFAKLTSVLNTILQMRNSKTGTEDYLISPVADVNGNFFDSRQAPKNMPQDA
AANGAYHIGLKGLMLLGRIKNQEGKKLNVIKNEEYFEPVQNRNN
```

In some embodiments, the guide nucleotide sequence-programmable DNA binding protein is a Cpf1 protein from a *Acidaminococcus* species (AsCpf1). Cpf1 proteins form *Acidaminococcus* species have been described previously and would be apparent to the skilled artisan. Exemplary ¹⁵ *Acidaminococcus* Cpf1 proteins (AsCpf1) include, without limitation, any of the AsCpf1 proteins provided herein.

Wild-type AsCpf1 - Residue R912 is indicated in bold underlining and residues 661-667 are indicated in italics and underlining.

(SEQ ID NO: 1060)

```
TQFEGFTNLYQVSKTLRFELIPQGKTLKHIEQQGFIEEDKARNDHYKELKPIIDRIYKTYADQCQLVQL
DWENLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNG
KVLQLGTVTTEHENALLRSDKFTTYFSGFYENRKNVFASEDISTAIPHRIVQDNFPKFKENCHIFTRL
ITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQIDLYNQLGGISREAGTEKIKGLNEVLNL
AIQKNDETAHIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLLRNENVLEAALFNEL
NSIDLTHIFSHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAEKVQRSLKHEDINLQEIISAG
KELSEAFKQKTSEILSHAALDQPLPTTMLKKQEEEILKSQLDSLLLGLYHLLDWFAVDESNEVDPEF
SARLTGIKLEMEPSLSFYNKARNYATKKPSVEKFKLNFQMPTLAGWDVNKENNGAILFVKNGLY
YLGIMPKQKGRYALSPEPETEKTSEGFDMYYDFPDAKMIPKCSTQLKAVTAHFQTHTPILSNNF
IEPLEIUTKEIYDLNNPEKEPKKQTAYAKKTGQKGYREALCKWIDFTRDLSKTKTTSIDLSLRPSQ
YKDLGEEYAENPLLYHISFQRIAEKEIMDAVETGLYLFQIYNKDFAKGHGKPNLHTLYWTGLFSP
NLAKTSIKLNQAELFYRPKSRMKRMAHRLGEKMLNKLKDQKTPIPDTLYQELYDYVNHRLSHDLS
DEARALLPNVITKEVSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGE
RNLIYIYUTVIDTGKILEQRSLMTIQQFDYQKLDNREKERVAARQAWSVGTIKDLKQGYLSQVIHEIVDE
LMIHYQAVVVLENLNFGKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLDQ
FTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHERKFLEGFDFLHYDVKTGDFILHFK
MNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVENHRFGRDLYPANELIALLE
EKGIVFRDGSNILPKLENDDSHAIDTMVALIRSVLQMRNSAATGEDYINSPVRDLNGVCFDSRFQNP
EWPMDADANGAYHIALKGQLLNHLKESKDLKQONGISNQDWLAYQELRN
```

AsCpf1 (R912A) - Residue A912 is indicated in bold underlining and residues 661- 667 are indicated in italics and underlining.

(SEQ ID NO: 1061)

```
TQFEGFTNLYQVSKTLRFELIPQGKTLKHIEQQGFIEEDKARNDHYKELKPIIDRIYKTYADQCQLVQL
DWENLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNG
KVLQLGTVTTEHENALLRSDKFTTYFSGFYENRKNVFASEDISTAIPHRIVQDNFPKFKENCHIFTRL
ITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQIDLYNQLGGISREAGTEKIKGLNEVLNL
AIQKNDETAHIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLLRNENVLEAALFNEL
NSIDLTHIFSHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAEKVQRSLKHEDINLQEIISAG
KELSEAFKQKTSEILSHAALDQPLPTTMLKKQEEEILKSQLDSLLLGLYHLLDWFAVDESNEVDPEF
```

- continued

```

SARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNQMPPTLASGVWDVNKEKNNGAILFVKNGLY
YLGIMPQKGRYKALSFEPETEKTSEGFDKMYDYFPDAAKMIPKCSTQLKAVTAHFQTHTPILLSNNF
IEPLEITKEIYDLNNPEKEPKFQTYAKKTDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQ
YKDLGEYYAELNPLLYHISFQRIAEEKEMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPE
NLAKTSIKLNGQAELFYRPKSRMKRMRAHRLGEKMLNKKLDQKTPIPDTLYQELYDYNVNHLRSHDLS
DEARALLPNVITKEVSHEIICKDRRTSDKFFFHVPI TLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGE
ANLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNRKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVD
LMIHYQAVVVLLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLDQ
FTSFAKMGTQSGLFYVPAPYTSKIDPLTGFDVDPFWKTIKNHESRKHFLEGFDLHYDVKTGDFILHFK
MNRLSFLQRGLPGFMPAWDIVFEKNETQFDAKGTPIAGKRIVPVIENHRFTGRYRDLYPANELIALLE
EKGIVFRDGSNLPLKLLENDD SHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLNGCFDSRFQNP
EWPMADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQELRN

```

In some embodiments, the nucleic acid programmable DNA binding protein is a Cpf1 protein from a Lachnospiraceae species (LbCpf1). Cpf1 proteins form Lachnospiraceae species have been described previously have been

described previously and would be apparent to the skilled artisan. Exemplary Lachnospiraceae Cpf1 proteins (LbCpf1) include, without limitation, any of the LbCpf1 proteins provided herein.

Wild-type LbCpf1
(SEQ ID NO: 1062)

```

MSKLEKFTNCYSLSKTLRFKAI PVGKTQENIDNKRLLV EDEKRAEDYKGVKKLLDRYYLSFINDVLHSI
KLKNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGNEG YKSLFKDII ETILPEFLDDKDEIAL
VNSFNGTTAFTGFFDNRENMPSEEAKSTSIAFRCINENLTRYISNMDIFEKVDIAFDKHEVQEIKEKILN
SDYDVEDPFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTQKLPFKPLYKQVL
SDRESLSFYGEGYTSDEEVLEVFRNLTNKNSEIFSSIKKLEKLFKNFDEYSSAG-
IFVKNGPAISTSKDIEG
EWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLKEII IQK
VDEIYKVGSSSEKLFADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYIKAFFGEGKETNRDESFYGD
FVLAYDILLKVDHIYDAIRNYVTQKPYSKDKF KLYFQNPQFMGGWDKDKETDYRATILRGSKYLA
MDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKWMAYYNPSEDIQKIKYKNGTFKK
GDMFNLNDCHKLIDFFKDISRYPKWSNAYDFNFSETEKYKDIAGFYREVEQGYKVSFESASKKEVD
KLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMFKLLFDENNHGQIRLSGGAEFLMRRASLKKEELVVH
PANSPIANKNPDNPKTTTSLSYDVKDKRFSQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPVVI
GIDRGERNLILYIVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWT SIENIKELK
AGYISQVVKICELVEKYDAVIALEDLNSGFKNRSRKVEKQVYQKPEKMLIDKLN YMVDKSNPCAT
GGALKGYQITNKFESPKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIMYV
EEDLFEPALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPKKNVFDWEEVCLTSAYKELFNKGINY
QQGDIRALLCEQSDKAFYSSFMALMSMLQMRNSITGRTDVFILISPVKNSDQGIFYDSRNYEAOENAIL
PKNADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKh

```

LbCpf1 (R836A)
(SEQ ID NO: 1063)

```

MSKLEKFTNCYSLSKTLRFKAI PVGKTQENIDNKRLLV EDEKRAEDYKGVKKLLDRYYLSFINDVLHSI
KLKNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGNEG YKSLFKDII ETILPEFLDDKDEIAL
VNSFNGTTAFTGFFDNRENMPSEEAKSTSIAFRCINENLTRYISNMDIFEKVDIAFDKHEVQEIKEKILN
SDYDVEDPFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTQKLPFKPLYKQVL

```

107

108

- continued

SDRESLSFYGEGYTSDEEVLEVFRNLTNKNSEIFSSIKKLEKLFKNFDEYSSAG-
IFVKNGPAISTISKDIFG

EWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLKEIIIQK
VDEIYKVYGSSEKLFADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYIKAFFGEGKETNRDESFYGD
FVLAYDILLKVDHIYDAIRNYTQKPYSKDKFKLYPQNQPQFMGGWDKDKETDYRATILRGSKYALI
MDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKWMAYYNPSEDIQKIKNGTFKK
GDMFNLNDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYDIAGFYREVEEQGYKVSFESASKKEVD
KLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMYFKLLFDENNHGQIRLSGGAEFLMRRASLKEELVVH
PANSPIANKNPDNPKKTTLSYDVYKDKRFSQYELHIPIAINKCPKNIFKINTEVRLVLLKHDDNPYVI
GIDRGEANLLYIVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKEL
KAGYISQVVKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQKPEKMLIDKLNYMVDKSNPC
TGGALKGYQITNKFESPKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIMYV
PEEDLFEPALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPKNNVFDWEEVCLTSAYKELFNKG
YQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENAI
LPKNADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTsvkh

LbCpf1 (R1138A)

(SEQ ID NO: 1064)

MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAEDYKGVKKLLDRYYLSFINDVLHSI
KLKNLNYYISLFRKKTRTEKENKELENLEINLRKEIAKFKGNEGYKSLFKDIIETILPEFLDDKDEIAL
VNSFNGFTTAFTGFFDNRENMFSEEAKSTSIAFCRINCENLTRYISNMDFEKAIFDKHEVQEIKEKILN
SDYDVEDFFEGEFFNFVLTQEGIDVYNIAIIGGFVTESGEKIKGLNEYINLYNQTKQKLPKFKPLYKQVL
SDRESLSFYGEGYTSDEEVLEVFRNLTNKNSEIFSSIKKLEKLFKNFDEYSSAG-
IFVKNGPAISTISKDIFG

EWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLKEIIIQK
VDEIYKVYGSSEKLFADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYIKAFFGEGKETNRDESFYGD
FVLAYDILLKVDHIYDAIRNYTQKPYSKDKFKLYPQNQPQFMGGWDKDKETDYRATILRGSKYALI
MDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKWMAYYNPSEDIQKIKNGTFKK
GDMFNLNDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYDIAGFYREVEEQGYKVSFESASKKEVD
KLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMYFKLLFDENNHGQIRLSGGAEFLMRRASLKEELVVH
PANSPIANKNPDNPKKTTLSYDVYKDKRFSQYELHIPIAINKCPKNIFKINTEVRLVLLKHDDNPYVI
GIDRGERNLLYIVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKEL
AGYISQVVKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQKPEKMLIDKLNYMVDKSNPC
GGALKGYQITNKFESPKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIMYV
EEDLFEPALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPKNNVFDWEEVCLTSAYKELFNKG
YQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENAIL
PKNAADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTsvkh

In some embodiments, the Cpf1 protein is a crippled Cpf1 protein. As used herein a “crippled Cpf1” protein is a Cpf1 protein having diminished nuclease activity as compared to a wild-type Cpf1 protein. In some embodiments, the crippled Cpf1 protein preferentially cuts the target strand more efficiently than the non-target strand. For example, the Cpf1 protein preferentially cuts the strand of a duplexed nucleic acid molecule in which a nucleotide to be edited resides. In some embodiments, the crippled Cpf1 protein preferentially cuts the non-target strand more efficiently than the target strand. For example, the Cpf1 protein preferen-

tially cuts the strand of a duplexed nucleic acid molecule in which a nucleotide to be edited does not reside. In some embodiments, the crippled Cpf1 protein preferentially cuts the target strand at least 5% more efficiently than it cuts the non-target strand. In some embodiments, the crippled Cpf1 protein preferentially cuts the target strand at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50%, 60%, 70%, 80%, 90%, or at least 100% more efficiently than it cuts the non-target strand.

In some embodiments, a crippled Cpf1 protein is a non-naturally occurring Cpf1 protein. In some embodi-

109

ments, the crippled Cpf1 protein comprises one or more mutations relative to a wild-type Cpf1 protein. In some embodiments, the crippled Cpf1 protein comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 mutations relative to a wild-type Cpf1 protein. In some embodiments, the crippled Cpf1 protein comprises an R836A mutation mutation as set forth in SEQ ID NO: 763, or in a corresponding amino acid in another Cpf1 protein. It should be appreciated that a Cpf1 comprising a homologous residue (e.g., a corresponding amino acid) to R836A of SEQ ID NO: 763 could also be mutated to achieve similar results. In some embodiments, the crippled Cpf1 protein comprises a R1138A mutation as set forth in SEQ ID NO: 763, or in a corresponding amino acid in another Cpf1 protein. In some embodiments, the crippled Cpf1 protein comprises an R912A mutation mutation as set forth in SEQ ID NO: 762, or in a corresponding amino acid in another Cpf1 protein. Without wishing to be bound by any particular theory, residue R838 of SEQ ID NO: 763 (LbCpf1) and residue R912 of SEQ ID NO: 762 (AsCpf1) are examples of corresponding (e.g., homologous) residues. For example, a portion of the alignment between SEQ ID NO: 762 and 763 shows that R912 and R838 are corresponding residues.

In some embodiments, any of the Cpf1 proteins provided herein comprises one or more amino acid deletions. In some

110

embodiments, any of the Cpf1 proteins provided herein comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acid deletions. Without wishing to be bound by any particular theory, there is a helical region in Cpf1, which includes residues 661-667 of AsCpf1 (SEQ ID NO: 762), that may obstruct the function of a deaminase (e.g., APOBEC) that is fused to the Cpf1. This region comprises the amino acid sequence KKTGDQK. Accordingly, aspects of the disclosure provide Cpf1 proteins comprising mutations (e.g., deletions) that disrupt this helical region in Cpf1. In some embodiments, the Cpf1 protein comprises one or more deletions of the following residues in SEQ ID NO: 762, or one or more corresponding deletions in another Cpf1 protein: K661, K662, T663, G664, D665, Q666, and K667. In some embodiments, the Cpf1 protein comprises a T663 and a D665 deletion in SEQ ID NO: 762, or corresponding deletions in another Cpf1 protein. In some embodiments, the Cpf1 protein comprises a K662, T663, D665, and Q666 deletion in SEQ ID NO: 762, or corresponding deletions in another Cpf1 protein. In some embodiments, the Cpf1 protein comprises a K661, K662, T663, D665, Q666 and K667 deletion in SEQ ID NO: 762, or corresponding deletions in another Cpf1 protein.

AsCpf1 (deleted T663 and D665)
(SEQ ID NO: 1065)

TQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQQFIEEDKARNDHYKELKPIIDRIYKTYADQCLQLVQL
DWENLSAAIDSYSRKETEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNG
KVLKQLGTVTTEHENALLRSFDKFTTYSFGFYENRKNVSAEDISTAIPHRIQDNFPKFKENCHIFTRL
ITAVPSLREHFENVKKAIGIFVSTSIIEVFSFPFYNQLLTQIDLYNQLGGISREAGTEKIKGLNEVNL
AIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLRNENVLETAEALFNL
NSIDLTHIFSHKKLETISSALCDHWDTLRNALYERRISELTGKITSAKEVKVQRSLKHEDINLQEIISAAG
KELSEAFQKTSEILSHAHAALDQPLPTMLKKQEEKEILSQLDSLLGLYHLLDWFADVDESNEVDPEF
SARLTGIKLEMPEPSLSFYNKARNYATKKPYSVEKFPLKLNQMPMLASGVNKEKNNGAILFVKNGLY
YLGIMPQKGRYKALSFEPTEKTSSEGFDKMYDYFPDAAKMIPKCSTQLKAVTAHFQTHTPILLSNNF
IEPLEITKEIYDLNNPEKEPKKFQTYAKKGQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQY
KDLGEYYAELNPLLYHISFQRIAKEIMDAVETGKLYLFQIYNKDFAKHHGKPNLHTLYWTGLFSPEN
LAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKKLDQKTPIPDTLYQELYDYVNHLRSHDLSD
EARALLPNVITKEVSHEIICKDRRTSDKFFFHVPIITLNQYQANSPSKFNQRVNAYLKEHPETIIGIDRGER
NLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREKERVAARQAWSVVGTLKDLKQGQLSQQVIHEIVDL
MIHYQAVVYLENLNGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQF
TSFAKMGTSQSGFLFYVPAPYTSKIDPLTGFDVFWKTIKNHESRKHFLEGFDLHYDVKTGDFILHFK
MNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIEHRFTGRYRDLYPANELIALLE
EKGIVPRDSNILPKLLENDDSHAITDMVALIRSVLQMRNSNAATGEDYIINSPVRDLNGVCFDSRFQNP
EWPMADANGAYHIALKGQLLNHLKESKDLKLQNGISNQDWLAYIQELRN

AsCpf1 (deleted K662, T663, D665, and Q666)
(SEQ ID NO: 1066)

TQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQQFIEEDKARNDHYKELKPIIDRIYKTYADQCLQLVQL
DWENLSAAIDSYSRKETEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNG
KVLKQLGTVTTEHENALLRSFDKFTTYSFGFYENRKNVSAEDISTAIPHRIQDNFPKFKENCHIFTRL
ITAVPSLREHFENVKKAIGIFVSTSIIEVFSFPFYNQLLTQIDLYNQLGGISREAGTEKIKGLNEVNL

- continued

AIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFSDEEVQSPCKYKTLLRNENVLETAEALFNL
 NSIDLTHIFSHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEVQRSLKHEDINLQEIIISAAG
 KELSEAFKQKTSEILSHAHAALDQPLPTTMLKKQEEKEILKSQDSLLGLYHLLDWFADVDESNEVDPEF
 SARLTGIKLEMESLSPYNKARNYATKKPYSVEFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLY
 YLGIMPQKGRYKALSPEPTEKTSEGFDKMYDYFPDAAKMIPKCSTQLKAVTAHFQTHTPILLSNNF
 IEPLEITKEIYDLNNPEKEPKFQTYAKGKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDL
 GEYYAELNPPLYHISFQRIAEKIMDAVETGKLYLFQIYNKDFAKGHGKPNLHTLYWTGLFSPENLAK
 TSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKKLDQKTPIDPLTLYQELYDYVNHRSLHDLSDEAR
 ALLPNVITKEVSHEIIKDRRTSDKFFFHVPI TLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGERNL
 YITVIDSTGKILEQRSLNTIQQFDYQKQLDNREKERAARQAWSVVGTIKDLQGQLSQVIHEIVDLM
 HYQAVVVLENLNPGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTS
 FAKMGTQSGFLFYVPAPYTSKIDPLTGFDPPFWKTIKNHESRKHFLEGFDLHYDVKTGDFILHFKM
 RNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIEHRFTGRYRDLYPANELIALEEKG
 IVFRDGSNILPKLLENDSSHAIITMVALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWP
 MDADANGAYHIALKGQLLNHLESKDLKLQNGISNQDWLAYIQELRN

AsCpf1 (deleted K661, K662, T663, D665, Q666, and K667)

(SEQ ID NO: 1067)

TQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQQFIEEDKARNDHYKELKPIIDRIYKTYADQCLQLVQL
 DWENLSAAIDSYRKETEETRNALIEEQATYRNAIHDFIGRTDNLTDAINKRHAEIYKGLFKAELFNG
 KVLKQLGTVTTEHENALLRSFDKFTTYFSGFYENRKNVFAEDI STAI PHRIVQDNFPKFKENCHIFTRL
 ITAVPSLREHFENVKKAIGIFVSTSIEEVFSPPFYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVNL
 AIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFSDEEVQSPCKYKTLLRNENVLETAEALFNL
 NSIDLTHIFSHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEVQRSLKHEDINLQEIIISAAG
 KELSEAFKQKTSEILSHAHAALDQPLPTTMLKKQEEKEILKSQDSLLGLYHLLDWFADVDESNEVDPEF
 SARLTGIKLEMESLSPYNKARNYATKKPYSVEFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLY
 YLGIMPQKGRYKALSPEPTEKTSEGFDKMYDYFPDAAKMIPKCSTQLKAVTAHFQTHTPILLSNNF
 IEPLEITKEIYDLNNPEKEPKFQTYAGGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGE
 YYAELNPPLYHISFQRIAEKIMDAVETGKLYLFQIYNKDFAKGHGKPNLHTLYWTGLFSPENLAKTS
 IKLNGQAELFYRPKSRMKRMAHRLGEKMLNKKLDQKTPIDPLTLYQELYDYVNHRSLHDLSDEARAL
 LPNVITKEVSHEIIKDRRTSDKFFFHVPI TLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGERNL
 VIDSTGKILEQRSLNTIQQFDYQKQLDNREKERAARQAWSVVGTIKDLQGQLSQVIHEIVDLMHYQ
 VVVLENLNPGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAK
 MGTQSGFLFYVPAPYTSKIDPLTGFDPPFWKTIKNHESRKHFLEGFDLHYDVKTGDFILHFKMNRNL
 SFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIEHRFTGRYRDLYPANELIALEEKGIVF
 RDGSNILPKLLENDSSHAIITMVALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPMD
 ADANGAYHIALKGQLLNHLESKDLKLQNGISNQDWLAYIQELRN

113

In some embodiments, the guide nucleotide sequence-programmable DNA-binding protein domain of the present disclosure has no requirements for a PAM sequence. One example of such guide nucleotide sequence-programmable DNA-binding protein may be an Argonaute protein from *Natronobacterium gregoryi* (NgAgo). NgAgo is a ssDNA-guided endonuclease. NgAgo binds 5' phosphorylated ssDNA of ~24 nucleotides (gDNA) to guide it to its target site and will make DNA double-strand breaks at gDNA site. In contrast to Cas9, the NgAgo-gDNA system does not require a protospacer-adjacent motif (PAM). Using a nuclease inactive NgAgo (dNgAgo) can greatly expand the codons that may be targeted. The characterization and use of NgAgo have been described in Gao et al., *Nat Biotechnol.* *Epub* 2016 May 2. PubMed PMID: 27136078; Swarts et al., *Nature.* 507(7491) (2014):258-61; and Swarts et al., *Nucleic Acids Res.* 43(10) (2015):5120-9, each of which are incorporated herein by reference. The sequence of *Natronobacterium gregoryi* Argonaute is provided in SEQ ID NO: 270.

Wild type *Natronobacterium gregoryi* Argonaute
(SEQ ID NO: 270)
MTVIDLDSTTTADELTSGYTDISVTLTGVYDNTDEQHPRMSLAFEQDN
GERRYITLWKNTTPKDVFTYATGSTYIFTNIDYEVKDGYENLTATQY
TTVENATAQEVTGTTDEDETFAAGGEPLDHLLDALNETPDDAESDSGH
VMTSFASRDQLPEWTLHTYTLTADGAKTDTEYARRTLAYTVRQELEYTD
HDAAPVATDGLMLLTPEPLGETPLDLDCGVRVEADETRTLDDYTTAKDRL
LARELVEEGLKRSLSWDDYLVRGIDEVLSKEPVLTCDFFDLHERYDLSVE
VGHSGRAYLHINFRHRFPVKLTADIDDDNIYPGLRVKTYPRRRGHIV
WGLRDECATDSLNTLGNQSVVAYHRNNQTPINTDLLDAIEAADRRVVET
RRQGHGDDAVSFPQELLAVEPNTHQIKQFASDGFHQQARSKTRLSASRC
SEKAQAFAPAERLDPVRLNGSTVEFSSEFTGNNEQQQLRLLYENGESVLTF

114

-continued

RDGARGAHPDETFSGKIVNPPEFEVAVVLPEQQADTCKAQWDTMADLL
NQAGAPPTRSETVQYDAFSSPESISLNVAGAIDPSEVDAAFVVLPPDQE
5 GFADLASPTETYDELKKALANMGYIYSQMYFDRFRDAKIFYTRNVALGL
LAAAGGVAFTTEHAMPGDADMFIGIDVSRSYPEDGASGQINIAATATAV
YKDGTILGHSSTRPQLGKEKLQSTDVRDIMKNAILGYQQVTGESPTHIVI
10 HRDGFMNEDLDPATEFLNEQGVEYDIVEIRKQPQTRLLAVSDVQYDTPV
KSIAAINQNEPRATVATFGAPEYLATRDGGGLPRPIQIERVAGETDIET
LTRQVYLLSQSHIQVHNSTARLPITTAYADQASTHATKGYLVQTGAFES
15 NVGFL

Also provided herein are Cas9 variants that have relaxed PAM requirements (PAMless Cas9). PAMless Cas9 exhibits an increased activity on a target sequence that does not include a canonical PAM (e.g., NGG) at its 3'-end as compared to *Streptococcus pyogenes* Cas9 as provided by SEQ ID NO: 1, e.g., increased activity by at least 5-fold, at least 10-fold, at least 50-fold, at least 100-fold, at least 500-fold, at least 1,000-fold, at least 5,000-fold, at least 10,000-fold, at least 50,000-fold, at least 100,000-fold, at least 500,000-fold, or at least 1,000,000-fold. Such Cas9 variants that have relaxed PAM requirements are described in U.S. Provisional applications 62/245,828, 62/279,346, 62/311,763, 62/322,178, and 62/357,332, each of which is incorporated herein by reference. In some embodiments, the dCas9 or Cas9 nickase useful in the present disclosure may further comprise mutations that relax the PAM requirements, e.g., mutations that correspond to A262T, K294R, S409I, E480K, E543D, M694I, or E1219V in SEQ ID NO: 1.

Other on-limiting, exemplary Cas9 variants (including dCas9, Cas9 nickase, and Cas9 variants with alternative PAM requirements) suitable for use in the nucleobase editors described herein and their respective sequence are provided below.

VRER-nCas9 (D10A/D1135V/G1218R/R1335E/T1337R) *S. pyogenes* Cas9 Nickase
(SEQ ID NO: 821)
MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAATRLKRTAR
RRYTRRKRNRICYLQEIFSNEMAVKDDSFTHRLEESFLVEEDKKHERHPFGNIVDEVAYHEKYTIYHLR
KKLVDSTDKAIDLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEEKNGLFGNLIALSGLGTPNPKNSNFDAEAKLQLSKDTYDDDLDN
LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEK
YKEIFFDQSNGYAGYIDGGSQEEFYKFIKPILEKMDGTEELLVQLNRDLLRKQRTFDNGSIPHQIHL
GELHAILRQQEDFYPFLKDNRKREKIEKILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEVVDKGA
SAQSFIERMNTFDKLNPKNEVLPKHSLLYELYFTVYNELTKVVKYVTEGMRKP AFLSGEQKKAIVDLLFK
TNRKVTVQQLKEDYFKKIECFDSVEISGVEDRFNASLGYHDLLKIICKDKFLDNEENEDILEDIVLTLT
LFEDREMIEERLKTYAHLFDDKVMKQLKRRRTGWRSLRKLINGIRDQSGKTIIDFLKSDGFANRN
FMQLIHDDSLTFKFEDIQKAQVSGQGDLSHEHIANLAGSPAIIKGILQTVKVVDELVKVMGRHKPENIVI
EMARENQTTQKGOKNSRERMKRIEEGIKELGSQILKEHPVENTQLONEKLYLYLQNGRDMYVDQEL
DINRLSDYDVEDHVVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKROLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSD
FRKDFQFYKVERINNYHHAHDAYLNAVVGTLAKKYPKLESEFVYGDYKVDVRKMIAKSEQEIGKA

- continued

TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGDRFATVRKVLSMPQVNIVKKTEVQ
TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAVSVLVAKEVGKSKKLKSVKELLGITIME
RSSFEKNPIDFLEAKGYKEVKDIIKLPKSYLFELENGRKMLASARELQGNELALPSKYVNFLYLA
SHYEKLGSPEDNEQKQLFVEQHKYLDEIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENI
IHLFTLTNLGAPAAFKYFDTTIDRKKEYRSTKEVLDATLIHQSTITGLYETRIDSQLGGD
(single underline: HNH domain; double underline: RuvC domain)
VQR-nCas9 (D10A/D1135V/R1335Q/T1337R) *S. pyogenes* Cas9 Nickase
(SEQ ID NO: 822)
MDKKYSIGLAIGTNsvgwavitdeykvpskkfkvlgntdrhsikknligallfdsgetaeatrlkrtar
RRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR
KKLVDSTDKAIDLRIYLALAHMICKFRGHFLIEGDLNPDNSDVKDFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEAKLQLSKDTYDDLDN
LIAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQQLPEK
YKEIFFDQSCKNGYAGYIDGGASQEEFYKFIPILEKMDGTEELLVVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEVVVDKGA
SAQSFIERMTNFDKNLPNEKVLPHSLLYEFVYNELTKVKVYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAASLGTYHDLIKIIDKDFLDNEENEDILEDIVLTLT
LFEDREMIEERLKTYAHLFDDKVMQKLKRRRTGWGRSLRKLINGIRDQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVI
EMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMDYVQEL
DINRLSDYDVHDIVPQFLKDDSIDNKVLTRSDKNGKSDNVPSEEVVKMKNYWRQQLNAKLITORK
FDNLTKAERGGLSELDKAGFIKQQLTREVKVITLKSCLVSD
FRKDFQFYKVREINNNYHAHDAYLNAVVGTTALIKKPKLESEFVYGDYKVYDVRKMIAKSEQEIGKA
TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGDRFATVRKVLSMPQVNIVKKTEVQ
TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAVSVLVAKEVGKSKKLKSVKELLGITIME
RSSFEKNPIDFLEAKGYKEVKDIIKLPKSYLFELENGRKMLASAGELQGNELALPSKYVNFLYLA
SHYEKLGSPEDNEQKQLFVEQHKYLDEIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENI
IHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQSTITGLYETRIDSQLGGD
(single underline: HNH domain; double underline: RuvC domain)
EQR-nCas9 (D10A/D1135E/R1335Q/T1337R) *S. pyogenes* Cas9 Nickase
(SEQ ID NO: 823)
MDKKYSIGLAIGTNsvgwavitdeykvpskkfkvlgntdrhsikknligallledsgetaeatrlkrtar
RRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR
KKLVDSTDKAIDLRIYLALAHMICKFRGHFLIEGDLNPDNSDVKDFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEAKLQLSKDTYDDLDN
LIAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQQLPEK
YKEIFFDQSCKNGYAGYIDGGASQEEFYKFIPILEKMDGTEELLVVKLNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEVVVDKGA
SAQSFIERMTNFDKNLPNEKVLPHSLLYEFVYNELTKVKVYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAASLGTYHDLIKIIDKDFLDNEENEDILEDIVLTLT
LFEDREMIEERLKTYAHLFDDKVMQKLKRRRTGWGRSLRKLINGIRDQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVI

- continued

EMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMDYDQEL
DINRLSDYDVHDIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNWRLQLLNAKLITQRK
FDMLTKAERGGLELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVS D
FRKDFQFYKVREINNYHHAHDAYLNAAVGTLAKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKA
TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGDFATVRKVL SMPQVNIVKKTEVQ
TGGFSKESILPKRNSDKLIARKKDWDPKYGGFESP T VAYS VLVAKVEKGSKKLKSVKELLGITIME
RSSFEKNPIDFLEAKGYKEVKDLI I KLPKYS LFELENGRKRLASAGELQKGNE LALPSKYVNPL YLA
SHYEKLKGSPEDNEQKOLFVEQHKYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIRQAE NI
IHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATL IHQSITGLYETRIDLSQLGGD

 (single underline: HNH domain; double underline: RuvC domain)
 KKH-nCas9 (D10A/E782K/N968K/R1015H) *S. aureus* Cas9 Nickase
 (SEQ ID NO: 268)
 MKRNYI LGLAIGITSVGYGIIDYETRDVIDAGVRLPKEANVENNEGRRSKRGARRLKRRRRHRIQRVKK
 LLFDYNLLTDHS ELS GINPYEARVKGLSQKLSEEEPSAALLHLAKRRGVHN VNEEEDTGNELSTKEQI
 SRNSKALEEKYVAELQLERLKKDGEV RGSINRFKTS DYVKEAKQLLKVQKAYHQLDQS FIDTYIDLLE
 TRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSV KYAYNADLYNALNDNNLVITRDENE
 KLEYYEKFQI I ENVFKQKKP TLKQIAKEI LVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEII EN
 AELLDQIAKILTIYQSSEDIQEEELTNLNSELTQEEIEQISNLKG YTGT HNL S KAINL LDELWHTNDNQIA
 IFNRLKLVPKKVDSL SQKEIPTTLVDDFILSPVV KRSFI QSIKVINAI I KK YGLPN DIII ELAREKNSKDAQK
 MINEMQKRNRQTNERIEEI I RTGKENAKYLI EKIKLHDMQEGKCLYSLEAPILEDLLNNPFN YEV DHI IP
 RSVSF DNSFNNKVLVKQEE NSKKG NRT P FQYLS S SD SKI SYETFKKHILNLAKGKGRISKTKKEYLLEER
 DINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGTSFLRRKWF KKERNKGY
 KHHAE DALIIANAD FIF KEWK KLD KAK KV M ENQM FEEKQAESMPEIETEQEYKEI FITPHQIKH KDFK
 DYK YSHRVDKKPNRKLINDTLYSTRKDKGNTLIVNNLNGLYDKDNDKLKLINKSPEKLLMYHDP
 QTYQKLKLIMEQYQGDEKNPL YKYYEETGNYLTKYSKKDNGPV IKKIKYYGNKLNAHLDITDDYPNSR
 NKVV KLSLKPYRFDVYLDNGVYKFVTVKNLDV IKKENYYEVNSKCYEEAKLKKISNQAEFIASFYKN
 DLIKINGELYR VIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPHI IKTIASTQSIKKYSTDILGNLYE
 VKSKKHPQI I KKG
Streptococcus thermophilus CRISPR1 Cas9 (St1Cas9) Nickase (D9A)
 (SEQ ID NO: 269)
 MSDLVLG LAIGIGS VGV GILN KVTGEI I HKNSRIFPAQAENNLV RRTNRQG RRLTRRKKH RR VRLN RL
 FEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVH RGI SYLDDAS DDGNSIGDYAQIVK
 ENSKQLETKTPGQIQLERYQTYQGLRG DFTVEKDGGKHLRILNVFP TSAYRSEALRILQTOQQEFNPQITDE
 FINRYLEILT GKRKYYHGP GNEKSRTD YGRYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNL
 LNDLNNLTVP TETK KLSKEQK NQI INYV KNEKAMGP A KLFK YIA KLLSCD VADI KGYR IDKSGK AEI HT
 FEAYRKMK TLETLDIEQMDRETLDK LAYVLT LNT ERGI QEA LEHEFADGSFSQKV D E LVQFRKANS
 SIFGKGWHNF SVKLM MELIPELYETSEEQMTILTRLGKQKTTSSN KTKYIDEKLLTEE IYNPVVAKS VR
 QAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSV
 FHGHKQLATKIRLW HQQGERC LYTGKTISIHD LINN NSNQF EVDH ILPLS ITFDDSLANKVLVYATANQE
 KGQ RTPYQALDSMDDA WS FRELKA FVRESK TLSNKK KEYLL TEED ISKFDVRKKFIERNLV DTRYAS R
 TLVSYSEDQ LLDIETGELISDDEYKESVFKAPYQHFVDTLKSKF EDSI LFSYQVDSKFN RKSIDATIYAT
 RQAKVGKD KADETYVLGKIKD IYTQDGYD AFM KIYKKDKSKFLM YRHD PQT F EKVIEPILE NYPNQI
 NEKGKEVPCNPF LKYKEEHGYI RKYSKKG NGPEI KSLKYYDSKLG NHIDITPKD SNNKVV LQSVSPWR

- continued

```

ADVFNFNKTGKYEILGLKYADLQFEKGTGTYKISQEKYNDIKKKEGVSDSEFKFTLYKNDLLLKVDT
ETKEQQLFRFLSRMPKQKHYVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNI SIYKVRTD
VLGNQHIIKNEGDPKPLDF
Streptococcus thermophilus CRISPR3Cas9 (St3Cas9) Nickase (D10A)
(SEQ ID NO: 824)
MTKPYSIGLAIGTNNSVGWAVITDNYKVPSSKKMVKVLGNTSKYIKKNLLGVLLFDGITAEGRRLKRTA
RRRYTRRRNRIRYLQEIFSTEMATLDDAFFQRLLDSFLVPDDKRDSKYPFGNLVEEKVYHDEFPTIYHL
RKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQ
LEEIVKDKISKLEKKDRILKLFPGKEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHF SKESYDEDLET
LGYIGDDYSDVFLKAKKLYDAIILSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYN
EVFKDDTKNGYAGYIDGKTNQEDFYVYLNLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPQIHLQ
EMRAILDQAKFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKNEKITPWNFEDVIDKESS
AAEAFINRMTSFIDLYLPPEKVLPHSLLYETFNVNELTKVRFIAESMRDYQFLDSKQKKDIVRLYFKDK
RKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLLNIINDKEFLDDSSNEAIIIEEIHTLTIFEDRE
MIKQRLSKFENIFDKSVLKKLSRRHYTGWLGSKAKLINGIRDEKSGNTILDYLIIDDGISNRNFMQLIHDD
ALSFKKKIQKAQIIQGEDKGNIKEVVVKSLPGSPAIIKGILQSIKIVDELVKVMGRKPESIVVEMARENQ
YTQNQGKSNQQLRKLEKSLKELGSKILKENI PAKLSKIDNNNALQNDRLYLYLQNGKDMYTGDDLDI
DRLSNYDIDHITPQAFLKDNSIDNKVLVSSASN RGKSDDFPSLEVKKRKTFWYQLLKS KLISQRKFDNL
TKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDEKENNKDDENNRAVTVKIITLKLSTLVSQFRKD
FELYKVREINDPHHAHDAYLNAVIASALLKKYKPLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNI
FKKSISLADGRVIERPLIEVNEETGESVWNKESDLATVRRVLSYPQVNKKVEEQNHGLDRGPKGL
FMANLSSKPKPNNSNENLVGAKEYLDPKKYGGYAGISNSFAVLVKGTIEGAKKKITNVLEFQGISILDRI
NYRKDKLNFNLLEKGYKDI ELLIELPKYSLFELSDGSRRMLASILSTNNKRGEIHKGQIPLSQKFVKLLYH
AKRISNTINENHRKYVENHKKEFEELFYIILEFNENYVGAKKNGKLLNSAFQSWQNHSIDECCSFIGPT
GSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLKDATALIHQSVTGLYETRIDLAKLGE

```

In some embodiments, the nucleobase editors useful in the present disclosure comprises: (i) a guide nucleotide sequence-programmable DNA-binding protein domain; and (ii) a deaminase domain. In some embodiments, the deaminase domain of the fusion protein is a cytosine deaminase. In some embodiments, the deaminase is an APOBEC1 deaminase. In some embodiments, the deaminase is a rat APOBEC1. In some embodiments, the deaminase is a human APOBEC1. In some embodiments, the deaminase is an APOBEC2 deaminase. In some embodiments, the deaminase is an APOBEC3A deaminase. In some embodiments, the deaminase is an APOBEC3B deaminase. In some embodiments, the deaminase is an APOBEC3C deaminase. In some embodiments, the deaminase is an APOBEC3D deaminase. In some embodiments, is an APOBEC3F deaminase. In some embodiments, the deaminase is an APOBEC3G deaminase. In some embodiments, the deaminase is an APOBEC3H deaminase. In some embodiments, the deaminase is an APOBEC4 deaminase. In some embodiments, the deaminase is an activation-induced deaminase (AID). In some embodiments, the deaminase is a Lamprey CDA1 (pmCDA1). In some embodiments, the deaminase is a human APOBEC3G or a functional fragment thereof. In some embodiments, the deaminase is an APOBEC3G variant comprising mutations correspond to the D316R/D317R mutations in the human APOBEC3G. Exemplary, non-

limiting cytosine deaminase sequences that may be used in accordance with the methods of the present disclosure are provided in Example 1 below.

45 In some embodiments, the cytosine deaminase is a wild type deaminase or a deaminase as set forth in SEQ ID NOS: 271-292, 303, and 1072-1083. In some embodiments, the cytosine deaminase domains of the fusion proteins provided herein include fragments of deaminases and proteins homologous to a deaminase. For example, in some embodiments, a deaminase domain may comprise a fragment of the amino acid sequence set forth in any of SEQ ID NOS: 271-292, 303, and 1072-1083. In some embodiments, a deaminase domain comprises an amino acid sequence 50 homologous to the amino acid sequence set forth in any of SEQ ID NOS: 271-292, 303, and 1072-1083, or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in any of SEQ ID NOS: 271-292, 303, and 1072-1083. In some embodiments, proteins comprising a 55 deaminase, a fragments of a deaminase, or homologs of a deaminase or a deaminase are referred to as "deaminase variants." A deaminase variant shares homology to a deaminase, or a fragment thereof. For example a deaminase variant is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99%

121

identical, at least about 99.5% identical, or at least about 99.9% to a wild type deaminase or a deaminase as set forth in any of SEQ ID NOs: 271-292, 303, and 1072-1083. In some embodiments, the deaminase variant comprises a fragment of the deaminase, such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% to the corresponding fragment of wild type deaminase or a deaminase as set forth in any of SEQ ID NOs: 271-292, 303, and 1072-1083. In some embodiments, the cytosine deaminase is at least at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to an APOBEC3G variant as set forth in SEQ ID NO: 291 or SEQ ID NO: 292, and comprises mutations corresponding to the D316E/D317R mutations in SEQ ID NO: 290.

In some embodiments, the cytosine deaminase domain is fused to the N-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain. For example, the fusion protein may have an architecture of NH₂-[cytosine deaminase]-[guide nucleotide sequence-programmable DNA-binding protein domain]-COOH. The “[]” used in the general architecture above indicates the presence of an optional linker sequence. The term “linker,” as used herein, refers to a chemical group or a molecule linking two molecules or moieties, e.g., two domains of a fusion protein, such as, for example, a dCas9 domain and a cytosine deaminase domain. Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (e.g., a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

In some embodiments, the cytosine deaminase domain and the Cas9 domain are fused to each other via a linker. Various linker lengths and flexibilities between the deaminase domain (e.g., APOBEC1) and the Cas9 domain can be employed (e.g., ranging from very flexible linkers of the form (GGGS)_n (SEQ ID NO: 337), (GGGGS)_n (SEQ ID NO: 308), (GGS)_n (SEQ ID NO: 784), and (G)_n (SEQ ID NO: 783) to more rigid linkers of the form (EAAAK)_n (SEQ ID NO: 309), SGSETPGTSESATPES (SEQ ID NO: 310) (see, e.g., Guilinger et al., *Nat. Biotechnol.* 2014; 32(6): 577-82; the entire contents are incorporated herein by reference), (SGGS)_nSGSETPGTSESATPES(SGGS)_n (SEQ ID NO: 1068), (XP)₁ (SEQ ID NO: 785), or a combination of any of these, wherein X is any amino acid and n is independently an integer between 1 and 30, in order to achieve the optimal length for deaminase activity for the specific application. In some embodiments, n is independently 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30, or, if more than one linker or more than one linker motif is present, any combination thereof. In some embodiments, the linker comprises a (GGS)_n (SEQ ID

122

NO: 784) motif, wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15. In some embodiments, the linker comprises a (GGS)_n (SEQ ID NO: 784) motif, wherein n is 1, 3, or 7. In some embodiments, the linker comprises the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 310), also referred to as the XTEN linker. In some embodiments, the linker comprises an amino acid sequence chosen from the group including, but not limited to, AGVF (SEQ ID NO: 825), GFLG (SEQ ID NO: 826), FK, AL, ALAL (SEQ ID NO: 827), or ALALA (SEQ ID NO: 828). In some embodiments, suitable linker motifs and configurations include those described in Chen et al., *Fusion protein linkers: property, design and functionality*. *Adv Drug Deliv Rev.* 2013; 65(10):1357-69, which is incorporated herein by reference. In some embodiments, the linker may comprise any of the following amino acid sequences: VPFLLEPDN-INGKTC (SEQ ID NO: 311), GSAGSAAGSGEF (SEQ ID NO: 312), SIVAQLSRPDPA (SEQ ID NO: 313), MKIIIEQLPSA (SEQ ID NO: 314), VRHKLKRVGS (SEQ ID NO: 315), GHGTGSTGSGSS (SEQ ID NO: 316), MSRPDPA (SEQ ID NO: 317), GSAGSAAGSGEF (SEQ ID NO: 312), SGSETPGTSESA (SEQ ID NO: 318), SGSETPGTSESATPEGGSGGSS (SEQ ID NO: 319), or GGSM (SEQ ID NO: 320). Additional suitable linker sequences will be apparent to those of skill in the art based on the instant disclosure.

In some embodiments, the nucleobase editor comprises a guide nucleotide sequence-programmable DNA-binding protein domain and an apolipoprotein B mRNA-editing complex 1 (APOBEC1) deaminase domain, where the deaminase domain is fused to the N-terminus of the napD-NAbp domain via a linker comprising the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 310). In some embodiments, the a guide nucleotide sequence-programmable DNA-binding protein domain comprises the amino acid sequence of any of the a guide nucleotide sequence-programmable DNA-binding protein domains provided herein. In some embodiments, the deaminase is rat APOBEC1 (SEQ ID NO: 288). In some embodiments, the deaminase is human APOBEC1 (SEQ ID NO: 286). In some embodiments, the deaminase is pmCDA1 (SEQ ID NO: 289). In some embodiments, the deaminase is human APOBEC3G (SEQ ID NO: 279). In some embodiments, the deaminase is a human APOBEC3G variant of any one of (SEQ ID NOs: 290-292). In some embodiments, the fusion protein comprises a guide nucleotide sequence-programmable DNA-binding protein domain and an apolipoprotein B mRNA-editing complex 1 catalytic polypeptide-like 3G (APOBEC3G) deaminase domain, wherein the deaminase domain is fused to the N-terminus of the a guide nucleotide sequence-programmable DNA-binding protein domain via a linker of any length or composition (e.g., an amino acid sequence, a peptide, a polymer, or a bond). In some embodiments, the linker comprises the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 310). In some embodiments, the linker comprises the amino acid sequence (SGGS)₂SGSETPGTSESATPES(SGGS)₂ (SEQ ID NO: 1069).

In some embodiments, the fusion protein comprises a guide nucleotide sequence-programmable DNA-binding protein domain and a cytidine deaminase 1 (CDA1) deaminase domain, wherein the deaminase domain is fused to the N-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain via a linker comprising the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 310). In some embodiments, the linker comprises the amino acid sequence (SGGS)₂SGSETPGTSESATPES

123

(SGGS)₂ (SEQ ID NO: 1069). In some embodiments, the guide nucleotide sequence-programmable DNA-binding protein domain comprises the amino acid sequence of any of the guide nucleotide sequence-programmable DNA-binding protein domains provided herein.

In some embodiments, the fusion protein comprises a guide nucleotide sequence-programmable DNA-binding protein and an activation-induced cytidine deaminase (AID) deaminase domain, where the deaminase domain is fused to the N-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain via a linker comprising the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 310). In some embodiments, the linker comprises the amino acid sequence (SGGS)₂SGSETPGTSESATPES (SGGS)₂ (SEQ ID NO: 1069). In some embodiments, the guide nucleotide sequence-programmable DNA-binding protein comprises the amino acid sequence of any of the guide nucleotide sequence-programmable DNA-binding protein domains provided herein.

Some aspects of the disclosure are based on the recognition that certain configurations of a guide nucleotide sequence-programmable DNA-binding protein, and a cytidine deaminase domain fused by a linker are useful for efficiently deaminating target cytidine residues. Other aspects of this disclosure relate to the recognition that a nucleobase editing fusion protein with an apolipoprotein B mRNA-editing complex 1 (APOBEC1) deaminase domain fused to the N-terminus of a guide nucleotide sequence-programmable DNA-binding protein via a linker comprising the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 310) was capable of efficiently deaminating target nucleic acids in a double stranded DNA target molecule. In some embodiments, the fusion protein comprises a guide nucleotide sequence-programmable DNA-binding protein domain and an apolipoprotein B mRNA-editing complex 1 (APOBEC1) deaminase domain, where the deaminase domain is fused to the N-terminus of the napDNAbp via a linker comprising the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 310). In some embodiments, the fusion protein comprises a guide nucleotide sequence-programmable DNA-binding protein domain and an apolipoprotein B mRNA-editing complex 1 (APOBEC1) deaminase domain, where the deaminase domain is fused to the N-terminus of the napDNAbp via a linker comprising the amino acid sequence (SGGS)₂SGSETPGTSESATPES(SGGS)₂ (SEQ ID NO: 1069).

To successfully edit the desired target C base, the linker between Cas9 and APOBEC may be optimized, as described in Komor et al., *Nature*, 533, 420-424 (2016), which is incorporated herein by reference. The numbering scheme for base editing is based on the predicted location of the target C within the single stranded stretch of DNA (R-loop) displaced by a programmable guide RNA sequence occurring when a DNA-binding domain (e.g. Cas9, nCas9, dCas9) binds a genomic site (see FIG. 6). Conveniently, the sequence immediately surrounding the target C also matches the sequence of the guide RNA. The numbering scheme for base editing is based on a standard 20-mer programmable sequence, and defines position “21” as the first DNA base of the PAM sequence, resulting in position “1” assigned to the first DNA base matching the 5'-end of the 20-mer programmable guide RNA sequence. Therefore, for all Cas9 variants, position “21” is defined as the first base of the PAM sequence (e.g. NGG, NGAN, NGNG, NGAG, NGCG, NNGRRT, NGRRN, NNNRRT, NNNGATT, NNAGAA, NAAAC). When a longer programmable guide RNA sequence is used (e.g. 21-mer) the 5'-end bases are assigned

124

a decreasing negative number starting at “-1”. For other DNA-binding domains that differ in the position of the PAM sequence, or that require no PAM sequence, the programmable guide RNA sequence is used as a reference for numbering. A 3-aa linker gives a 2-5 base editing window (e.g., positions 2, 3, 4, or 5 relative to the PAM sequence at position 21). A 9-aa linker gives a 3-6 base editing window (e.g., positions 3, 4, 5, or 6 relative to the PAM sequence at position 21). A 16-aa linker (e.g., the SGSETPGTSESATPES (SEQ ID NO: 310) linker) gives a 4-7 base editing window (e.g., positions 4, 5, 6, or 7 relative to the PAM sequence at position 21). A 21-aa linker gives a 5-8 base editing window (e.g., positions 5, 6, 7, 8 relative to the PAM sequence at position 21). Each of these windows can be useful for editing different targeted C bases. For example, the targeted C bases may be at different distances from the adjacent PAM sequence, and by varying the linker length, the precise editing of the desired C base is ensured. One skilled in the art, based on the teachings of CRISPR/Cas9 technology, in particular the teachings of U.S. Provisional applications 62/245,828, 62/279,346, 62/311,763, 62/322, 178, 62/357352, 62/370,700, and 62/398,490, and in Komor et al., *Nature*, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage, 533, 420-424 (2016), each of which is incorporated herein by reference, will be able to determine the window of editing for his/her purpose, and properly design the linker of the cytosine deaminase-dCas9 protein for the precise targeting of the desired C base.

To successfully edit the desired target C base, appropriate Cas9 domain may be selected to attached to the deaminase domain (e.g., APOBEC1), since different Cas9 domains may lead to different editing windows, as described in in U.S. Pat. No. 9,068,179, US Patent Application Publications US 2015/0166980, US 2015/0166981, US 2015/0166982, US20150166984, and US20150165054, and US Provisional Applications, U.S. Ser. No. 62/245,828, filed Oct. 23, 2015; 62/279,346, filed Jan. 15, 2016; 62/311,763, filed Mar. 22, 2016; 62/322,178, filed Apr. 13, 2016, 62/357,352, filed Jun. 30, 2016, U.S. Pat. No. 62,370,700, filed Aug. 3, 2016; 62/398,490, filed Sep. 22, 2016; 62/408,686, filed Oct. 14, 2016; PCT Application PCT/US2016/058344, filed Oct. 22, 2016, U.S. patent application Ser. No. 15/311,852, filed Oct. 22, 2016; and in Komor et al., *Nature*, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage, 533, 420-424 (2016), the entire contents of each of which are incorporated herein by reference. For example, APOBEC1-XTEN-SaCas9n-UGI gives a 1-12 base editing window (e.g., positions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 relative to the NNNRRT PAM sequence in positions 20-26). One skilled in the art, based on the teachings of CRISPR/Cas9 technology, will be able to determine the editing window for his/her purpose, and properly determine the required Cas9 homolog and linker attached to the cytosine deaminase for the precise targeting of the desired C base.

In some embodiments, the fusion protein useful in the present disclosure further comprises a uracil glycosylase inhibitor (UGI) domain. A “uracil glycosylase inhibitor” refers to a protein that inhibits the activity of uracil-DNA glycosylase. The C to T base change induced by deamination results in a U:G heteroduplex, which triggers cellular DNA-repair response. Uracil DNA glycosylase (UDG) catalyzes removal of U from DNA in cells and initiates base excision repair, with reversion of the U:G pair to a C:G pair as the most common outcome. Thus, such cellular DNA-repair response may be responsible for the decrease in nucleobase

125

editing efficiency in cells. Uracil DNA Glycosylase Inhibitor (UGI) is known in the art to potently blocks human UDG activity. As described in Komor et al., Nature (2016), fusing a UGI domain to the cytidine deaminase-dCas9 fusion protein reduced the activity of UDG and significantly enhanced editing efficiency.

Suitable UGI protein and nucleotide sequences are provided herein and additional suitable UGI sequences are known to those in the art, and include, for example, those published in Wang et al., Uracil-DNA glycosylase inhibitor gene of bacteriophage PBS2 encodes a binding protein specific for uracil-DNA glycosylase. J. Biol. Chem. 264: 1163-1171(1989); Lundquist et al., Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein. Role of specific carboxylic amino acids in complex formation with *Escherichia coli* uracil-DNA glycosylase. J. Biol. Chem. 272:21408-21419(1997); Ravishankar et al., X-ray analysis of a complex of *Escherichia coli* uracil DNA glycosylase (EcUDG) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG. Nucleic Acids Res. 26:4880-4887(1998); and Putnam et al., Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with *Escherichia coli* uracil-DNA glycosylase. J. Mol. Biol. 287:331-346(1999), each of which is incorporated herein by reference. In some embodiments, the UGI domain comprises the amino acid sequence of SEQ ID NO: 304 without the N-terminal methionine (M). In some embodiments, the UGI comprises the following amino acid sequence: *Bacillus* phage PBS2 (Bacteriophage PBS2)Uracil-DNA glycosylase inhibitor MTNLSDIIEKETGKQLVIQE-
SILMLPPEEVEEVIGNKPESDILVHTAYDEST-
DENVMLLTSDAPEYKPWAL VIQDSNGENIKML
(SEQ ID NO: 304)

In some embodiments, the UGI protein comprises a wild type UGI or a UGI as set forth in SEQ ID NO: 304. In some embodiments, the UGI proteins useful in the present disclosure include fragments of UGI and proteins homologous to a UGI or a UGI fragment. For example, in some embodiments, a UGI comprises a fragment of the amino acid sequence set forth in SEQ ID NO: 304. In some embodiments, a UGI comprises an amino acid sequence homologous to the amino acid sequence set forth in SEQ ID NO: 304 or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in SEQ ID NO: 304. In some embodiments, proteins comprising UGI or fragments of UGI or homologs of UGI or UGI fragments are referred to as "UGI variants." A UGI variant shares homology to UGI, or a fragment thereof. For example a UGI variant is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% to a wild type UGI or a UGI as set forth in SEQ ID NO: 304. In some embodiments, the UGI variant comprises a fragment of UGI, such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% to the corresponding fragment of wild type UGI or a UGI as set forth in SEQ ID NO: 304.

It should be appreciated that additional proteins may be uracil glycosylase inhibitors. For example, other proteins that are capable of inhibiting (e.g., sterically blocking) a uracil-DNA glycosylase base-excision repair enzyme are

126

within the scope of this disclosure. In some embodiments, a uracil glycosylase inhibitor is a protein that binds DNA. In some embodiments, a uracil glycosylase inhibitor is a protein that binds single-stranded DNA. For example, a uracil glycosylase inhibitor may be a *Erwinia tasmaniensis* single-stranded binding protein. In some embodiments, the single-stranded binding protein comprises the amino acid sequence (SEQ ID NO: 305). In some embodiments, a uracil glycosylase inhibitor is a protein that binds uracil. In some embodiments, a uracil glycosylase inhibitor is a protein that binds uracil in DNA. In some embodiments, a uracil glycosylase inhibitor is a catalytically inactive uracil DNA-glycosylase protein. In some embodiments, a uracil glycosylase inhibitor is a catalytically inactive uracil DNA-glycosylase protein that does not excise uracil from the DNA. For example, a uracil glycosylase inhibitor is a UdgX. In some embodiments, the UdgX comprises the amino acid sequence (SEQ ID NO: 306). As another example, a uracil glycosylase inhibitor is a catalytically inactive UDG. In some embodiments, a catalytically inactive UDG comprises the amino acid sequence (SEQ ID NO: 307). It should be appreciated that other uracil glycosylase inhibitors would be apparent to the skilled artisan and are within the scope of this disclosure.

Erwinia tasmaniensis SSB (themostable single-stranded DNA binding protein)

(SEQ ID NO: 305)

MASRGVINKVILVGNLQGDPEVRYMPNNGAVANITLATSESWRDQKQTGET

KEKTEWHRVVLFGKLAEGAVEGEYLRLKGSOVYIEGALQTRKWTQAGVEKY

TTEVVNVGGTMQMLGGRSQGGGASAGGQNNGSNNGWQQPQQPQGGNQF

SGGAQQQARPQQQPQQNNAPANNEPPIDEDDDIP

UdgX (binds to Uracil in DNA but does not excise)
(SEQ ID NO: 306)

MAGAQDFVPHADLAELAAAAGECRGCGLYRDATAQAVFGAGGRSARI MM

IGEQPGDKEDLAGLPFVGPAAGRLLDRALEAADIDRDALYVTNAVHKFKF

TRAAGGKRRIHKTPSRTEVVACRPWLIAEMTSVEPDVVVLLGATAAKAL

LGNDFRVTQHHRGEVLHVDDVPGDPALVATVHPSSLLRGPKEERESAFAG

LVDDLRLVAADVRP

UDG (catalytically inactive human UDG, binds to Uracil in DNA but does not excise)
(SEQ ID NO: 307)

MIGQKTLYSFFSPSPARKRHPSPPEPAVQGTGVAGVPEESGDAAAIPAK

KAPAGQEEPGTTPSSPLSAEQLDRIQRNKAAALLRLAARNVPVGFGESW

KKHLSGEFGKPYFIKLMGFVAEERKHYTVYPHQQFTWTQMCDIKDVK

VVILGQEPIHGPNQAHGLCFSVQRPVPPPSLENIYKELSTDIEDFVHP

GHGDLSGWAKQGVLLNAVLTVRAHQANSRKERGWEOFDAVWSLNQN

SNGLVFLWGSYAQKKGSAIDRKRRHHVLTQTAHPSPSLSVYRGFFGCRHFS

KTNELLQKSGKKPIDWKEL

In some embodiments, the UGI domain is fused to the C-terminus of the dCas9 domain in the fusion protein. Thus, the fusion protein would have an architecture of NH₂-[cytosine deaminase]-[guide nucleotide sequence-programmable DNA-binding protein domain]-[UGI]-COOH. In some embodiments, the UGI domain is fused to the N-terminus of the cytosine deaminase domain. As such, the fusion protein would have an architecture of NH₂-[UGI]-[cytosine deaminase]-[guide nucleotide sequence-programmable

DNA-binding protein domain]-COOH. In some embodiments, the UGI domain is fused between the guide nucleotide sequence-programmable DNA-binding protein domain and the cytosine deaminase domain. As such, the fusion protein would have an architecture of NH₂-[cytosine deaminase]-[UGI]-[guide nucleotide sequence-programmable DNA-binding protein domain]-COOH. The linker sequences described herein may also be used for the fusion of the UGI domain to the cytosine deaminase-dCas9 fusion proteins.

In some embodiments, the fusion protein comprises the structure:

[cytosine deaminase]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA binding protein]-[optional linker sequence]-[UGI];
 [cytosine deaminase]-[optional linker sequence]-[UGI]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA binding protein];
 [UGI]-[optional linker sequence]-[cytosine deaminase]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA binding protein];
 [UGI]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA binding protein]-[optional linker sequence]-[cytosine deaminase];
 [guide nucleotide sequence-programmable DNA binding protein]-[optional linker sequence]-[cytosine deaminase]-[optional linker sequence]-[UGI]; or
 [guide nucleotide sequence-programmable DNA binding protein]-[optional linker sequence]-[UGI]-[optional linker sequence]-[cytosine deaminase].

In some embodiments, the fusion protein comprises the structure:

[cytosine deaminase]-[optional linker sequence]-[Cas9 nickase]-[optional linker sequence]-[UGI];
 [cytosine deaminase]-[optional linker sequence]-[UGI]-[optional linker sequence]-[Cas9 nickase];
 [UGI]-[optional linker sequence]-[cytosine deaminase]-[optional linker sequence]-[Cas9 nickase];
 [UGI]-[optional linker sequence]-[Cas9 nickase]-[optional linker sequence]-[cytosine deaminase];
 [Cas9 nickase]-[optional linker sequence]-[cytosine deaminase]-[optional linker sequence]-[UGI]; or
 [Cas9 nickase]-[optional linker sequence]-[UGI]-[optional linker sequence]-[cytosine deaminase].

In some embodiments, fusion proteins provided herein further comprise a nuclear localization sequence (NLS). In some embodiments, the NLS is fused to the N-terminus of the fusion protein. In some embodiments, the NLS is fused to the C-terminus of the fusion protein. In some embodiments, the NLS is fused to the N-terminus of the UGI protein. In some embodiments, the NLS is fused to the C-terminus of the UGI protein. In some embodiments, the NLS is fused to the N-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain. In some embodiments, the NLS is fused to the C-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain. In some embodiments, the NLS is fused to the N-terminus of the cytosine deaminase. In some embodiments, the NLS is fused to the C-terminus of the deaminase. In some embodiments, the NLS is fused to the fusion protein via one or more linkers. In some embodiments, the NLS is fused to the fusion protein without a linker. Non-limiting, exemplary NLS sequences may be PKKKRKV (SEQ ID NO: 829) or MDSLLMNRKFLYQFKNVRWAKGRRE-TYLC (SEQ ID NO: 830).

In some embodiments, any of the fusion proteins provided herein comprise a second UGI domain. Fusion proteins

comprising two UGI domains are described in U.S. Provisional Applications, U.S. Ser. No. 62/475,830, filed Mar. 23, 2017; 62/490,587; 62/511,934, filed May 26, 2017; 62/551,951, filed Aug. 30, 2017; and Komor et al. (2017) Improved 5 Base Excision Repair Inhibition and Bacteriophage Mu Gam Protein Yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv*, 3: eaao4774; the entire contents of which is incorporated herein by reference. In some embodiments, the second UGI domain comprises a 10 wild-type UGI or a UGI as set forth in SEQ ID NO: 304. In some embodiments, the UGI proteins provided herein include fragments of UGI and proteins homologous to a UGI or a UGI fragment. For example, in some embodiments, the second UGI domain comprises a fragment of the amino acid sequence set forth in SEQ ID NO: 304. In some embodiments, a UGI fragment comprises an amino acid sequence that comprises at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid sequence as set forth in SEQ ID NO: 304. In some embodiments, the second UGI domain comprises an amino acid sequence homologous to the amino acid sequence set forth in SEQ ID NO: 304 or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in SEQ ID NO: 304. In some embodiments, proteins comprising UGI or fragments of UGI or homologs of UGI or UGI fragments are referred to as “UGI variants.” A UGI variant shares homology to UGI, or a fragment thereof. For example a UGI variant is at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 96% identical, at least 97% identical, at least 98% identical, at least 99% identical, at least 99.5% identical, or at least 99.9% identical to a wild type UGI or a UGI as set forth in SEQ ID NO: 304. In some embodiments, the UGI variant comprises a fragment of UGI, such that the fragment is at least 70% identical, at least 80% identical, at least 90% identical, at least 95% identical, at least 96% identical, at least 97% identical, at least 98% identical, at least 99% identical, at least 99.5% identical, or at least 99.9% to the corresponding fragment of wild-type UGI or a UGI as set forth in SEQ ID NO: 304.

In some embodiments, the fusion protein comprises the structure:

[deaminase]-[optional linker sequence]-[dCas9]-[optional linker sequence]-[first UGI]-[optional linker sequence]-[second UGI];
 [deaminase]-[optional linker sequence]-[Cas9 nickase]-[optional linker sequence]-[first UGI]-[optional linker sequence]-[second UGI]; or
 [deaminase]-[optional linker sequence]-[Cas9]-[optional linker sequence]-[first UGI]-[optional linker sequence]-[second UGI].

In some embodiments, the nucleobase editor comprises a 55 guide nucleotide sequence-programmable DNA-binding protein domain and an apolipoprotein B mRNA-editing complex 1 (APOBEC1) deaminase domain, wherein the deaminase domain is fused to the N-terminus of the guide nucleotide sequence-programmable DNA-binding protein domain via a linker comprising the amino acid sequence (SGGS)₂SGSETPGTSESATPES(SGGS)₂ (SEQ ID NO: 1069). In some embodiments, the a guide nucleotide sequence-programmable DNA-binding protein domain comprises the amino acid sequence of any of the a guide 60 nucleotide sequence-programmable DNA-binding protein domains provided herein. In some embodiments, the deaminase is rat APOBEC1 (SEQ ID NO: 288). In some embodiment 65

129

ments, the deaminase is human APOBEC1 (SEQ ID NO: 286). In some embodiments, the deaminase is a human APOBEC3G variant of any one of (SEQ ID NOs: 290-292). In some embodiments, the nucleobase editor comprises a first UGI domain fused to the C-terminus of a guide nucleotide sequence-programmable DNA-binding protein domain via a linker comprising the amino acid sequence (GGS)_n (SEQ ID NO: 784), wherein n is 3. In some embodiments, the nucleobase editor comprises a second UGI domain fused to the C-terminus of a first UGI domain via a linker comprising the amino acid sequence (GGS)_n (SEQ ID NO: 784), wherein n is 3.

In some embodiments, the fusion protein comprises the amino acid sequence of SEQ ID NO: 1084. In some embodiments, the fusion protein comprises an amino acid sequence that is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence as set forth in SEQ ID NO: 1084.

In some embodiments, any of the fusion proteins provided herein may further comprise a Gam protein. The term “Gam protein,” as used herein, refers generally to proteins capable

130

of binding to one or more ends of a double strand break of a double stranded nucleic acid (e.g., double stranded DNA). In some embodiments, the Gam protein prevents or inhibits degradation of one or more strands of a nucleic acid at the site of the double strand break. In some embodiments, a Gam protein is a naturally-occurring Gam protein from bacteriophage Mu, or a non-naturally occurring variant thereof. Fusion proteins comprising Gam proteins are described in Komor et al. (2017) Improved Base Excision Repair Inhibition and Bacteriophage Mu Gam Protein Yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv*, 3: eaao4774; the entire contents of which is incorporated by reference herein. In some embodiments, the Gam protein comprises an amino acid sequence that is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence provided by SEQ ID NO: 3027. In some embodiments, the Gam protein comprises the amino acid sequence of SEQ ID NO: 3027. In some embodiments, the fusion protein (e.g., BE4-Gam of SEQ ID NO: 3028) comprises a Gam protein, wherein the Cas9 domain of BE4 is replaced with any of the Cas9 domains provided herein.

Gam from bacteriophage Mu:

(SEQ ID NO: 1070)

AKPAKRIKSAAAAYVPQNRDAVITDIKRIGDLQREASRLETEMNDAIAEITEKFAARIAPIKTIDETLSKGVQGW
CEANRDELTNGGKVKTANLVTDVSWRVRPPSVSIRGMDAVMETLERLGLQRFIRTQEIINKEAILLEPKAVAGV
AGITVKSGIEDFSIIPFEQEAGI

BE4 -Gam:

(SEQ ID NO: 1071)

MAKPAKRIKSAAAAYVPQNRDAVITDIKRIGDLQREASRLETEMNDAIAEITEKFAARIAPIKTIDETLSKGVQGW
WCEANRDELTNGGKVKTANLVTDVSWRVRPPSVSIRGMDAVMETLERLGLQRFIRTQEIINKEAILLEPKAVAG
VAGITVKSGIEDFSIIPFEQEAGISGSETPGTSESATPESSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKET
CLLYEINWGGRHSIWRHTSQNTNKHVEVNFIIEKFTTERYFCPNTRCSITWFLSWSPCGECRAITEFLSRYPHV
LFIFYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNSNEAHWPRPHLWVRLYVLEYCII
LGLPPCLNLRRKQPQLTFITALQSCHYQRLPPHILWATGLKSGGGSSGGSETPGTSESATPESSGGSSGG
DKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAETRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLDSTDKA
RLIYLALAHMIKFRGHFLIEGDLNPNDNSVDKLFIQLVQTYNQLEENPINASGVDAKAILSARLSKSRRLENLI
AQLPGEKKNLFGNLIALSGLTPNFKSNFDLAEDAKLQLSKDTYDDLDNLLAQIGDQYADLFLAKNLSDAIL
LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGSQEEFYKFI
KPILEKMDGEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAI LRRQEDFYPFLKDNRREKIEKILTFRIPYY
VGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFERMTNFDKNLPNEKVLPHSLLYEYPTVYNELTK
VKVYTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYPKKIECFDSVIEGVSVDRENASLGTYHDLLKII
DKDKDFLDNEENEDILEDIVTLTLEFREMIERLKTYAHLEDDKVMKQLKRRRTGWGRLSRKLINGIRDQSG
KTILDPLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGLQTVKVVDELVKM
GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNQGRDMYVDQ
ELDINRLSDYDVHDIVPQSLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFDNLT
KAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVR
EINNYHHAHDAYLNAVVGTLAKKPKLESEFVYGDYKVDVRKMIAKSEQEIGKATAKYFFYSNIMNNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW

- continued

```

DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITMERSSFEKNPIDFLEAKGYKEVKKDLIKL PKY
SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLGS PEDNEQKQLFVEQHKHYLDEII EQI SEF
SKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRTSTKEVLDATLHQS
ITGLYETRIDSQLGGDGGSGGSGG STNLS DII EKETGKQLVI QESILMLP EEEV EIGNK PESDILVHTAYDE
STDENVMLLTS DAPEYKPW ALVI QDSNGEN KIKMLSGGSGGSGG STNLS DII EKETGKQLVI QESILMLP EEEV E
VIGNK PESDILVHTAYDE STDENVMLLTS DAPEYKPW ALVI QDSNGEN KIKMLSGGSPKKRK

```

Some aspects of the present disclosure provide nucleobase editors associated with a guide nucleotide sequence (e.g., a guide RNA or gRNA). gRNAs can exist as a complex of two or more RNAs, or as a single RNA molecule. gRNAs that exist as a single RNA molecule may be referred to as single-guide RNAs (sgRNAs), though “gRNA” is used interchangeably to refer to guide RNAs that exist as either single molecules or as a complex of two or more molecules. Typically, gRNAs that exist as a single RNA species comprise two domains: (1) a domain that shares homology to a target nucleic acid (e.g., and directs binding of the Cas9 complex to the target); and (2) a domain that binds the Cas9 protein. In some embodiments, domain (2) corresponds to a sequence known as a tracrRNA, and comprises a stem-loop structure. For example, in some embodiments, domain (2) is identical or homologous to a tracrRNA as provided in Jinek et al., *Science* 337:816-821(2012), which is incorporated herein by reference. Other examples of gRNAs (e.g., those including domain 2) can be found in U.S. Provisional Patent Application, U.S. Ser. No. 61/874,682, filed Sep. 6, 2013, entitled “Switchable Cas9 Nucleases And Uses Thereof,” and U.S. Provisional Patent Application, U.S. Ser. No. 61/874,746, filed Sep. 6, 2013, entitled “Delivery System For Functional Nucleases,” each of which is incorporated herein by reference in their entirety. The gRNA comprises a nucleotide sequence that complements a target site, which mediates binding of the nuclease/RNA complex to said target site, providing the sequence specificity of the nucleic acid:RNA complex. These proteins are able to be targeted, in principle, to any sequence specified by the guide RNA.

Methods of using RNA-programmable nucleases, such as Cas9, for site-specific cleavage (e.g., to modify a genome) are known in the art (see e.g., Cong, L. et al. *Science* 339, 819-823 (2013); Mali, P. et al. *Science* 339, 823-826 (2013); Hwang, W. Y. et al. *Nature biotechnology* 31, 227-229 (2013); Jinek, M. et al. *eLife* 2, e00471 (2013); Dicarlo, J. E. et al. *Nucleic acids research* (2013); Jiang, W. et al. *Nature Biotechnology* 31, 233-239 (2013); each of which are incorporated herein by reference). In particular, examples of guide nucleotide sequences (e.g., sgRNAs) that may be used to target the fusion protein of the present disclosure to its target sequence to deaminate the targeted C bases are described in Komor et al., *Nature*, 533, 420-424 (2016), which is incorporated herein by reference.

The specific structure of the guide nucleotide sequences (e.g., sgRNAs) depends on its target sequence and the relative distance of a PAM sequence downstream of the target sequence. One skilled in the art will understand, that no unifying structure of guide nucleotide sequence is given, for that the target sequences are different for each and every C targeted to be deaminated.

However, the present disclosure provides guidance in how to design the guide nucleotide sequence, e.g., an sgRNA, so that one skilled in the art may use such teaching to a target sequence of interest. An gRNA typically comprises a tracrRNA framework allowing for Cas9 binding, and a guide sequence, which confers sequence specificity to fusion proteins disclosed herein. In some embodiments, the guide RNA comprises a structure 5'-[guide sequence]-tracrRNA-3'. Non-limiting, exemplary tracrRNA sequences are shown in Table 11.

TABLE 11

TracrRNA orthologues and sequences		
Organism	tracrRNA sequence	SEQ ID NO
<i>S. pyogenes</i>	GUUUUAGAGCUAUGCUGGAAAGCCACGGUGAAAA GUUCAACUAUUGCCUGAU CGGAUAAA UUUGAACG AUACGACAGUCGGUGCUUUUUUU	322
<i>S. pyogenes</i>	GUUUUAGAGCUAGAAAAGCAAGUUAAAAGGC UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU	323
<i>S. thermophilus</i> CRISPR1	GUUUUUGUACUCUCAAGAUCAAAUACUUGCAGA AGCUACAAGAUAGGCUCUAUGCGAAUCAACA CCCUGUCAUUUAUGGCAGGGUGUUUU	324
<i>S. thermophilus</i> CRISPR3	GUUUUAGAGCUGUGUUGUUGUAAAACAACACAG CGAGUAAAAGGCUUAGUCGGUACUCAACUUG AAAAGGUGGCACCGAUUCGGUGUUUU	325
<i>C. jejuni</i>	AAGAAA UUUAAAAGGGACUAAAAGAGUUUG CGGGACUCUGCGGGGUUACAUCCCUAAAACCGCU UUU	326

TABLE 11-continued

TracrRNA orthologues and sequences		
Organism	tracrRNA sequence	SEQ ID NO
<i>F. novicida</i>	AUCUAAAUAUAAAUGUACCAAAUAAAUGCU CUGUAUCAUUAAAAGUAUUUGAACGGACCUCU GUUUGACACGUCUGAAUACUAAAA	327
<i>S. thermophilus2</i>	UGUAAGGGACGCCUUACACAGUUACUAAAUCUUG CAGAAGCUACAAGAUAAAGGUUCUCAUGCAGCGAAAUC AACACCCUGCUAUUUUAUGGCAGGGUGUUUUCGUU AUUU	328
<i>M. mobile</i>	UGUAUUUCGAAAACAGAUGUACAGUUAGAAAUC AUAAGAAUAGAUACAUACACUAAAAGGGUUUUAUG CCGUAAACUACUACUAAAAGGUAGUAGUU UUUU	329
<i>L. innocua</i>	AUUGUUAGUAUCAAAAACAUAGCAAGUUAAAA UAAGGCUUUUGGUUCAACUUUAAAAGUA GCGCUGUUUCGGCGCUUUUUU	330
<i>S. pyogenes</i>	GUUGGAACCAUUCAAAACAGCAUAGCAAGUUAAAA UAAGGCUAGUCCGUUACACUUGAAAAAGUGGCA CCGAGUCGGUGCUUUUUU	331
<i>S. mutans</i>	GUUGGAACAUUCGAAACACACAGCAAGUUAAAA UAAGGCAGUGAUUUUUAUCCAGCGUACACAAC UGAAAAAGUGCGCACCGAUUCGGUGCUUUUUU UU	332
<i>S. thermophilus</i>	UUGUGGUUUGAAAACAUUCGAAACACACAGCGAG UUAAAUAAGGCUUAGUCCGUACUACUUGAAAA GGUGGCACCGAUUCGGUGUUUUUUU	333
<i>N. meningitidis</i>	ACAUAUUGUCGCACUGCGAAAUGAGAACCGUUGCU ACAAUAAGGCCUGUAAAAGAUGUGCCGAAACGC UCUGCCCCUAAAAGCUUCUGCUUUUAGGGGCA	334
<i>P. multocida</i>	GCAUAUUGGUUGCACUGCGAAAUGAGAGACGUUGCU ACAAUAAGGCCUUCUGAAAAGAUGACCGUACGC CUGCCCCUUGUAAAUGCAAGGGGCAUCG UUUU	335

The guide sequence of the gRNA comprises a sequence that is complementary to the target sequence. The guide sequence is typically about 20 nucleotides long. For example, the guide sequence may be 15-25 nucleotides long. In some embodiments, the guide sequence is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides long. In some embodiments, the guide sequence is more than 25 nucleotides long. Such suitable guide RNA sequences typically comprise guide sequences that are complementary to a nucleic sequence within 50 nucleotides upstream or downstream of the target nucleotide to be edited.

In some embodiments, the guide RNA is about 15-100 nucleotides long and comprises a sequence of at least 10 contiguous nucleotides that is complementary to a target sequence. In some embodiments, the guide RNA is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides long. In some embodiments, the guide RNA comprises a sequence of 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 contiguous nucleotides that is complementary to a target sequence.

Compositions

Some aspects of the present disclosure relate to compositions that may be used for generating cancer vaccines in vivo or ex vivo. In some embodiments, the composition comprises: (i) a nucleobase editor or a nucleic acid molecule

40 encoding the nucleobase editor described herein; and (ii) a guide nucleotide sequence targeting the nucleobase editor to a tumor specific antigen-encoding polynucleotide. The guide nucleotide sequence that may be used to generate hetero-clitic or cryptic epitopes may be selected from SEQ ID NOS: X-X. Guide nucleotide sequences for generating specific heteroclitic or cryptic epitopes may be found in Tables 5 and 6.

45 In some embodiments, the composition described herein further comprises a pharmaceutically acceptable carrier.

As used here, the term "pharmaceutically-acceptable carrier" means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, manufacturing aid (e.g., lubricant, talc magnesium, calcium or zinc stearate, or steric acid), or solvent encapsulating material, involved in carrying or transporting the compound from one site (e.g., the delivery site) of the body, to another site (e.g., organ, tissue or portion of the body). A pharmaceutically acceptable carrier is "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the tissue of the subject (e.g., physiologically compatible, sterile, physiologic pH, etc.). Some examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, meth-

135

ylcellulose, ethyl cellulose, microcrystalline cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium stearate, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol (PEG); (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates and/or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids (23) serum component, such as serum albumin, HDL and LDL; (22) C2-C12 alcohols, such as ethanol; and (23) other non-toxic compatible substances employed in pharmaceutical formulations. Wetting agents, coloring agents, release agents, coating agents, sweetening agents, flavoring agents, perfuming agents, preservative and antioxidants can also be present in the formulation. The terms such as "excipient", "carrier", "pharmaceutically acceptable carrier" or the like are used interchangeably herein.

In some embodiments, the nucleobase editors and the guide nucleotides of the present disclosure in a composition is administered by injection, by means of a catheter, by means of a suppository, or by means of an implant, the implant being of a porous, non-porous, or gelatinous material, including a membrane, such as a sialastic membrane, or a fiber. In some embodiments, the injection is directed to the liver.

In other embodiments, the nucleobase editors and the guide nucleotides are delivered in a controlled release system. In one embodiment, a pump may be used (see, e.g., Langer, 1990, *Science* 249:1527-1533; Sefton, 1989, *CRC Crit. Ref Biomed. Eng.* 14:201; Buchwald et al., 1980, *Surgery* 88:507; Saudek et al., 1989, *N. Engl. J. Med.* 321:574). In another embodiment, polymeric materials can be used. (See, e.g., *Medical Applications of Controlled Release* (Langer and Wise eds., CRC Press, Boca Raton, Fla., 1974); *Controlled Drug Bioavailability, Drug Product Design and Performance* (Smolen and Ball eds., Wiley, New York, 1984); Ranger and Peppas, 1983, *Macromol. Sci. Rev. Macromol. Chem.* 23:61. See also Levy et al., 1985, *Science* 228:190; During et al., 1989, *Ann. Neurol.* 25:351; Howard et al., 1989, *J. Neurosurg.* 71:105.) Other controlled release systems are discussed, for example, in Langer, *supra*.

In typical embodiments, the pharmaceutical composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous or subcutaneous administration to a subject, e.g., a human. Typically, compositions for administration by injection are solutions in sterile isotonic aqueous buffer. Where necessary, the pharmaceutical can also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the pharmaceutical is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the pharmaceutical is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients can be mixed prior to administration.

136

A pharmaceutical composition for systemic administration may be a liquid, e.g., sterile saline, lactated Ringer's or Hank's solution. In addition, the pharmaceutical composition can be in solid forms and re-dissolved or suspended immediately prior to use. Lyophilized forms are also contemplated.

The pharmaceutical composition can be contained within a lipid particle or vesicle, such as a liposome or microcrystal, which is also suitable for parenteral administration. The particles can be of any suitable structure, such as unilamellar or plurilamellar, so long as compositions are contained therein. Compounds can be entrapped in 'stabilized plasmid-lipid particles' (SPLP) containing the fusogenic lipid dioleoylphosphatidylethanolamine (DOPE), low levels (5-10 mol %) of cationic lipid, and stabilized by a polyethylene glycol (PEG) coating (Zhang Y. P. et al., *Gene Ther.* 1999, 6:1438-47). Positively charged lipids such as N-[1-(2,3-dioleyloxy)propyl]-N,N-trimethyl-amoniummethylsulfate, or "DOTAP," are particularly preferred for such particles and vesicles. The preparation of such lipid particles is well known. See, e.g., U.S. Pat. Nos. 4,880,635; 4,906,477; 4,911,928; 4,917,951; 4,920,016; and 4,921,757.

The pharmaceutical compositions of this disclosure may be administered or packaged as a unit dose, for example. The term "unit dose" when used in reference to a pharmaceutical composition of the present disclosure refers to physically discrete units suitable as unitary dosage for the subject, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; i.e., carrier, or vehicle.

In some embodiments, the nucleobase editors or the guide nucleotides described herein may be conjugated to a therapeutic moiety, e.g., an anti-inflammatory agent. Techniques for conjugating such therapeutic moieties to polypeptides, including e.g., Fc domains, are well known; see, e.g., Amon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld et al. (eds.), 1985, pp. 243-56, Alan R. Liss, Inc.); Hellstrom et al., "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson et al. (eds.), 1987, pp. 623-53, Marcel Dekker, Inc.); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera et al. (eds.), 1985, pp. 475-506); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin et al. (eds.), 1985, pp. 303-16, Academic Press; and Thorpe et al. (1982) "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates," *Immunol. Rev.*, 62:119-158.

Further, the compositions of the present disclosure may be assembled into kits. In some embodiments, the kit comprises nucleic acid vectors for the expression of the nucleobase editors described herein. In some embodiments, the kit further comprises appropriate guide nucleotide sequences (e.g., gRNAs) or nucleic acid vectors for the expression of such guide nucleotide sequences, to target the nucleobase editors to the desired target sequences.

The kit described herein may include one or more containers housing components for performing the methods described herein and optionally instructions of uses. Any of the kit described herein may further comprise components needed for performing the assay methods. Each component of the kits, where applicable, may be provided in liquid form (e.g., in solution), or in solid form, (e.g., a dry powder). In certain cases, some of the components may be reconstituted

137

able or otherwise processible (e.g., to an active form), for example, by the addition of a suitable solvent or other species (for example, water or certain organic solvents), which may or may not be provided with the kit.

In some embodiments, the kits may optionally include instructions and/or promotion for use of the components provided. As used herein, "instructions" can define a component of instruction and/or promotion, and typically involve written instructions on or associated with packaging of the disclosure. Instructions also can include any oral or electronic instructions provided in any manner such that a user will clearly recognize that the instructions are to be associated with the kit, for example, audiovisual (e.g., videotape, DVD, etc.), Internet, and/or web-based communications, etc. The written instructions may be in a form prescribed by a governmental agency regulating the manufacture, use, or sale of pharmaceuticals or biological products, which can also reflect approval by the agency of manufacture, use or sale for animal administration. As used herein, "promoted" includes all methods of doing business including methods of education, hospital and other clinical instruction, scientific inquiry, drug discovery or development, academic research, pharmaceutical industry activity including pharmaceutical sales, and any advertising or other promotional activity including written, oral and electronic communication of any form, associated with the disclosure. Additionally, the kits may include other components depending on the specific application, as described herein.

The kits may contain any one or more of the components described herein in one or more containers. The components may be prepared steriley, packaged in a syringe and shipped refrigerated. Alternatively it may be housed in a vial or other container for storage. A second container may have other components prepared steriley. Alternatively the kits may include the active agents premixed and shipped in a vial, tube, or other container.

The kits may have a variety of forms, such as a blister pouch, a shrink wrapped pouch, a vacuum sealable pouch, a sealable thermoformed tray, or a similar pouch or tray form, with the accessories loosely packed within the pouch, one or more tubes, containers, a box or a bag. The kits may be sterilized after the accessories are added, thereby allowing the individual accessories in the container to be otherwise unwrapped. The kits can be sterilized using any appropriate sterilization techniques, such as radiation sterilization, heat sterilization, or other sterilization methods known in the art. The kits may also include other components, depending on the specific application, for example, containers, cell media, salts, buffers, reagents, syringes, needles, a fabric, such as gauze, for applying or removing a disinfecting agent, disposable gloves, a support for the agents prior to administration, etc.

Therapeutics

The compositions or cancer vaccines (e.g., a whole-cell vaccine comprising a modified tumor cell) described herein may be administered to a subject in need thereof, in a therapeutically effective amount, to treat cancer. The compositions and cancer vaccines described herein induce tumor-specific adaptive responses. It is known that cancer cells exploit immune checkpoints to evade immune surveillance. Thus, in some embodiments, in addition to the compositions or cancer vaccines described herein, agents that modulate the activities of immune checkpoints are also administered to boost the tumor-specific immune response elicited by the cancer vaccines.

In some embodiments, the agents that modulate the activities of immune checkpoints are immune checkpoint inhibi-

138

tors. "Immune checkpoints" are proteins in the immune system that either enhance an immune response signal (co-stimulatory molecules) or reduce an immune response signal. Many cancers protect themselves from the immune system by exploiting the inhibitory immune checkpoint proteins to inhibit the T cell signal. Such inhibitory checkpoint proteins include, without limitation, Cytotoxic T-Lymphocyte-Associated protein 4 (CTLA-4), Programmed Death 1 receptor (PD-1), T-cell Immunoglobulin domain and Mucin domain 3 (TIM3), Lymphocyte Activation Gene-3 (LAG3), V-set domain-containing T-cell activation inhibitor 1 (VTVN1 or B7-H4), Cluster of Differentiation 276 (CD276 or B7-H3), B and T Lymphocyte Attenuator (BTLA), Galectin-9 (GAL9), Checkpoint kinase 1 (Chk1), Adenosine A2A receptor (A2aR), Indoleamine 2,3-dioxygenase (IDO), Killer-cell Immunoglobulin-like Receptor (KIR), Lymphocyte Activation Gene-3 (LAG3), and V-domain Ig suppressor of T cell activation (VISTA).

Some of these immune checkpoint proteins need their cognate binding partners, or ligands, for their immune inhibitory activity. For example, A2aR is the receptor of adenosine A2A and binding of A2A to A2aR activates a negative immune feedback loop. As another example, PD-1 associates with its two ligands, PD-L1 and PD-L2, to down regulate the immune system by preventing the activation of T-cells. PD-1 promotes the programmed cell death of antigen specific T-cells in lymph nodes and simultaneously reduces programmed cell death of suppressor T cells, thus achieving its immune inhibitory function. As yet another example, CTLA4 is present on the surface of T cells, and when bound to its binding partner CD80 or CD86 on the surface of antigen-present cells (APCs), it transmits an inhibitory signal to T cells, thereby reducing the immune response.

Cancer cells are known to exploit the immune checkpoint proteins to escape being attacked by the immune system. Therefore, the use of immune checkpoint inhibitors to enhance an immune response against cancer, and thus treating cancer, have been described. The immunotherapeutic agents in the compositions of the present disclosure may also be immune checkpoint inhibitors. In some embodiments, the immune checkpoint inhibits any one or more of Cytotoxic T-Lymphocyte-Associated protein 4 (CTLA-4), Programmed Death 1 receptor (PD-1), T-cell Immunoglobulin domain and Mucin domain 3 (TIM3), Lymphocyte Activation Gene-3 (LAG3), V-set domain-containing T-cell activation inhibitor 1 (VTVN1 or B7-H4), Cluster of Differentiation 276 (CD276 or B7-H3), B and T Lymphocyte Attenuator (BTLA), Galectin-9 (GAL9), Checkpoint kinase 1 (Chk1), Adenosine A2A receptor (A2aR), Indoleamine 2,3-dioxygenase (IDO), Killer-cell Immunoglobulin-like Receptor (KIR), Lymphocyte Activation Gene-3 (LAG3) and V-domain Ig suppressor of T cell activation (VISTA).

An "immune checkpoint inhibitor" is a molecule that prevents or weakens the activity of an immune checkpoint inhibitor. For example, an immune checkpoint inhibitor may inhibit the binding of the immune checkpoint protein to its cognate binding partner, e.g., PD-1, CTLA-4, or A2aR. In some embodiments, the immune checkpoint inhibitor is a small molecule. In some embodiments, the immune checkpoint inhibitor is a nucleic acid aptamer (e.g., a siRNA targeting any one of the immune checkpoint proteins). In some embodiments, the immune checkpoint inhibitor is a recombinant protein. In some embodiments, the immune checkpoint inhibitor is an antibody. In some embodiments, the antibody comprises an anti-CTLA-4, anti-PD-1, anti-PD-L1, anti-TIM3, anti-LAG3, anti-B7-H3, anti-B7-H4,

139

anti-BTLA, anti-GAL9, anti-Chk, anti-A2aR, anti-IDO, anti-KIR, anti-LAG3, anti-VISTA antibody, or a combination of any two or more of the foregoing antibodies. In some embodiments, the immune checkpoint inhibitor is a monoclonal antibody. In some embodiments, the immune checkpoint inhibitor comprises anti-PD1, anti-PD-L1, anti-CTLA-4, or a combination of any two or more of the foregoing antibodies. For example, the anti-PD-1 antibody is pembrolizumab (Keytruda®) or nivolumab (Opdivo®) and the anti-CTLA-4 antibody is ipilimumab (Yervoy®). Thus, in some embodiments, the immune checkpoint inhibitor comprises pembrolizumab, nivolumab, ipilimumab, or any combination of two or more of the foregoing antibodies. The examples described herein are not meant to be limiting and that any immune checkpoint inhibitors known in the art and any combinations thereof may be used in accordance with the present disclosure.

In some embodiments, the immune checkpoint may be inhibited by disrupting any one of the immune checkpoint genes (e.g., CTLA-4, PD-1, PD-L1, TIM3, LAG3, B7-H3, B7-H4, BTLA, GAL9, Chk1, or A2aR) using any of the gene editing methods known in the art (e.g., CRISPR/Cas9 mediated cleavage of any of the genes).

In some embodiments, an adjuvant is further administered to the subject. Adjuvants are substances which enhance the immune response when administered together with an immunogen or antigen. Adjuvants are thought to function in several ways, including by increasing the surface area of antigen, prolonging the retention of the antigen in the body thus allowing time for the lymphoid system to have access to the antigen, slowing the release of antigen, targeting antigen to macrophages, activating macrophages, or otherwise eliciting non-specific activation of the cells of the immune system see, e.g., H. S. Warren et al, Annu. Rev. Immunol., 4:369 (1986). Currently, an essential role of adjuvants in vaccines is to direct CD4+ T cell subset differentiation, although how adjuvants perform this function is poorly understood.

The ability of an adjuvant to induce and increase a specific type of immune response and the identification of that ability is thus a key factor in the selection of particular adjuvants for vaccine use against a particular pathogen. Typical adjuvants include water and oil emulsions, e.g., Freund's adjuvant, and chemical compounds such as aluminum hydroxide or alum. At present, alum is the only adjuvant approved in the United States for human vaccines; it has been determined that alum induces the production of TH 2 cells.

Many of the most effective adjuvants include bacteria or their products, e.g., microorganisms such as the attenuated strain of *Mycobacterium bovis*, *bacillus Calmette-Guerin* (BCG); microorganism components, e.g., alum-precipitated diphtheria toxoid, bacterial lipopolysaccharide and endotoxins. However, the role that the bacteria play is ill-defined. Recently, it has been noted that many bacteria or their products, lipopolysaccharide, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, and *C. parvum*, stimulate IL-12 production by macrophages A. D'Andrea et al, J. Exp. Med., 176:1387 (1992).

Cancers or tumors include but are not limited to neoplasms, malignant tumors, metastases, or any disease or disorder characterized by uncontrolled cell growth such that it would be considered cancerous. The cancer may be a primary or metastatic cancer. Cancers include, but are not limited to, biliary tract cancer; bladder cancer; brain cancer including glioblastomas and medulloblastomas; breast cancer; cervical cancer; choriocarcinoma; colon cancer; endometrial cancer; esophageal cancer; gastric cancer; hemato-

140

logical neoplasms including acute lymphocytic and myelogenous leukemia; multiple myeloma; AIDS-associated leukemias and adult T-cell leukemia lymphoma; intraepithelial neoplasms including Bowen's disease and Paget's disease; liver cancer; lung cancer; lymphomas including Hodgkin's disease and lymphocytic lymphomas; neuroblastomas; oral cancer including squamous cell carcinoma; ovarian cancer including those arising from epithelial cells, stromal cells, germ cells and mesenchymal cells; pancreatic cancer; prostate cancer; rectal cancer; sarcomas including leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma, and osteosarcoma; skin cancer including melanoma, Kaposi's sarcoma, basocellular cancer, and squamous cell cancer; testicular cancer including germinal tumors such as seminoma, non-seminoma, teratomas, choriocarcinomas; stromal tumors and germ cell tumors; thyroid cancer including thyroid adenocarcinoma and medullary carcinoma; and renal cancer including adenocarcinoma and Wilms' tumor. Commonly encountered cancers include breast, prostate, lung, ovarian, colorectal, and brain cancer.

"A therapeutically effective amount" as used herein refers to the amount of each base-editing agent of the present disclosure required to confer therapeutic effect on the subject, either alone or in combination with one or more other therapeutic agents. Effective amounts vary, as recognized by those skilled in the art, depending on the particular condition being treated, the severity of the condition, the individual subject parameters including age, physical condition, size, gender and weight, the duration of the treatment, the nature of concurrent therapy (if any), the specific route of administration and like factors within the knowledge and expertise of the health practitioner. These factors are well known to those of ordinary skill in the art and can be addressed with no more than routine experimentation. It is generally preferred that a maximum dose of the individual components or combinations thereof be used, that is, the highest safe dose according to sound medical judgment. It will be understood by those of ordinary skill in the art, however, that a subject may insist upon a lower dose or tolerable dose for medical reasons, psychological reasons or for virtually any other reasons. Empirical considerations, such as the half-life, generally will contribute to the determination of the dosage. For example, therapeutic agents that are compatible with the human immune system, such as polypeptides comprising regions from humanized antibodies or fully human antibodies, may be used to prolong the half-life of the polypeptide.

Frequency of administration may be determined and adjusted over the course of therapy, and is generally, but not necessarily, based on treatment and/or suppression and/or amelioration and/or delay of a disease. Alternatively, sustained continuous release formulations of a polypeptide or a polynucleotide may be appropriate. Various formulations and devices for achieving sustained release are known in the art. In some embodiments, dosage is daily, every other day, every three days, every four days, every five days, or every six days. In some embodiments, dosing frequency is once every week, every 2 weeks, every 4 weeks, every 5 weeks, every 6 weeks, every 7 weeks, every 8 weeks, every 9 weeks, or every 10 weeks; or once every month, every 2 months, or every 3 months, or longer. The progress of this therapy is easily monitored by conventional techniques and assays.

The dosing regimen (including the polypeptide used) can vary over time. In some embodiments, for an adult subject of normal weight, doses ranging from about 0.01 to 1000 mg/kg may be administered. In some embodiments, the dose is between 1 to 200 mg. The particular dosage regimen, i.e.,

141

dose, timing and repetition, will depend on the particular subject and that subject's medical history, as well as the properties of the polypeptide or the polynucleotide (such as the half-life of the polypeptide or the polynucleotide, and other considerations well known in the art).

For the purpose of the present disclosure, the appropriate dosage of a therapeutic agent as described herein will depend on the specific agent (or compositions thereof) employed, the formulation and route of administration, the type and severity of the disease, whether the polypeptide or the polynucleotide is administered for preventive or therapeutic purposes, previous therapy, the subject's clinical history and response to the antagonist, and the discretion of the attending physician. Typically the clinician will administer a polypeptide until a dosage is reached that achieves the desired result.

Administration of one or more agents can be continuous or intermittent, depending, for example, upon the recipient's physiological condition, whether the purpose of the administration is therapeutic or prophylactic, and other factors known to skilled practitioners. The administration of an agent (e.g., cancer vaccine) may be essentially continuous over a preselected period of time or may be in a series of spaced dose, e.g., either before, during, or after developing a disease. "Treat," as used herein, means to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve, or affect the disease, the symptom of the disease, or the predisposition toward the disease. "Treating," as used herein refers to the application or administration of a composition or a cancer vaccine described herein to a subject in need thereof.

"A subject in need thereof", refers to an individual who has a disease, a symptom of the disease, or a predisposition toward the disease. In some embodiments, the subject is a mammal. In some embodiments, the subject is a non-human primate. In some embodiments, the subject is human. Alleviating a disease includes delaying the development or progression of the disease, or reducing disease severity. Alleviating the disease does not necessarily require curative results.

As used therein, "delaying" the development of a disease means to defer, hinder, slow, retard, stabilize, and/or postpone progression of the disease. This delay can be of varying lengths of time, depending on the history of the disease and/or individuals being treated. A method that "delays" or alleviates the development of a disease, or delays the onset of the disease, is a method that reduces probability of developing one or more symptoms of the disease in a given time frame and/or reduces extent of the symptoms in a given time frame, when compared to not using the method. Such comparisons are typically based on clinical studies, using a number of subjects sufficient to give a statistically significant result.

“Development” or “progression” of a disease means initial manifestations and/or ensuing progression of the disease. Development of the disease can be detectable and assessed using standard clinical techniques as well known in the art. However, development also refers to progression that may be undetectable. For purpose of this disclosure, development or progression refers to the biological course of the symptoms. “Development” includes occurrence, recurrence, and onset.

As used herein "onset" or "occurrence" of a disease includes initial onset and/or recurrence. Conventional methods, known to those of ordinary skill in the art of medicine, can be used to administer the isolated polypeptide or pharmaceutical composition to the subject, depending upon the

142

type of disease to be treated or the site of the disease. This composition can also be administered via other conventional routes, e.g., administered orally, parenterally, by inhalation spray, topically, rectally, nasally, buccally, vaginally or via
5 an implanted reservoir.

The term "parenteral" as used herein includes subcutaneous, intracutaneous, intravenous, intramuscular, intraarticular, intraarterial, intrasynovial, intrasternal, intrathecal, intralesional, and intracranial injection or infusion techniques. In addition, it can be administered to the subject via injectable depot routes of administration such as using 1-, 3-, or 6-month depot injectable or biodegradable materials and methods.

Without further elaboration, it is believed that one skilled in the art can, based on the above description, utilize the present disclosure to its fullest extent. The following specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. All publications cited herein are incorporated by reference for the purposes on subject matter referenced herein.

EXAMPLES

25 In order that the disclosure described herein may be more fully understood, the following examples are set forth. The synthetic examples described in this application are offered to illustrate the compounds and methods provided herein and are not to be construed in any way as limiting their scope.

Example 1: Guide Nucleotide Sequence-Programmable DNA-Binding Protein Domains, Deaminases, and Base Editors

Non-limiting examples of suitable guide nucleotide sequence-programmable DNA-binding protein domain s are provided. The disclosure provides Cas9 variants, for example, Cas9 proteins from one or more organisms, which may comprise one or more mutations (e.g., to generate dCas9 or Cas9 nickase). In some embodiments, one or more of the amino acid residues, identified below by an asterisk, of a Cas9 protein may be mutated. In some embodiments, the D10 and/or H840 residues of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 11-260, are mutated. In some embodiments, the D10 residue of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 11-260, is mutated to any amino acid residue, except for D. In some embodiments, the D10 residue of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 11-260, is mutated to an A. In some embodiments, the H840 residue of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding residue in any of the amino acid sequences provided in SEQ ID NOs: 11-260, is an H. In some embodiments, the H840 residue of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 11-260, is mutated to any amino acid residue, except for H. In some embodiments, the H840 residue of the amino acid sequence provided in SEQ ID NO: 1, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 11-260, is mutated to an A. In some embodiments, the D10 residue of the amino acid sequence provided

143

in SEQ ID NO: 1, or a corresponding residue in any of the amino acid sequences provided in SEQ ID NOs: 11-260, is a D.

A number of Cas9 sequences from various species were aligned to determine whether corresponding homologous amino acid residues of D10 and H840 of SEQ ID NO: 1 or SEQ ID NO: 11 can be identified in other Cas9 proteins, allowing the generation of Cas9 variants with corresponding mutations of the homologous amino acid residues. The alignment was carried out using the NCBI Constraint-based Multiple Alignment Tool (COBALT (accessible at st-va.ncbi.nlm.nih.gov/tools/cobalt), with the following parameters. Alignment parameters: Gap penalties -11, -1; End-Gap penalties -5, -1. CDD Parameters: Use RPS BLAST on; Blast E-value 0.003; Find Conserved columns and Recompute on. Query Clustering Parameters: Use query clusters on; Word Size 4; Max cluster distance 0.8; Alphabet Regular.

144

An exemplary alignment of four Cas9 sequences is provided below. The Cas9 sequences in the alignment are: Sequence 1 (S1): SEQ ID NO: 11 WP_0109222511|gi 4992247111 type II CRISPR RNA-guided endonuclease 5 Cas9 [*Streptococcus pyogenes*]; Sequence 2 (S2): SEQ ID NO: 12|WP_039695303|gi 746743737|type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus gallolyticus*]; Sequence 3 (S3): SEQ ID NO: 13|WP_045635197|gi 10 782887988|type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mitis*]; Sequence 4 (S4): SEQ ID NO: 14|5AXW_Algi 924443546|*Staphylococcus Aureus* Cas9. The HNH domain (bold and underlined) and the RuvC domain (boxed) are identified for each of the four sequences. 15 Amino acid residues 10 and 840 in S1 and the homologous amino acids in the aligned sequences are identified with an asterisk following the respective amino acid residue.

S1	1	--MDKK-YSIGLD*IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI--GALLFDSG--ETAEATRLKRTARRYT	73
S2	1	--MTKKNYSIGLD*IGTNSVGWAVITDDYKVPAKKMKVLGNTDKYIKKNLL--GALLFDSG--ETAEATRLKRTARRYT	74
S3	1	--M-KKGYSIGLD*IGTNSVGFAVITDDYKVPSKKMKVLGNTDKRFIKKNLI--GALLFDEG--TTAEARRLKRTARRYT	73
S4	1	GSHMKRNYILGLD*IGITSVGYGII--DYET-----RDVIDAGVRLKEANVENNEGRRSKRGARRLKR	61
S1	74	RRKNRICYLQEIFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL	153
S2	75	RRKNRRLRYLQEIFFANEIAKVDFFQRLDESFLTDKTFDSHPIFGNKAEEDAYHQKFPTIYHLRKHLADSSEKADLRL	154
S3	74	RRKNRRLRYLQEIFSEEMSKVDSFFFHRLDDSLIPEKDRESKYPPIFATLTEEKEYHKQFPTIYHLRKQLADSKEKDLRL	153
S4	62	RRRHRIQRVKLL-----FDYNLLTD-----HSELSGINPYEARVKGLSQKLSEEE	107
S1	154	IYLALAHMIKFRGHFLIEGDLNPNDNSVDKLFIQLVQTYNOLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEK	233
S2	155	VYLALAHMIKFRGHFLIEGELNAENTDVQKIFADFVGVYNRTFDDSHLSEITVDVASILTEKISKSRRLLENLIKYYPTEK	234
S3	154	IYLALAHMIKYRGHFLYEEAFDIKNNDIQKIFNEFISIYDNTFEGSSLGQNAQVEAIFTDKISKSAKRERVLKLPDEK	233
S4	108	FSAALLHLAKRGG-----VHNVNEVEEDT-----	131
S1	234	KNGLPGNLIALSLGLTPNFKNFDSLQDAEAKLQLSKDTYDDLDNLQAQIGDQYADLPLAAKNLSDAILLSDILRVNTEIT	313
S2	235	KNTLPGNLIALALGLQPNFKTNFKLSEDALKQFSKDTYEEDEELLGKIGDDYADLFTSAKNLYDAILLSGILTVDNST	314
S3	234	STGLFSEFLKLIVGNQADFKKHFDLEDKAPLQFSKDTYDEDLENLLGQIGDDFTDLFVSAKKLYDAILLSGILTVDNST	313
S4	132	----GNELS-----TKEQISR-----	144
S1	314	KAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSNGYAGYIDGGASQEEFYKFKPILEKM--DGTEELLV	391
S2	315	KAPLSASMIKRYVEHHEDLEKLKEFIKANKSELYHDIFPKDKNKNGYAGYIENGVKQDEFYKYLKNILSKIKIDGSDFLD	394
S3	314	KAPLSASMIERYENHQNDLAALKQFIKNNLPEKYDEVFSQSKDGYAGYIDGKTTQETFYKYIKNLLSKF--EGTDYFLD	391
S4	145	----SKALEEKYVAELQ-----LERLKKDG-----	165
S1	392	KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRROEDFYPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAMTRKSEE	471
S2	395	KIEREDFLRKQRTFDNGSIPHQIHLQEMHAIHLRQGDYPFLKEKQDRIEKILTFRIPIYYVGPLVRKDSRFAWEYRSDE	474
S3	392	KIEREDFLRKQRTFDNGSIPHQIHLQEMNAILRROGEYPFLKDNKEKIEKILTFRIPIYYVGPLARGNRDFAWLTRNSDE	471
S4	166	--EVGRSINRFKTS-----YVKEAKQLLKVKQAKYHQLDQSFDITYIDLLETTRRTYYEGP--GEGSPFGW-----K	227
S1	472	TITPWNFEEVDKGASAQSFIGERMTNFDFKNLPNEKVLPKHSLLYEYFTVYNELTKVKVYVTEGMRKPAFLSGEQKKAIVDL	551
S2	475	KITPWNFDKVIDKEKSAEKFITRMTLNDLYLPEEKVLPKHSVYETYAVYNELTKIKYVNEQGKE-SFFDSNMKQEIFDH	553
S3	472	AIRPWNFEEIVDKASSAEDFINKMTNYDLYLPEEKVLPKHSLLYEYFTVYNELTKVKFIAEGLRDYQFLDSGQKKQIVNQ	551
S4	228	DIKEW-----YEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEK---LEYYEKFQIIEN	289
S1	552	LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR---FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLFED	628
S2	554	VFKENRKVTKEKLLNYLNKEFPEYRIKDLIGLDKENKSFNASLGTYHDLKKIL-DKAFLDDKVNEEVIEDIIKTLTLFED	632

-continued

S3	552	LPKENRKVTEKDIHYLHN-VDGYDIELKGIEKQ--FNASLSTYHDLKIIKDKEFMDDAKNEAILENIVHTLTIFED	627
S4	290	VFKQKKPCTLQIAKEILVNEEDIKYRVTSTGKPEF---TNLKVVHDIKDITARKEI---ENAEELDQIAKILTIYQS	363
S1	629	REMIEERLKTYAHLFDDKVMKQLKR-RRTYGWGRLSRKLINGIRDQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKED	707
S2	633	KDMIHERLQKYSIDFTANQLKKLER-RHYTGWGRLSYKLINGIRNKENNKTILDYLIDDGSANRNFMQLINDDTLPFKQI	711
S3	628	REMIKQRLAQYDSLFDKVKALTR-RHYTGWGLSAKLINGICDKQTGNTILDYLIDDGKINRNFMQLINDDGLSFKEI	706
S4	364	SEDIQEELTNLSELTQEEIEQISNLKGYGTGTHNLSKAINLILDE-----LWHTNDNQIAIFNRLKLVP-----	428
S1	708	IQKAQVSGQC <u>D</u> SLHEHIANLAGSPAIIKGILQTVKVVDELVKVMGRHKPENIVIEMA <u>R</u> ENQTT-----Q <u>G</u> <u>K</u> <u>N</u> <u>S</u> <u>R</u> <u>E</u> <u>M</u>	781
S2	712	IQKSQVVGDU <u>D</u> IEAVVHDLPGSPAIIKGILQSVKIVDELVKVMG-GNPDNIVIEMA <u>R</u> ENQTT-----NR <u>G</u> <u>R</u> <u>S</u> <u>Q</u> <u>O</u> <u>R</u> <u>L</u>	784
S3	707	IQKAQVIGKT <u>D</u> DKVQVQELSGSPAIIKGILQSVKIVDELVKVMG-HAPESIVIEMA <u>R</u> ENQTT-----AR <u>G</u> <u>K</u> <u>N</u> <u>S</u> <u>Q</u> <u>O</u> <u>R</u> <u>Y</u>	779
S4	429	-KKVDSLQQK <u>E</u> PTTLVDDFLSPVVKRSFTQSIKVINAIIKKYQ--LPNDIIIELA <u>R</u> EKNSKDAQKMINEM <u>Q</u> <u>K</u> <u>R</u> <u>N</u> <u>R</u> <u>Q</u> <u>T</u> <u>N</u>	505
S1	782	<u>K</u> RIEEGIKELGSQIL-----KEHPVENTQLQNEKLYLYLQNQGRDMYV <u>D</u> QELDINRLSD----YD <u>V</u> DH*IPQASF <u>L</u> KDD	850
S2	785	<u>K</u> KLQNSLKELGSNILNEEKPSYIEDKVENSHLQNDQLFLYYIQNGKDMYT <u>G</u> DELDIDHLSD----YDIDH*II <u>P</u> QAFIKDD	860
S3	780	<u>K</u> RIEDSILKILASGL---DSNILKENPTDNNQLQNDRLFLYYLQNQGRDMYT <u>G</u> EALDINGLSS---YDIDH*II <u>P</u> QAFIKDD	852
S4	506	<u>E</u> RIEEIIRTTGK-----ENAKYLIKEKIKLHDMQEGKCLYSLEAPILEDLLNNPFPNYEVDH*II <u>P</u> RSVSFDN	570
S1	851	<u>S</u> IDNKVLT <u>R</u> SDKNRGKSDNVPS <u>E</u> EVVKMKNYWRQLLNAKLIT <u>R</u> QKF <u>D</u> N-LTKAER <u>G</u> L-SELD-----KAGFIKRQLV	922
S2	861	<u>S</u> IDNRVLTSSAKNRGKSDDVPSLDIVRARKAEWVRLYKSG <u>L</u> ISKRF <u>D</u> N-LTKAER <u>G</u> L-TEAD-----KAGFIKRQLV	932
S3	853	<u>S</u> LDNRVLTSSKDNRGKSDNVPSIEVVQKR <u>A</u> F <u>W</u> QQLLDSKLISERK <u>E</u> NN-LTKAER <u>G</u> L-DERD-----KVGFIKRQLV	924
S4	571	<u>S</u> FNNKVVLVKQEEASKGNRTP <u>F</u> QYLSSSDSKISYET <u>F</u> KKHILNLAKGKGRISKTK <u>E</u> YLLERDINRFSVQKDFINRNLV	650
S1	923	<u>E</u> TRQITKHVAQILDARFNTEHDENDKLIREVKVITL <u>K</u> SKLVSDFRKDFQFYKVREINDYHHAHDAYLNAVVG <u>T</u> ALIKKYP	1002
S2	933	<u>E</u> TRQITKHVAQILDARFNTEHDENDKLIREVKVITL <u>K</u> SKLVSDFRKDFQFYKVREINDYHHAHDAYLNAVVG <u>T</u> ALLKKYP	1012
S3	925	<u>E</u> TRQITKHVAQILDARYNTEVNE <u>K</u> DKKNRTVKIITL <u>K</u> SKLVSFRKEFRLYKVREINDYHHAHDAYLNAVVA <u>K</u> AILKKYP	1004
S4	651	<u>D</u> TRYATRGLMNLLRSYFRVN-----NLDVKVKSINGGFTSFLRRWKFKERNKGKHYKHAEDALIA-----	712
S1	1003	<u>K</u> LESEFVYGDYKVDVRKMIAKSEQ--EIGKATAKYFFYSNIMNFFKTEITLANGEI <u>R</u> KPLIETNGETGEI <u>V</u> WDKG---	1077
S2	1013	<u>K</u> LADEFVYGEYKKYDIRKFITNSSD----KATAKYFFYSNLMNFFKTKV <u>Y</u> ADGTU <u>Y</u> FERPIIETNAD-GEIAWNQ---	1083
S3	1005	<u>K</u> LEPEFVYGEYQKYDLKRYISRSKDPKEVEKATEKYFFYSNLLNFFKEEVHYADGTIVKRENIEYSKDTGEIAWNKE---	1081
S4	713	--NADFI <u>F</u> KEWK <u>K</u> L <u>D</u> KAKKVMENQM-----FEEKQAESMPEIETEQEYKEIFITPHQIK	764
S1	1078	-----RDFATV <u>R</u> KVLSMPQVNIVKKTEVQT <u>G</u> GFSKESILPKRNSDKLIA <u>R</u> KKD--WDPKKYGGFDSP <u>T</u> VAYSVLVVAKV	1149
S2	1084	-----IDFEKVRKVLSPQVNIVKKVETQT <u>G</u> GFSKESILPKGDS <u>D</u> KLIPR <u>K</u> TKV <u>Y</u> WDTKKYGGFDSP <u>T</u> VAYSVFVVADV	1158
S3	1082	-----KDFATI <u>K</u> KKVLSLPQVNIVKKREVQT <u>G</u> GFSKESILPKGN <u>S</u> DKLIPR <u>K</u> TKD <u>I</u> LLDTTKYGGFDSP <u>V</u> IAYSILLIADI	1156
S4	765	<u>H</u> IKDFKD <u>K</u> YSHRV <u>D</u> DKPNRELINDTLYST <u>R</u> KDDKGNTLIVNNLNGLYDKDND <u>K</u> L---KKL <u>I</u> N-KSP---EKLLMYHH	835
S1	1078	EKGKSKKLKSVKELLGITMERSSFEKNI-DFLEAKG----YKEVK <u>K</u> DLI <u>I</u> KL <u>P</u> K <u>Y</u> SLFELENGRKRMLASAGELQKG	1223
S2	1084	EKGKAKKLKTVKELVGISIMERSFFEE <u>N</u> P-EFL <u>E</u> NGK----YHN <u>I</u> RED <u>K</u> L <u>I</u> KL <u>P</u> K <u>Y</u> SLFE <u>E</u> FE <u>G</u> GR <u>R</u> LLASASELQKG	1232
S3	1157	EKGKAKKLKTVKTLVGITIMEKA <u>E</u> ENPI-TFL <u>E</u> NGK----YHN <u>V</u> R <u>K</u> EN <u>I</u> L <u>C</u> L <u>P</u> K <u>Y</u> SLFELENGR <u>R</u> LLASAKELQKG	1230
S4	836	DPQTYQKLK-----LIMEQYGDENP <u>L</u> YKYYEETGN <u>L</u> TKYSSKKDNGPV <u>I</u> KKIKYYGN <u>K</u> LN <u>A</u> HL <u>D</u> ITDDYP <u>N</u> SRNKV	907
S1	1224	NELALPSKYVNFLYLA <u>S</u> HYEKL <u>K</u> GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSK <u>R</u> V <u>I</u> LD <u>A</u> NLDK <u>V</u> LSAYNKH-----	1297
S2	1233	NEMVLP <u>G</u> YLVELLYHAHRADNF----NSTEYLNYVSEHK <u>K</u> FE <u>K</u> V <u>L</u> SCVEDFANLYV <u>D</u> V <u>E</u> KN <u>L</u> S <u>K</u> RAVADSM-----	1301
S3	1231	NEIVLPVYLTLLYHSKNVH <u>K</u> L----DEPGHLEYIQKHRNEFK <u>D</u> LLNLVSEFSQKYV <u>L</u> ADANLEKIK <u>K</u> SLYADN-----	1299
S4	908	VKLSLKPYRF <u>D</u> -VYLDNGVYKFV----TVKNU <u>D</u> VIK--KENYYEVNSKAYEEAKLKKISNQA <u>E</u> FIASFYNN <u>D</u> LIKING	979

-continued

S1	1298	RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLHQHQSIT--GLYETRI ---- DLSQL	1365
S2	1302	DNFSIEEIISNSFINLLTALGAPADFNFLGEKIPRKRYTSTKECLNATLHQHQSIT-GLYETRIDLSQL	1369
S3	1300	EQADIEILANSFINLLTALGAPAAFKFFGKDIDRKRYTTVSEILNATLHQHQSIT-GLYETWII	1369
S4	980	ELYRVIDGVNNNDLLNRIEVNMDITYR-EYLENMNDKRPPRIIKTIASKT --- QSIKKYSTDILGNLYEVK-SKKHPQIJK	1055
S1	1366	GGD	1368
S2	1370	GEE	1372
S3	1368	GED	1370
S4	4056	G--	1056

The alignment demonstrates that amino acid sequences and amino acid residues that are homologous to a reference Cas9 amino acid sequence or amino acid residue can be identified across Cas9 sequence variants, including, but not limited to Cas9 sequences from different species, by identifying the amino acid sequence or residue that aligns with the reference sequence or the reference residue using alignment programs and algorithms known in the art. This disclosure provides Cas9 variants in which one or more of the amino acid residues identified by an asterisk in SEQ ID NOs: 11-14 (e.g., S1, S2, S3, and S4, respectively) are mutated as described herein. The residues D10 and H840 in Cas9 of SEQ ID NO: 1 that correspond to the residues identified in SEQ ID NOs: 11-14 by an asterisk are referred to herein as "homologous" or "corresponding" residues. Such homologous residues can be identified by sequence alignment, e.g., as described above, and by identifying the

sequence or residue that aligns with the reference sequence or residue. Similarly, mutations in Cas9 sequences that correspond to mutations identified in SEQ ID NO: 1 herein, e.g., mutations of residues 10, and 840 in SEQ ID NO: 1, are referred to herein as "homologous" or "corresponding" mutations. For example, the mutations corresponding to the D10A mutation in SEQ ID NO: 1 or S1 (SEQ ID NO: 11) for the four aligned sequences above are D11A for S2, D10A for S3, and D13A for S4; the corresponding mutations for H840A in SEQ ID NO: 1 or S1 (SEQ ID NO: 11) are H850A for S2, H842A for S3, and H560A for S4.

A total of 250 Cas9 sequences (SEQ ID NOs: 11-260) from different species are provided. Amino acid residues homologous to residues 10, and 840 of SEQ ID NO: 1 may be identified in the same manner as outlined above. All of these Cas9 sequences may be used in accordance with the present disclosure.

```

WP_039695303.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus gallolyticus]
SEQ ID NO: 12

WP_045635197.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus mitis]
SEQ ID NO: 13

5AXW A Cas9, Chain A, Crystal Structure [Staphylococcus Aureus] SEQ ID NO: 14

WP_009880683.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 15

WP_010922251.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 16

WP_011054416.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 17

WP_011284745.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 18

WP_011285506.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 19

WP_011527619.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 20

WP_012560673.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 21

WP_014407541.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 22

WP_020905136.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 23

WP_023080005.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]
SEQ ID NO: 24

```

149**150**

-Continued

WP_023610282.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 25

WP_030125963.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 26

WP_030126706.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 27

WP_031488318.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 28

WP_032460140.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 29

WP_032461047.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 30

WP_032462016.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 31

WP_032462936.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 32

WP_032464890.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 33

WP_033888930.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 34

WP_038431314.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 35

WP_038432938.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 36

WP_038434062.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pyogenes*]
SEQ ID NO: 37

BAQ51233.1 CRISPR-associated protein, Csnl family [*Streptococcus pyogenes*]
SEQ ID NO: 38

KGE60162.1 hypothetical protein MGAS2111_0903 [*Streptococcus pyogenes* MGAS2111]
SEQ ID NO: 39

KGE60856.1 CRISPR-associated endonuclease protein [*Streptococcus pyogenes* SS1447]
SEQ ID NO: 40

WP_002989955.1 MULTISPECIES: type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus*]
SEQ ID NO: 41

WP_003030002.1 MULTISPECIES: type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus*]
SEQ ID NO: 42

WP_003065552.1 MULTISPECIES: type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus*]
SEQ ID NO: 43

WP_001040076.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 44

WP_001040078.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 45

WP_001040080.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 46

WP_001040081.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 47

WP_001040083.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 48

WP_001040085.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 49

WP_001040087.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 50

WP_001040088.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*]
SEQ ID NO: 51

151**152**

-Continued

WP_001040089.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 52

WP_001040090.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 53

WP_001040091.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 54

WP_001040092.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 55

WP_001040094.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 56

WP_001040095.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 57

WP_001040096.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 58

WP_001040097.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 59

WP_001040098.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 60

WP_001040099.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 61

WP_001040100.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 62

WP_001040104.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 63

WP_001040105.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 64

WP_001040106.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 65

WP_001040107.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 66

WP_001040108.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 67

WP_001040109.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 68

WP_001040110.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 69

WP_015058523.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 70

WP_017643650.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 71

WP_017647151.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 72

WP_017648376.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 73

WP_017649527.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 74

WP_017771611.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 75

WP_017771984.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 76

CFQ25032.1 CRISPR-associated protein [*Streptococcus agalactiae*] SEQ ID NO: 77

CFV16040.1 CRISPR-associated protein [*Streptococcus agalactiae*] SEQ ID NO: 78

KLJ37842.1 CRISPR-associated protein Csn1 [*Streptococcus agalactiae*] SEQ ID NO: 79

-Continued

- KLJ72361.1 CRISPR-associated protein Csn1 [*Streptococcus agalactiae*] SEQ ID NO: 80
- KLL207071.1 CRISPR-associated protein Csn1 [*Streptococcus agalactiae*] SEQ ID NO: 81
- KLL42645.1 CRISPR-associated protein Csn1 [*Streptococcus agalactiae*] SEQ ID NO: 82
- WP_047207273.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 83
- WP_047209694.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 84
- WP_050198062.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 85
- WP_050201642.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 86
- WP_050204027.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 87
- WP_050881965.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 88
- WP_050886065.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus agalactiae*] SEQ ID NO: 89
- AHN30376.1 CRISPR-associated protein Csn1 [*Streptococcus agalactiae* 138P] SEQ ID NO: 90
- EAO78426.1 reticulocyte binding protein [*Streptococcus agalactiae* H36B] SEQ ID NO: 91
- CCW42055.1 CRISPR-associated protein, SAG0894 family [*Streptococcus agalactiae* ILRI112] SEQ ID NO: 92
- WP_003041502.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus anginosus*] SEQ ID NO: 93
- WP_037593752.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus anginosus*] SEQ ID NO: 94
- WP_049516684.1 CRISPR-associated protein Csn1 [*Streptococcus anginosus*] SEQ ID NO: 95
- GAD46167.1 hypothetical protein ANG6_0662 [*Streptococcus anginosus* T5] SEQ ID NO: 96
- WP_018363470.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus caballi*] SEQ ID NO: 97
- WP_003043819.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus canis*] SEQ ID NO: 98
- WP_006269658.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus constellatus*] SEQ ID NO: 99
- WP_048800889.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus constellatus*] SEQ ID NO: 100
- WP_012767106.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus dysgalactiae*] SEQ ID NO: 101
- WP_014612333.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus dysgalactiae*] SEQ ID NO: 102
- WP_015017095.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus dysgalactiae*] SEQ ID NO: 103
- WP_015057649.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus dysgalactiae*] SEQ ID NO: 104
- WP_048327215.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus dysgalactiae*] SEQ ID NO: 105
- WP_049519324.1 CRISPR-associated protein Csn1 [*Streptococcus dysgalactiae*] SEQ ID NO: 106
- WP_012515931.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus equi*] SEQ ID NO: 107

155**156**

-Continued

WP_021320964.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus equi*]
SEQ ID NO: 108

WP_037581760.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus equi*]
SEQ ID NO: 109

WP_004232481.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus equinus*]
SEQ ID NO: 110

WP_009854540.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus gallolyticus*]
SEQ ID NO: 111

WP_012962174.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus gallolyticus*]
SEQ ID NO: 112

WP_039695303.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus gallolyticus*]
SEQ ID NO: 113

WP_014334983.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus infantarius*]
SEQ ID NO: 114

WP_003099269.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus iniae*]
SEQ ID NO: 115

AHY15608.1 CRISPR-associated protein Csn1 [*Streptococcus iniae*]
SEQ ID NO: 116

AHY17476.1 CRISPR-associated protein Csn1 [*Streptococcus iniae*]
SEQ ID NO: 117

ESR09100.1 hypothetical protein IUSA1_08595 [*Streptococcus iniae* IUSA1] SEQ ID NO: 118

AGM98575.1 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [*Streptococcus iniae* SF1]
SEQ ID NO: 119

ALF27331.1 CRISPR-associated protein Csn1 [*Streptococcus intermedius*] SEQ ID NO: 120

WP_018372492.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus massiliensis*]
SEQ ID NO: 121

WP_045618028.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mitis*]
SEQ ID NO: 122

WP_045635197.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mitis*]
SEQ ID NO: 123

WP_002263549.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 124

WP_002263887.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 125

WP_002264920.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 126

WP_002269043.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 127

WP_002269448.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 128

WP_002271977.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 129

WP_002272766.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 130

WP_002273241.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 131

WP_002275430.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 132

WP_002276448.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 133

157**158**

-Continued

WP_002277050.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 134

WP_002277364.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 135

WP_002279025.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 136

WP_002279859.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 137

WP_002280230.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 138

WP_002281696.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 139

WP_002282247.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 140

WP_002282906.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 141

WP_002283846.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 142

WP_002287255.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 143

WP_002288990.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 144

WP_002289641.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 145

WP_002290427.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 146

WP_002295753.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 147

WP_002296423.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 148

WP_002304487.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 149

WP_002305844.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 150

WP_002307203.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 151

WP_002310390.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 152

WP_002352408.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 153

WP_012997688.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 154

WP_014677909.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 155

WP_019312892.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 156

WP_019313659.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 157

WP_019314093.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 158

WP_019315370.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 159

WP_019803776.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 160

159**160**

-Continued

WP_019805234.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 161

WP_024783594.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 162

WP_024784288.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 163

WP_024784666.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 164

WP_024784894.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 165

WP_024786433.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus mutans*]
SEQ ID NO: 166

WP_049473442.1 CRISPR-associated protein Csn1 [*Streptococcus mutans*] SEQ ID NO: 167

WP_049474547.1 CRISPR-associated protein Csn1 [*Streptococcus mutans*] SEQ ID NO: 168

EMC03581.1 hypothetical protein SMU69_09359 [*Streptococcus mutans* NLML4] SEQ ID NO: 169

WP_000428612.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus oralis*]
SEQ ID NO: 170

WP_000428613.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus oralis*]
SEQ ID NO: 171

WP_049523028.1 CRISPR-associated protein Csn1 [*Streptococcus parasanguinis*] SEQ ID NO: 172

WP_003107102.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus parauberis*]
SEQ ID NO: 173

WP_054279288.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus phocae*]
SEQ ID NO: 174

WP_049531101.1 CRISPR-associated protein Csn1 [*Streptococcus pseudopneumoniae*]
SEQ ID NO: 175

WP_049538452.1 CRISPR-associated protein Csn1 [*Streptococcus pseudopneumoniae*]
SEQ ID NO: 176

WP_049549711.1 CRISPR-associated protein Csn1 [*Streptococcus pseudopneumoniae*]
SEQ ID NO: 177

WP_007896501.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus pseudoporcinus*]
SEQ ID NO: 178

EFR44625.1 CRISPR-associated protein, Csn1 family [*Streptococcus pseudoporcinus* SPIN 20026]
SEQ ID NO: 179

WP_002897477.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus sanguinis*]
SEQ ID NO: 180

WP_002906454.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus sanguinis*]
SEQ ID NO: 181

WP_009729476.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus* sp. F0441]
SEQ ID NO: 182

CQR24647.1 CRISPR-associated protein [*Streptococcus* sp. FF10] SEQ ID NO: 183

WP_000066813.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus* sp. M334]
SEQ ID NO: 184

WP_009754323.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus* sp. taxon 056]
SEQ ID NO: 185

WP_044674937.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus suis*]
SEQ ID NO: 186

WP_044676715.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus suis*]
SEQ ID NO: 187

WP_044680361.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus suis*]
SEQ ID NO: 188

US 12,390,514 B2

161

-continued

WP_044681799.1 type II CRISPR RNA-guided endonuclease Cas9 [*Streptococcus suis*]
SEQ ID NO: 189

WP_049533112.1 CRISPR-associated protein Csnl [*Streptococcus suis*] SEQ ID NO: 190

WP_029090905.1 type II CRISPR RNA-guided endonuclease Cas9 [*Brochothrix thermosphacta*]
SEQ ID NO: 191

WP_006506696.1 type II CRISPR RNA-guided endonuclease Cas9 [*Catenibacterium mitsukai*]
SEQ ID NO: 192

AIT42264.1 Cas9hc:NLS:A [Cloning vector pYB196] SEQ ID NO: 193

WP_034440723.1 type II CRISPR endonuclease Cas9 [*Clostridiales bacterium S5-A11*]
SEQ ID NO: 194

AKQ21048.1 Cas9 [CRISPR-mediated gene targeting vector p (bhsp68-Cas9)]
SEQ ID NO: 195

WP_004636532.1 type II CRISPR RNA-guided endonuclease Cas9 [*Dolosigranulum pigrum*]
SEQ ID NO: 196

WP_002364836.1 MULTISPECIES: type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus*]
SEQ ID NO: 197

WP_016631044.1 MULTISPECIES: type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus*]
SEQ ID NO: 198

EMS75795.1 hypothetical protein H318_06676 [*Enterococcus durans* IPLA 655]
SEQ ID NO: 199

WP_002373311.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 200

WP_002378009.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 201

WP_002407324.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 202

WP_002413717.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 203

WP_010775580.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 204

WP_010818269.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 205

WP_010824395.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 206

WP_016622645.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 207

WP_033624816.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 208

WP_033625576.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 209

WP_033789179.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecalis*]
SEQ ID NO: 210

WP_002310644.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 211

WP_002312694.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 212

WP_002314015.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 213

WP_002320716.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 214

162

-Continued

WP_002330729.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 215

WP_002335161.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 216

WP_002345439.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 217

WP_034867970.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 218

WP_047937432.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus faecium*]
SEQ ID NO: 219

WP_010720994.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus hirae*]
SEQ ID NO: 220

WP_010737004.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus hirae*]
SEQ ID NO: 221

WP_034700478.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus hirae*]
SEQ ID NO: 222

WP_007209003.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus italicus*]
SEQ ID NO: 223

WP_023519017.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus mundtii*]
SEQ ID NO: 224

WP_010770040.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus phoeniculicola*]
SEQ ID NO: 225

WP_048604708.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus sp. AM1*]
SEQ ID NO: 226

WP_010750235.1 type II CRISPR RNA-guided endonuclease Cas9 [*Enterococcus villorum*]
SEQ ID NO: 227

AII16583.1 Cas9 endonuclease [Expression vector pCas9] SEQ ID NO: 228

WP_029073316.1 type II CRISPR RNA-guided endonuclease Cas9 [*Kandleria vitulina*]
SEQ ID NO: 229

WP_031589969.1 type II CRISPR RNA-guided endonuclease Cas9 [*Kandleria vitulina*]
SEQ ID NO: 230

KDA45870.1 CRISPR-associated protein Cas9/Csn1, subtype II/NMEM1 [*Lactobacillus animalis*]
SEQ ID NO: 231

WP_039099354.1 type II CRISPR RNA-guided endonuclease Cas9 [*Lactobacillus curvatus*]
SEQ ID NO: 232

AKP02966.1 hypothetical protein ABB45_04605 [*Lactobacillus farciminis*] SEQ ID NO: 233

WP_010991369.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria innocua*]
SEQ ID NO: 234

WP_033838504.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria innocua*]
SEQ ID NO: 235

EHN60060.1 CRISPR-associated protein, Csn1 family [*Listeria innocua* ATCC 33091]
SEQ ID NO: 236

EFR89594.1 crispr-associated protein, Csn1 family [*Listeria innocua* FSL S4-378]
SEQ ID NO: 237

WP_038409211.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria ivanovii*]
SEQ ID NO: 238

EFR895520.1 crispr-associated protein Csn1 [*Listeria ivanovii* FSL F6-596]
SEQ ID NO: 239

WP_003723650.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 240

WP_003727705.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 241

WP_003730785.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 242

-continued

WP_003733029.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 243

WP_003739838.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 244

WP_014601172.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 245

WP_023548323.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 246

WP_031665337.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 247

WP_031669209.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 248

WP_033920898.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria monocytogenes*]
SEQ ID NO: 249

AKI42028.1 CRISPR-associated protein [*Listeria monocytogenes*] SEQ ID NO: 250

AKI50529.1 CRISPR-associated protein [*Listeria monocytogenes*] SEQ ID NO: 251

EFR83390.1 crispr-associated protein Csn1 [*Listeria monocytogenes* FSL F2-208]
SEQ ID NO: 252

WP_046323366.1 type II CRISPR RNA-guided endonuclease Cas9 [*Listeria seeligeri*]
SEQ ID NO: 253

AKE81011.1 Cas9 [Plant multiplex genome editing vector pYLCRISPR/Cas9Pubi-H]
SEQ ID NO: 254

CUO82355.1 Uncharacterized protein conserved in bacteria [*Roseburia hominis*]
SEQ ID NO: 255

WP_033162887.1 type II CRISPR RNA-guided endonuclease Cas9 [*Sharpea azabuensis*]
SEQ ID NO: 256

AGZ01981.1 Cas9 endonuclease [synthetic construct] SEQ ID NO: 257

AKA60242.1 nuclease deficient Cas9 [synthetic construct] SEQ ID NO: 258

AKS40380.1 Cas9 [Synthetic plasmid pFC330] SEQ ID NO: 259

4UN5_B Cas9, Chain B, Crystal Structure SEQ ID NO: 260

Non-limiting examples of suitable deaminase domains are provided.

Human AID

(SEQ ID NO: 303)

MDSLLMNRRKFLYQFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGYLRLRNKGCHVELLFLRYISDWD
LDPGRCYRVTWPTSWSPCYDCARHVADFLRGPNLSSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIAIMT
FKDYFYCWNTFVENHERTFKAWEGLHENSVRLSRQLRILLPLPLYEVDDLRDAFRTLGL
(underline: nuclear localization signal; double underline: nuclear export signal)

Mouse AID

(SEQ ID NO: 271)

MDSLLMKQKKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSCSLDFGHLRNKGCHVELLFLRYISDWD
LDPGRCYRVTWPTSWSPCYDCARHVADFLRGPNLSSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIGIMT
FKDYFYCWNTFVENRERTFKAWEGLHENSVRTRQLRILLPLPLYEVDDLRDAFRMLGF
(underline: nuclear localization signal; double underline: nuclear export signal)

Dog AID

(SEQ ID NO: 272)

MDSLLMKQRKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGHLRNKGCHVELLFLRYISDWD
LDPGRCYRVTWPTSWSPCYDCARHVADFLRGPNLSSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIAIMT
FKDYFYCWNTFVENREKTFKAWEGLHENSVRLSRQLRILLPLPLYEVDDLRDAFRTLGL
(underline: nuclear localization signal; double underline: nuclear export signal)

US 12,390,514 B2

167

-continued

Bovine AID

(SEQ ID NO: 273)

MDSLLKKQRFQFLYQFKNVRWAKGRHETYLCYVVKRRDSPTSFSLDFGHLRNKAGCHVELLFLRYISDWD
LDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLRLRIFTARLYFCDKERKAEPEGLRRLHRAGVQIAIM
TFKDYFYCWNTFVENHERTFKAWEGLHENSVRSLRQLRRILLPLYEVDDLRDAFRTLGL
 (underline: nuclear localization signal; double underline: nuclear export signal)

Rat AID

(SEQ ID NO: 1072)

MAVGSKPKAALVGPHWERERIWCFLCSTGLTQQTGQTSRWRPAATQDPVSPPRSLLMQRFKLYHFK
 NVRWAKGRHETYLCYVVKRRDSATSFSLSDFGYLRNKGCHVELLFLRYISDWLDLPGRCYRVTWFTSWS
 PCYDCARHVADFLRGNPNLSLRIFTARLTGWLGAAPGLMSPARPSDYFYCWNTFVENHERTFKAWEGLHE
 NSVRLSRRRLRILLPLYEVDDLRLDAFRTLGL

Mouse APOBEC-3

(SEQ ID NO: 274)

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKCDSPVSLHHGVFKNKDNIA
EETCFLYWFDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNIRDPEPQONLCRL
LVQEQAQVAAMDLYEFKKCWKKFVDNGGRRFRPKRLLTNFRYQDSKLQEI LRPCYIPVPSSSSTLSNIC
LTGKLPETRVCVERGRRMDPLSSEEFFYSQFYNQRVKHL CYYHMRMKPYLCYQLEQFNQAPLKGCLLSEKGK
QHAETILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLA AFKRDRPDLLHIYTTSRLYFHWRPQKGLCSLWQ
SGLLVVDVMDLPQFTDCWTNFVNPKRPFWPKGLEII SRRTQRRLRIKESWGLQDVLVNDFGNLQLGPPMS
 (italic: nucleic acid editing domain)

Rat APOBEC-3

(SEQ ID NO: 275)

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKCDSPVSLHHGVFKNKDNIA
EICFLYWFDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNIRDPEPQONLCRL
VOEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPKRLLTNFRYQDSKLQEI LRPCYIPVPSSSSTLSNIC
TKGLPETRVCVERGRRMDPLSSEEFFYSQFYNQRVKHL CYYHGVKPYLCYQLEQFNQAPLKGCLLSEKGK
HAEILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLA AFKRDRPDLLHIYTTSRLYFHWRPQKGLCSLWQ
ILVDVMDLPQFTDCWTNFVNPKRPFWPKGLEII SRRTQRRLRIKESWGLQDVLVNDFGNLQLGPPMS
 (italic: nucleic acid editing domain)

Rhesus macaque APOBEC-3G

(SEQ ID NO: 276)

MVEPMDPRTFVSNFNNRPILSGLNTVWLCCVEVTKDPGSPPLDAKIFQGKVYSKAKYHPEMRFRLRFHKW
RQLHDQEKVTVWYVSWSPCTRCANSVATFLAKDPKVTLTIFVARLYYFWKPDYQQALRILCQKRGPHAT
MKJMNMYNEFQDCWNKFVDGRGPKFPKNLPKHYTLLQATLGELLRHLMDPGFTTSNFNNKPVWSQHE
TYLCYKVERLHNHTDWVPLNQHGRFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQYRVTCTFTSWSPCFS
CAQEMAKFISNNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYS EFEYCWDTFVDRQGRPFQP
WDGLDEHSQALSGRLRAI (italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

Chimpanzee APOBEC-3G

(SEQ ID NO: 277)

MKPHFRNPVERMYQDTSFSDNFYNRPILSRNTVWLCCVEVTKDPGSPPLDAKIFQGKLYPEAKDHPEMRF
FHWFSKWRKLHRDQEKVTVWYVSWSPCTKCTRDVATFLAEDPKVTLTIFVARLYYFWKPDYQQALRILCQK
DGPRATMKIMNYDEFQHCWSKVFSQRELPEWNNLPKYYILLHIMLGEILRHSMDPPTFTSFNFNNELWVR
GRHETYLYCVERLHNHTDWVLLNQHGRFLCNQAPHKHGFLERGRHAELCFLDVIPFWKLDLHQDYRVTCTFTS
WSPCFSCAQEMAKFISNPKHVSCLCIFAARIYDDQGRQCEGLRTLAKAGAKISIMTYSEFKHCDTFVDHQG
CPFPQPWDGLEEHSQALSGRLRAILQNQGN (italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

Green monkey APOBEC-3G

(SEQ ID NO: 278)

MNPQIRNMVEQMPDIFVYYFNNRPILSGRNTVWLCCVEVTKDPGSPPLDAKIFQGKLYPEAKDHPEMRF
HWFRKWRKLHRDQEKVTVWYVSWSPCTKCTRDVATFLAEDPKVTLTIFVARLYYFWKPDYQQALRILCQK
GGPHATMKIMNYDEFQHCWSKVFSQRELPEWNNLPKYYILLHIMLGEILRHSMDPPTFTSFNFNNELWVR
VSGQRETYLYCVERSHNDTWVLLNQHGRFLCNQAPHKHGFLERGRHAELCFLDVIPFWKLDLHQDYRVTCTFTS
WSPCFSCAQEMAKFISNPKHVSCLCIFTARIYDDQGRQCEGLRTLAKAGAKISIMTYSEFKHCDTFVDHQG
QGRPFQPWDGLDEHSQALSGRLRAI (italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

Human APOBEC-3G

(SEQ ID NO: 279)

MKPHFRNTVERMYRDTESYNFYNRPILSRNTVWLCCVEVTKDPGSPPLDAKIFQGQVYSELKYHPEMRF
HWFNSKWRKLHRDQEKVTVWYVSWSPCTKCTRDVATFLAEDPKVTLTIFVARLYYFWKPDYQQALRILCQK
DGPRATMKIMNYDEFQHCWSKVFSQRELPEWNNLPKYYILLHIMLGEILRHSMDPPTFTSFNFNNELWVR
GRHETYLYCVERMHNDTWVLLNQHGRFLCNQAPHKHGFLERGRHAELCFLDVIPFWKLDLHQDYRVTCTFTS
WSPCFSCAQEMAKFISNPKHVSCLCIFTARIYDDQGRQCEGLRTLAKAGAKISIMTYSEFKHCDTFVDHQG
CPFPQPWDGLDEHSQALSGRLRAILQNQEN (italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

Human APOBEC-3F

(SEQ ID NO: 280)

MKPHFRNTVERMYRDTESYNFYNRPILSRNTVWLCCVEVTKDPGSPPLDAKIFQGQVYSELKYHPEMRF
SWFCGNQLPAYKCFQITWFWVSWTPCPDCVAKLAEFLAEHPNVTLTISAARLYYWERDYRRALCRLSQAGA
RVKIMDDEEFAYCWENFVYSEGQPMPWYKFDNDYAFHLRTLKEILRNPMEAMYPIFYFHKNLRKAY
GRNESWLCTMEVVKHHSPVSWKRGVFRNQVDPETHCHAERCFSLWFCDILSPNTNYEVWTWYTSWSPCPE

-continued

CAGEVAEFLARHSNVNLTI~~T~~ARLYYFWDTDYQEGRLSLSQEGASVEIMGYKDFKYCWENFVNNDDEPFK
PWKGLKYNFLFLDSKLQEILE (italic: nucleic acid editing domain)

Human APOBEC-3B

(SEQ ID NO: 281)

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQVFKPKQYHAEM
CFLSWFCGNQLPAYKCFQITWFSWTPCPDCVAKLAEFLSEHPNVTLTISAA~~R~~LYYWERD~~Y~~RRALCRLSQA
GARVTIMDYEEFAYCWENFVYNEQQFMPWYKF~~D~~ENYAFLHRTLKEILRLMDPDTFTFNFNNDPL VLRR
RQTYLCYEVERLDNGTVLMDQHMGFLCNEAKNLLCGFYGRHAE~~L~~RFLDLVPSLQLDPAQIYRVTFW~~F~~ISWS
PCFSWG~~C~~AGEVRAFLQENTHVR~~L~~RIFAARIYDPLYKEALQMLRDAGAQVSIMTYDEF~~E~~YCWDTFVYRQ
GCPFPQWPDGLEEHQS~~A~~S~~G~~R~~L~~AILQNQGN (italic: nucleic acid editing domain)

Rat APOBEC-3B:

(SEQ ID NO: 1073)

MQPQGLGPNA~~M~~G~~P~~VCLGC~~S~~H~~R~~RPYSP~~I~~R~~N~~PLKKLYQQT~~F~~YFHF~~K~~N~~V~~YAWGRKNNFLCYEVNGMD~~C~~~~A~~
PVPLRQGVFRKQGHIAELCFIYWFHDKVLRVLSPMEEFKV~~T~~WYMSW~~S~~PCSKCAEQVARFLAHRNL~~S~~LA
IFSSRLYYL~~R~~NP~~N~~YQQ~~K~~LC~~R~~LIQEGVHVAAMDLPEFKKCWNKFVDNDGQ~~P~~FRP~~M~~RLRINFSFYDCKLQ
E~~I~~FSRMNLLREDVFYLFQFNNSHVR~~K~~PV~~Q~~NR~~Y~~Y~~R~~RSYLCYQLERANGQ~~E~~PLKG~~Y~~LLYKKGEQHVEILFLE
KMRSMELSQVR~~I~~TCYL~~T~~WSPCPNCARQ~~A~~FKKDHPD~~L~~ILRIYTSLYFWRKK~~Q~~KGLCTLWRSGI~~H~~V~~D~~
VMDLPQFADCWTNFVN~~P~~Q~~R~~PF~~R~~P~~W~~NELEKNSWRIQ~~R~~LLR~~I~~KE~~S~~EWGL

Bovine APOBEC-3B:

(SEQ ID NO: 1074)

DGWEVAFRSGTVLKAGVLGVSMTEGWAGSGH~~P~~GQACVWTPGTRNTMNL~~L~~RE~~V~~LFKQ~~Q~~GNQ~~P~~RV~~P~~P~~A~~
YYR~~R~~KT~~L~~CYQLQ~~K~~QRNDLTLDRGC~~F~~RN~~K~~Q~~R~~HAE~~I~~R~~F~~ID~~K~~IN~~S~~LDLN~~P~~SQSY~~K~~I~~C~~Y~~I~~TWSP~~C~~P~~N~~C~~A~~EL~~V~~N
FITRNNHLK~~L~~E~~F~~AS~~R~~LYFHW~~I~~K~~S~~FK~~M~~GLQ~~D~~LNAGIS~~V~~AVM~~T~~HTEF~~D~~C~~W~~EQ~~F~~V~~D~~NQ~~S~~R~~P~~F~~Q~~WD~~K~~LE~~Q~~Y
SASIRRLQ~~R~~IL~~T~~API

Chimpanzee APOBEC-3B:

(SEQ ID NO: 1075)

MNPQIRNPMEMWYQRTFY~~Y~~NNFENEPILYGRSYTWLCYEVKIRRG~~H~~SNLLWDTGVFRGQ~~M~~YSQ~~P~~EH~~H~~AM
CFLSWFCGNQLSAYKCFQITWFSWTPCPDCVAKLAKFLAEHPNVTLTISAA~~R~~LYYWERD~~Y~~RRALCRLS
QAGARVKIMDDEFAYCWENFVYNEQQFMPWYKF~~D~~ENYAFLHRTLKEIIRHLM~~D~~P~~T~~FTFNFNNDPLVL
R~~R~~HQTYLCYEVERLDNGTVLMDQHMGFLCNEAKNLLCGFYGRHAE~~L~~RFLDLVPSLQLDPAQIYRVTFW
ISWSPCPFWGCAGQVRAFLQENTHVR~~L~~RIFAARIYDPLYKEALQMLRDAGAQVSIMTYDEF~~E~~YCWDTF
VYRQGCPFPQWPDGLEEHQS~~A~~S~~G~~R~~L~~AILQVRASSLCMVP~~H~~RP~~PP~~QQ~~S~~PC~~P~~CL~~L~~C~~S~~E~~P~~PL~~G~~SL~~P~~T~~G~~P~~A~~P
SLPFLLTASFS~~P~~PPP~~P~~ASL~~P~~PL~~P~~SL~~S~~SPG~~H~~LP~~V~~PSFH~~S~~LT~~S~~CIQ~~P~~CC~~S~~R~~I~~RE~~T~~EG~~W~~AS~~V~~SK~~E~~GR~~D~~LG

Human APOBEC-3C:

(SEQ ID NO: 282)

MNPQIRNPMKAMYPGT~~F~~QFKNLWEANDRNETWL~~C~~FTVEGIKRRSVV~~S~~WKTGVFRNQ~~V~~D~~S~~ETHCHAER
CFLSWFCDDILSP~~N~~TYQ~~V~~TW~~T~~TSWSP~~C~~P~~D~~C~~A~~GEVAEFLARHSNVNL~~T~~IF~~T~~ARLYYFQ~~Y~~PC~~Y~~Q~~E~~GLR~~S~~LSQ~~E~~
VAVEIMDYEDFKYC~~W~~ENFVYNDNEPFKP~~W~~K~~G~~L~~K~~T~~N~~F~~R~~LLK~~R~~RL~~R~~ESL~~Q~~ (italic: nucleic
acid editing domain)

Gorilla APOBEC-3C:

(SEQ ID NO: 1076)

MNPQIRNPMKAMYPGT~~F~~QFKNLWEANDRNETWL~~C~~FTVEGIKRRSVV~~S~~WKTGVFRNQ~~V~~D~~S~~ETHCHAER
CFLSWFCDDILSP~~N~~TYQ~~V~~TW~~T~~TSWSP~~C~~P~~D~~C~~A~~GEVAEFLARHSNVNL~~T~~IF~~T~~ARLYYFQ~~Y~~PC~~Y~~Q~~E~~GLR~~S~~LSQ~~E~~
EGVAVKIMDYKDFK~~C~~YCWENFVYNDDEPFKP~~W~~K~~G~~L~~K~~T~~N~~F~~R~~LLK~~R~~RL~~R~~Q~~E~~ILE

Human APOBEC-3A:

(SEQ ID NO: 283)

MEASPASGPRHLM~~D~~PHI~~F~~TSN~~F~~NN~~G~~IGR~~H~~K~~T~~LCYEVERLDNGTSV~~K~~MDQH~~R~~GF~~L~~H~~N~~Q~~A~~KN~~N~~LLCGFYGRH
AELRF~~L~~DLVPSLQLDPAQIYV~~V~~TW~~F~~ISWSPCPFWGCAG~~V~~RAFLQ~~E~~NTHVR~~L~~RIFAARIYDPLYKEALQML
RDAGAQVSIMTYDEFKHC~~W~~D~~T~~F~~V~~DHQ~~G~~CP~~F~~Q~~P~~WD~~G~~L~~D~~E~~H~~SQ~~A~~S~~G~~R~~L~~AILQ~~N~~Q~~G~~N (italic:
nucleic acid editing domain)

Rhesus macaque APOBEC-3A:

(SEQ ID NO: 1077)

MDGSPASPRHLM~~D~~PD~~N~~TF~~F~~N~~N~~DL~~S~~VR~~G~~R~~H~~Q~~T~~LCYEVERLDNGT~~W~~VPM~~D~~ERR~~G~~FLCN~~K~~AKN~~V~~PC~~G~~
Y~~G~~CH~~V~~EL~~R~~FL~~C~~EV~~P~~S~~Q~~L~~D~~PA~~Q~~TY~~V~~TW~~F~~ISWSPCP~~R~~RG~~C~~AG~~Q~~VR~~V~~FL~~Q~~EN~~K~~VR~~L~~RIFAARIYDPLY
QE~~A~~RL~~T~~LRDAGAQVSIMTYEEFKHC~~W~~D~~T~~F~~V~~D~~Q~~GR~~P~~F~~Q~~P~~W~~D~~G~~L~~D~~E~~H~~SQ~~A~~S~~G~~R~~L~~AILQ~~N~~Q~~G~~N

Bovine APOBEC-3A:

(SEQ ID NO: 1078)

MDEYTF~~T~~EN~~F~~NNQ~~G~~W~~P~~S~~K~~TYLCYEMERLDG~~D~~AT~~I~~PL~~D~~EYK~~G~~FR~~N~~K~~G~~DQ~~P~~E~~K~~CHAEIYFL~~G~~K~~I~~HS~~W~~N~~L~~
DR~~N~~Q~~H~~Y~~R~~LT~~C~~FI~~S~~W~~S~~PC~~Y~~CD~~A~~Q~~K~~L~~T~~FL~~K~~EN~~H~~H~~I~~L~~A~~S~~R~~I~~Y~~TH~~N~~R~~F~~G~~C~~HS~~Q~~GL~~C~~EL~~Q~~AAG~~A~~R~~I~~TI~~M~~F~~E~~
FK~~H~~C~~W~~E~~T~~F~~V~~D~~H~~K~~G~~K~~P~~Q~~P~~W~~E~~GL~~N~~V~~K~~S~~Q~~AL~~C~~TEL~~Q~~AI~~L~~K~~T~~Q~~G~~N

Human APOBEC-3H:

(SEQ ID NO: 284)

MALLTAET~~F~~RLQ~~F~~NNK~~R~~LLRR~~P~~Y~~P~~R~~K~~ALLCYQLTP~~Q~~NG~~S~~PT~~R~~GYF~~E~~N~~K~~K~~CH~~AEI~~C~~FIN~~E~~I~~K~~SM~~G~~LD~~E~~
CYQ~~V~~T~~C~~Y~~L~~WT~~S~~PC~~S~~SC~~A~~WE~~L~~V~~D~~FIKAHD~~H~~LN~~L~~GI~~F~~AS~~R~~LYYHW~~C~~K~~P~~Q~~K~~GL~~R~~LL~~C~~GS~~Q~~V~~P~~VE~~V~~M~~G~~PK~~F~~AD
CWENF~~V~~D~~H~~E~~K~~PL~~S~~FS~~N~~PY~~K~~MLE~~L~~D~~K~~N~~S~~R~~A~~I~~K~~R~~R~~LER~~I~~K~~I~~PG~~V~~R~~A~~Q~~G~~RY~~M~~D~~I~~LC~~A~~E~~V~~
(italic: nucleic acid editing domain)

-continued

Rhesus macaque APOBEC-3H:

(SEQ ID NO: 1079)

MALLTAKTFSLQFNKNRVRVNPYPRKALLCYQLTPQNGSTPTRGHLKNKKDHAEIRFINKIKSMGLDET
 QCYQVTCYLTWSWPCPSCAGELVDFIAKHRHNLRFASRLYYHWRPNYQEGLLLLGSQVPVEMLPEFT
 DCWENFVDHKEPSPSFPNPKLEELDKNSQAICKRRLERIKSRSDVLENGLRSQQLGPVTSSSIRNSR

Human APOBEC-3D

(SEQ ID NO: 285)

MNPQIRNPMPERMRYRDTFYDNFNEPILYGRSYTWLCYEVKIKRGRSNLLWDGTGFRGPVLPKRQSNHRQE
 VYFRFENHAEMCFLSWFCGNRLPANRPFQITWFVSWNPCLPCVVVKFLAEPNVTLTISAARLYYYRDRD
 WRWVLLRLHKGARVKIMDYEDFAYCWENFVCNEGOPFMPWYKFDDNYASLHRTLKEILRNPMEMAP
 HIFYFHFKNLLKACGRNESWLCTMEVTKHHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTN
 YEVTVWYTWSWSPCPECAGEVAEFLARHSNVNLITIFTARLCYFWTDYQEGLCSLSQEGASVKIMGYKDFVSC
 WKNFVYSDDEPFKPWKLQTNFRLKRRRLREILQ (italic: nucleic acid editing domain)

Human APOBEC-1

(SEQ ID NO: 286)

MTSEKGPGSTGDPTRLRRRIEPWFEDVYDPRELRKEAICLLYEIKWGMRSRKIWIWRSRGKNTTNHVEVNFIKKFTS
 ERDFHPMSMCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIIYVARLFWHMDQQRNRQGLRDLVNSGVTIQI
 MRASEYYHCWRNFVNYPGDEAHWPQYPPWMLYALELHCIIILSLPPCLKISRRWQNHLTFRLHLQNC
 HYQTIPPHILLATGLIHPSAWR

Mouse APOBEC-1

(SEQ ID NO: 287)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGRHSVWRHTSQNTSNHVEVNFIKKFTT
 ERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRHPVVTLFIFIYIARLYHHTDQRNRQGLRDLISSGVTIQIMTE
 QEYCFCWRNFVNYPSPSNEAYWPRYPHLWVRLVYLEYCIILGLPPCLKILRRKQPQLTFFTIALQSCHYQRLP
 PHILWWATGLK

Rat APOBEC-1

(SEQ ID NO: 288)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGRHSIWRTSQNTNKHVENVFIEKFTT
 RYFCPNTRCSITWFLSWSPCGECSRAITEFLSRHPVVTLFIFIYIARLYHADPRNRQGLRDLISSGVTIQIMTE
 ESGYCFCWRNFVNYPSPSNEAHWPYRPHLWVRLVYLEYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLP
 PHILWWATGLK

Human APOBEC-2 :

(SEQ ID NO: 1080)

MAQKEEEAAVATEAASQNGDELNLDDPEKLIKELIELPPFEIVTGERLPANFFKFQFRNVEYSSGRNKTFLCY
 VVEAQSKGGQVQASRGYLEDEHAAAEEAFFNTILPAFDPALRYNTWVYSSPCAACADRILKTLK
 NLRLLILVGRLFMWEEPEIQAALKLKEAGCKLRIMKPQDFEYVWQNVEQEEGESKAFEPWEDIQENFL
 YYEEKLADILK

Mouse APOBEC-2 :

(SEQ ID NO: 1081)

MAQKEEEAAEAAAPASQNGDDLENLEDPEKLIKELIDLPPFEIVTGVRLPVNFFKFQFRNVEYSSGRNKTFLC
 YVVEVQSKGGQQATQGYLEDHAGAHAAEEAFFNTILPAFDPALRYNTWVYSSPCAACADRILKTLK
 TKNLRLLILVSLRFMWEPEVQAALKLKEAGCKLRIMKPQDFEYIWNQFVEQEEGESKAFEPWEDIQENFL
 LYEEKLADILK

Rat APOBEC-2 :

(SEQ ID NO: 1082)

MAQKEEEAAEAAAPASQNGDDLENLEDPEKLIKELIDLPPFEIVTGVRLPVNFFKFQFRNVEYSSGRNKTFLC
 YVVEAQSCKGGQVQASRGYLEDEHATNAEEAFFNSIMPTFDPALRYMTWVYSSPCAACADRILKTLK
 TKNLRLLILVSLRFMWEPEVQAALKLKEAGCKLRIMKPQDFEYIWNQFVEQEEGESKAFEPWEDIQENFL
 LYEEKLADILK

Bovine APOBEC-2 :

(SEQ ID NO: 1083)

MAQKEEEAAAAPASQNGGEVENLEDPEKLIKELIELPPFEIVTGERLPAHYFKFQFRNVEYSSGRNKTFLC
 VVEAQSKGGQVQASRGYLEDEHATNAEEAFFNSIMPTFDPALRYMTWVYSSPCAACADRILKTLK
 KNLRLLILVGRLFMWEPEIQAALKLKEAGCKLRIMKPQDFEYIWNQFVEQEEGESKAFEPWEDIQENFL
 YYEEKLADILK

Petromyzon marinus CDA1 (pmCDA1)

(SEQ ID NO: 289)

MTDAEYVRIHEKLDIYTFKKQFNNKKSVSRCYVLFELKRRGERRACFWGYAVNKPQSGTERGIHAEIFS
 RKVEEYLDRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRGNGHTLKIWACKLYYEKNARNQIGLWNL
 RDNGVGLNMVSEHYQCRKIFIQSSHNLNENRWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAV

Human APOBEC3G D316R_D317R

(SEQ ID NO: 290)

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTWLCYEVTKGPSRPPLDAKIFRGQVYSELKYHPEMRFF
 HWSKWRKLHRDQEYEWYWLSWSPCTKCTRDATFLAEPDKVTLTIFVARLYYFWDPEYQEARLRSLCQ
 KRDGPRATMKIMNYDEFQHCKSFVYSQRELFPWNLPKYYILLHIMLGEILRHSMDDPPTFTFNFNNEPW
 VRGRHETYLCYEVERMHDWTWLLNQRGFLCNQAPHKHGPLEGRHAELCFLDVIPFWKLDDQDYRVT
 CFTSWSPCFSCAQEMAKFISKNHVSLCIFTARIYRQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVD
 HQGCPFQFWGLDEHSQDLSGRLRAILQNQEN

-continued

Human APOBEC3G chain A

(SEQ ID NO: 291)

MDPPTFTFNFNNEP WVGRHETYL CYEVERM HNDT WVL NQRRGFLCNCAPHKHG FLEGRHAE LCF LDV
 IPFWKLDL DQDYR VTCFTS WSP CFSCA QEMAK FIS KNKH VSL CIFTARIY DDQ GRC QEG LTL AEAGAKI S
 MTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQDLSGRL RAIL Q

Human APOBEC3G chain A D120R_D121R

(SEQ ID NO: 292)

MDPPTFTFNFNNEP WVGRHETYL CYEVERM HNDT WVL NQRRGFLCNCAPHKHG FLEGRHAE LCF LDV
 IPFWKLDL DQDYR VTCFTS WSP CFSCA QEMAK FIS KNKH VSL CIFTARIY RRQ GRC QEG LTL AEAGAKI S
 MTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQDLSGRL RAIL Q

Non-limiting examples of fusion proteins/nucleobase editors are provided.
His6-rAPOBEC1-XTN-dCas9 for *Escherichia coli* expression

(SEQ ID NO: 293)

MGSSHHHHHHMSSETGPVAVDPTLRRRIEPEHEFEVFFDPRELRKETCLLYEINWGRHSIWRHTSQNTNKH
 VEVNFIKEFTTERYFCPNTRCSITWFLSWSPCGECRSAITEFLSRYPHTLFYIARLYHHADPRNRQGLRDLI
 SSGVTIQIMTEQESGYCRNFVNYSNSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNLRLRKQPOLTFFTI
 ALQSCHYORLPHILWATGLKGSGSETPGTESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG
 NTDRHSIKKNLIGALLFDSGETAETTRLKRTARRRYTRRKNRICYLQEIIFSNEMAKVDDSFTHRLEESFLVE
 EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DRLIYLALAHMIKFRGHFLIEGDLNP DNSD
 VDKLFQLVQTYNQLFEENPINASGVDAKIALSRSRLENLIAQLPGEKKNGLFGNLIALSGLTPNFK
 SNFDLAEDAKLQLSKDTYD DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHHQDLTLLKALV RQOLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIFKPILEKMDGTEELLVKLNR
 DLLRKQRTFDNGSIPHQIHLGELHAILRQEDFYPFLKDNRK EKIEKILTFRIPYYVGPLARGNSRFAWMTRKS
 EFTITPWNFEEVVDKGASAQSIFERMTNFDKLPNEKVKLPKHSLLYEYFTVYNELTKVVKVYVTEGMRKP AFL
 SGEQKKAIIVDLFKTNRKVTQVQLKIECFDSVEISGVEDRFNASLGTYHDLLKI KDKDFLDNEEN
 EDILEDIVLTLTLEDREMIERLKTYAHLFDKVMQKLRRTGWGRLSRKLINGIRDQSGKTILDPLK
 SDGFANRNFMOLIHDDS LT FKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKV VDELVKVGMGRH
 KOPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYV
 QELDINRLSDYDVAIPQSQSLKDDSDNPKVQLKTDNSKVRNRGKSDNVPSSEEVVKMKNYWRQLLNAKLITQR
 KFDNLTKAERGGLS ELDKAGFIKRQLVETRQITKVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFR
 KDFQFYKVERINNNYHHA DAYLNAV VGTALIKKPKLESEFVYGDYKVDVRKMIAKSEQEIGKATAKYF
 FYSNIMNFFKT EITLANGEIRKP LIETNGETGEIVWDKG R DFA T VRKVLSMPQVNIVKKT EVQTGGFSKESI
 LPKRNSDKLIAK KKDWDPKK YGGFDSP TVA YSVL VVAKVEKGKSKKLKSVKELLG ITIMER SFEKN PIDF
 LEAKGYKEVKKD LIIKLPKYSIFELENGRKRMLASAGELQKG NELALPSKVNFYLYASHYEKLKG SPEDN
 EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPA AFK
 YFDTTIDRKRYTSTKEVLDATLHQ SITG LYETRIDLSQLGGDSGGSPKKKRKV

rAPOBEC1-XTN-dCas9-NLS for Mammalian expression

(SEQ ID NO: 294)

MSSETGPVAVDPTLRRRIEPEHEFEVFFDPRELRKETCLLYEINWGRHSIWRHTSQNTNKHVEVNFIKEFTT
 ERYFCPNTRCSITWFLSWSPCGECRSAITEFLSRYPHTLFYIARLYHHADPRNRQGLRDLI SSGVTIQIMTE
 QESGYCRNFVNYSNSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNLRLRKQPOLTFFTI ALQSCHYORL
 PPHILWATGLKGSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKN
 LIGALLFDSGETAETTRLKRTARRRYTRRKNRICYLQEIIFSNEMAKVDDSFTHRLEESFLVEEDKKHERHPIF
 GNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DRLIYLALAHMIKFRGHFLIEGDLNP DNSD VDKLFQIQLVQ
 YNQLFEENPINASGVDAKIALSRSRLENLIAQLPGEKKNGLFGNLIALSGLTPNFKSNFDLAEDAKL
 QLSKDTYD DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIFKPILEKMDGTEELLVKLNR EDLLRKQRTFDN
 GSIPHQIHLGELHAILRQEDFYPFLKDNRK EKIEKILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEV
 VDKGASQAFSFERMTNFDKNLPEVKLPKHSLLYEYFTVYNELTKVVKVYVTEGMRKP AFLS GEQKKAI VDL
 LFKTNRKVTQVQLKIECFDSVEISGVEDRFNASLGTYHDLLKI KDKDFLDNEENEDIVLTLT
 LFEDREMIERLKTYAHLFDKVMQKLRRTGWGRLSRKLINGIRDQSGKTILDPLKSDGFANRNFM
 QLIHDDS LT FKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKV VDELVKVGMGRKPE NIVIEMARE
 NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVQELDINRLSDY
 DVDAIPQSQSLKDDSDNPKVQLKTDNSKVRNRGKSDNVPSSEEVVKMKNYWRQLLNAKLITQDKFDNLTKAERG
 GLSELDKAGFIKRQLVETRQITKVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREI
 NNYYHA DAYLNAV VGTALIKKPKLESEFVYGDYKVDVRKMIAKSEQEIGKATAKYF YSNI MNFFKT
 EITLANGEIRKP LIETNGETGEIVWDKG R DFA T VRKVLSMPQVNIVKKT EVQTGGFSKESI LPKRNSDKLIA
 RKKDWDPKK YGGFDSP TVA YSVL VVAKVEKGKSKKLKSVKELLG ITIMER SFEKN PIDF LEAKGYKEV
 KDLIIKLPKYSIFELENGRKRMLASAGELQKG NELALPSKVNFYLYASHYEKLKG SPEDNEQKQLFVEQH
 KHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPA AFK YFDTTIDRK
 TSTKEVLDATLHQ SITG LYETRIDLSQLGGDSGGSPKKKRKV

hAPOBEC1-XTN-dCas9-NLS for Mammalian expression

(SEQ ID NO: 295)

MTSEKG PSTGDP TLRRRIE PWEFVYD PPRELRKEAC ALLY EIKWGM SRK IWI RSSG KNT TNH VEVNFI KKFTSE
 RD PHP SMC SITWFL SWSP CWC SQA IREFL RHPG VTLV IYV ARLF WFM DQ QN RQGL RD LV NSG VT IQI
 MRASEYYHCWRFN VNY PGD EAHW PQY PPLW MM LYALEL H C I ILS LPP CL K I S R R W QN H LTF F RL H L Q N C
 HYQTIP PHILLATGLIHP SVA RSG SET PGT SESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGN
 TDRHSIKKNLIGALLFD SGETAETTRLKRTARRRYTRRKNRICYLQEIIFSNEMAKVDDSFTHRLEESFLVEEDKK
 HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DRLIYLALAHMIKFRGHFLIEGDLNP DNSD VDKLFQ
 LVQTYNQLFEENPINASGVDAKIALSRSRLENLIAQLPGEKKNGLFGNLIALSGLTPNFKSNFDLAEDA
 KLSKDTYD DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 LVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIFKPILEKMDGTEELLVKLNR EDLLRKQRTFDNGSIPH

- continued

QIHLGELHAILRRQQEDFYPFLKDNREKIEKILTFRIPIYYVGPARGNSRFAWMTRKSEETITPWNFEVVVDKGA
 SAQSFIERMNTNFDKNLNEKVLPKHSLLYEFYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVT
 VQLKEDYFKKECFDSVEISGVEDRFNASLGTYHDLLKI KDKDLDNEENEDILED -
 VLTTLFEDMIEERLK
 Tyahlfddkvmqkrrytwgrlsrklngirdkosgktildflksdgcfa nrfmqlih dslt fkd i qk
 aqvgqgds lhehi anlag spaik gk lqtvkv vde lkv mrg h k peniv emare nq t qk gq k ns rem
 kri eeg i kelg s q l k e h p v e n t q l o n e k l y l y l q n g r d m v d q e l d i n r l s d y d v a i v p q s f l k d d s i n k
 vlt r s d k n r g k s d n v p s e e v v k k m k n y w r q l l n a k l i t o r k f d n l t k a e r g l s e l d k a g f i k r o l v e t r o i t
 khvaqil d s r m n t k y d e n d k l i r e v k t l k s k l v s d f r k d f q f y k v r e i n n y h h a d y l n a v v g t a l i k k y p
 kles e f v y g d y k v y d v r k m i a k s e q e i g k a t a k y f f y s n i m n f f k t e i t l a n g e i r k r p l i e t n e g e b i v w d k g
 rd fat v r k v l s m p q v n i v k k t e v q t g g f s k e s i l p k r n s d k l i a r k k d w d p k y g f d s p t v a y s v l v v a k v e
 kgk s k k l s k v e l l g i t i m e r s f e k n p i d f l e a k g y k e v k k d l i i k l p k y s l f e l e n g r k r m l a s a g e l q k g n e l
 alpskyvnflyashyeklgspedneqkqlfveqkhyldei i eqi sefskrv il a dan ldk v l s a y n k h r d k p i
 alpskyvnflyashyeklgspedneqkqlfveqkhyldei i eqi sefskrv il a dan ldk v l s a y n k h r d k p i
 rqaeniihlftltnlgapaa fk y f d t t i d r k r y t s t k e v l d a t l i h q s i t g l y e t r i d l s q l g g d s g g p k k r k
 v

rAPOBEC1-XTEN-dCas9-UGI-NLS

(SEQ ID NO: 296)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGRHSIWRTSQNTNKHVENVFIEKFTT
 ERYFCPNTRCSITWFLSWSPCGECRSAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTE
 QESGYCWRNFVNYSPESEAHWPYRPHLWVRLYVLELYCIIILGLPPCLNILRRKQPLTFTTIALQSCHYQRL
 PPHILWATGLKGSSETPGTSESATPESDKKYSIGLAIGTNSGWAVITDEYKVPSKKFKVLGNTDRHSI KKN
 LIGALLFDGETAEATRLKRTARRRYTRKRNRCIYLQEI FSNEMAKVDDSFTHRLEESFLVEEDKKHERHP IF
 GNIVDEVAYHEKPYTIYHLRKKLVDSTDKDADRLLIYLA LAHMIKFRGHF LIEGDLNP DNSDV DKLFIQLVQT
 YNQLFEENPINASGVDAKILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKL
 QLSKDTYDDLDNL LAQIDQYADLFLAAKNLSDA ILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQOLPEKYE IFFDQSCKNGYAGYIDGGASQEEFYKFIPKILEMDGTEELLVKLNRD L RQRTFDN
 GSIPHQIHLGELHAILRRQQEDFYPFLKDNREKIEKILTFRIPIYYVGPARGNSRFAWMTRKSEETITPWNFEV
 VDKGASAQSFIERMNTNFDKNLNEKVLPKHSLLYEFYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDL
 LFKTNRKVTVKQLKEDYFKKECFDSVEISGVEDRFNASLGTYHDLLKI KDKDLDNEENEDILEDIVLTLT
 LFEDREMIERLKTYAHLFDDKVMQKRRYTWGRLSRKLNGIRDKOSGKTILDFLKS DGCFA NRFM
 QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKGK LQTVKV VDELVKVMGRHK PENIVI MARE
 NQTTQKGOKNSRERMKRIEEGIKELGSQLKEH PVENTQLONEKLYL YLQNGRDMV DQELD INR LSDY
 DVDAIPQPSFLKQDSEATRKLKRTARRYTRKRNRCIYLQEI FSNEMAKVDDSFTHRLEESFLVEEDKKHERHP IF
 GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTK YDENDKLIREVKVITLKS KLVSDFRKDFQFYK VREI
 NNHHAAHDAYLNAVVGTLALIKKPKLESEFVYGDYK VYDVRKMIAKS E QEIGKATAK YFFY SNI MNNF K
 ETTLANGEI RKRPLIETNGETEIVWDKGDFATVRKVLSPQVNI VVKTEVQ T GGF SKES I LPK RNS DKL IA
 RKKWDPKYGGFDSPVAYSVL VVAKVEKGSKKL SKV ELLG I T M E R S F E K N P I D F L E A K G Y K E V K
 KDLI I KLPKYSLFELENGRKRMLASAGELQKGNELAPSKVNVFLYASHYEKLGSPEDNEQKQLFVEQH
 KHYLDEIIEQI SEFSKRV IL A DAN LDK V L S A Y N K H R D K P I R Q A E N I I H L F T L T N L G A P A A F K Y F D T T I D R K R Y
 TSTKEVLDATL I H Q S I T G L Y E T R I D L S Q L G G D S G G S T N L S D I I E K E T G K Q L V I Q E S I L M L P E E V E V I G N K P E S D
 LVHTAYDESTDENVMLLTSDAPEYKPVWALVIQDSNGENKIKMLSGGSPKKRKV

rAPOBEC1-XTEN-Cas9 nickase-UGI-NLS)

(BE3, SEQ ID NO: 297)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGRHSIWRTSQNTNKHVENVFIEKFTT
 ERYFCPNTRCSITWFLSWSPCGECRSAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTE
 QESGYCWRNFVNYSPESEAHWPYRPHLWVRLYVLELYCIIILGLPPCLNILRRKQPLTFTTIALQSCHYQRL
 PPHILWATGLKGSSETPGTSESATPESDKKYSIGLAIGTNSGWAVITDEYKVPSKKFKVLGNTDRHSI KKN
 LIGALLFDGETAEATRLKRTARRYTRKRNRCIYLQEI FSNEMAKVDDSFTHRLEESFLVEEDKKHERHP IF
 GNIVDEVAYHEKPYTIYHLRKKLVDSTDKDADRLLIYLA LAHMIKFRGHF LIEGDLNP DNSDV DKLFIQLVQT
 YNQLFEENPINASGVDAKILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKL
 QLSKDTYDDLDNL LAQIDQYADLFLAAKNLSDA ILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQOLPEKYE IFFDQSCKNGYAGYIDGGASQEEFYKFIPKILEMDGTEELLVKLNRD L RQRTFDN
 GSIPHQIHLGELHAILRRQQEDFYPFLKDNREKIEKILTFRIPIYYVGPARGNSRFAWMTRKSEETITPWNFEV
 VDKGASAQSFIERMNTNFDKNLNEKVLPKHSLLYEFYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDL
 LFKTNRKVTVKQLKEDYFKKECFDSVEISGVEDRFNASLGTYHDLLKI KDKDLDNEENEDILEDIVLTLT
 FEDEMEERLKTYAHLFDDKVMQKRRYTWGRLSRKLNGIRDKOSGKTILDFLKS DGCFA NRFM
 LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKGK LQTVKV VDELVKVMGRHK PENIVI MARE
 QT T QKGOKNSRERMKRIEEGIKELGSQLKEH PVENTQLONEKLYL YLQNGRDMV DQELD INR LSDY
 DVHIVPQPSFLKQDSEATRKLKRTARRYTRKRNRCIYLQEI FSNEMAKVDDSFTHRLEESFLVEEDKKHERHP IF
 LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTK YDENDKLIREVKVITLKS KLVSDFRKDFQFYK VREI
 NYHHAAHDAYLNAVVGTLALIKKPKLESEFVYGDYK VYDVRKMIAKS E QEIGKATAK YFFY SNI MNNF K
 ETTLANGEI RKRPLIETNGETEIVWDKGDFATVRKVLSPQVNI VVKTEVQ T GGF SKES I LPK RNS DKL IA
 RKKWDPKYGGFDSPVAYSVL VVAKVEKGSKKL SKV ELLG I T M E R S F E K N P I D F L E A K G Y K E V K
 KDLI I KLPKYSLFELENGRKRMLASAGELQKGNELAPSKVNVFLYASHYEKLGSPEDNEQKQLFVEQH
 KHYLDEIIEQI SEFSKRV IL A DAN LDK V L S A Y N K H R D K P I R Q A E N I I H L F T L T N L G A P A A F K Y F D T T I D R K R Y
 TSTKEVLDATL I H Q S I T G L Y E T R I D L S Q L G G D S G G S T N L S D I I E K E T G K Q L V I Q E S I L M L P E E V E V I G N K P E S D I
 LVHTAYDESTDENVMLLTSDAPEYKPVWALVIQDSNGENKIKMLSGGSPKKRKV

pmCDA1-XTEN-dCas9-UGI (bacteria)

(SEQ ID NO: 298)

MTDAEYVRIHEKLDIYTFFKKQFFNNKKS VSHRCYVLFELKRRERRACFWGYAVNKPQ
 SGTERGIHAEIFSIRKVEEYL RDNP GQFTI NWYSSWSPCA DCA EK I LEWYNQELRGNGHT
 LKI WACKLYYEKNARNQIGLWNL RDNGVGLNMVMEHYQCCR KIFI QSSH NQLNENR
 WLEKTLKRAEKRSELSIMI QV KILH TT KSPAVSGSETPGTSESATPESDKKYSIGLAIGT
 NSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDGETAEATRLKRTARRY
 TRKRNRCIYLQEI FSNEMAKVDDSFTHRLEESFLVEEDKKHERHP I FGNIVDEVAYHEK
 PTIYHLRKKLVDSTDKDADRLLIYLA LAHMIKFRGHF LIEGDLNP DNSDV DKLFIQLVQTY
 YNQLFEENPINASGVDAKILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFK
 SNFDLAEDAKLQLSKDTYDDLDNL LAQIDQYADLFLAAKNLSDA ILLSDILRVNTEIT

- continued

KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSCKNGYAGYIDGGASQEEF
 YKFIKPITLEKMDGTEELLVKLNRDPLLKRQRTFDNGSIPHQIHLGELHAILRQEDFYPFL
 KDNREKIEKILTFRIPIYYVGPALARNSRFAAMTRKSEETITTPWNFEVVDKGASAQSIE
 RMTNFDKNLNPNEVKLPKHSSLYEYFTVYNELTKVVKVTEGMRKPAFLSGEQKKAIVDL
 LFKTNRKTVKQLKEDYFKKIICFDSEVEISGVEDRPNASLGTYHDLLKIIKDKDFLDNEE
 NEDILEDIVLTLTFEDREMIERLKTYAHLFFDKVMKQLKRRRTGWRSLRKLINGIR
 DKQSGKTIIDFLKSDGFANRFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP
 AIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE
 LGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKD
 DSIDNKVLTRSDKNRGSNDVPSEEVVKMKNYWRQLLNALKLITQRKFDSLTKAERGG
 LSELDAKFGIKRQLVETRQITKVAQILDLSRMTKYDENDKLIREVKVITLKSCLVSDPR
 KDFQFYKVREINNNYHHAHDAYLNAVVTALIKKPKLESEFVYGDYKVDVRKMIAKS
 EOEIGKATAKYFFYSNIMNFFKETITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
 VLSMPQVNIVKTEVQTGGFSKESILPKRNSDKLIAKRDWDPKKYGGFSPTVAYSVL
 VVAKVEKGSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI1KLPKYSIF
 LENGRKMLASAGELOQGNELALPSKYVNFPLYASHYEKLKGSPEDNEQQLFVBQHK
 HYLDEIEEQISEFSKRVILADANLDKVLSAVKNHRDKPIREQAENIIHLFTLTNLGAPAAFK
 YFDTTIDRKRYTSTKEVLDAATLHQSIITGLYETRIDLSQLGGDSGGSMTNLSDIIEKETGK
 QLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPVWALVIQD
 SNGENKIKML

pmCDA1-XTEN-nCas9-UGI-NLS (mammalian construct) (SEQ ID NO: 299) :
 MTDAEYVRIHEKLDIYTFFKQFFNNKKSVCNSHRCYVLFELKRGERRACFWGYAVNKPQ
 SGTERGIHAEISIRKVEEYLDRDNGPQFTINWYSSWSPCADCAEKILEWYNQELRGNGHT
 LKWWACKLYYEKARNQIGLWNLDRDNGVGLNMVSEHYQCCRKIFIQSSHNLQNENR
 WLEKTLKRAEKRRSELSIMIQVKILHHTKSPAVSGSETPGTSESATPESDKYSIGLAIGT
 NSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDGETAEATRLKRTARRYY
 TRRKNRICYLQEIFSNEMAKVDDSSFFHRLAESFLVEEDKKHERHPIFGNIVDEVAYHEKY
 PTIYHLRKVLSTDADLRLIYLALAHMIKFRGHPLIEGDLNPDPNSDVKLFIQLVQTY
 NQJFEEPINASGVDAKLARLSRSRLENLIAJLPGEEKNGLFGNLIALSGLTPNPK
 SNPDLAEDAKLQSLKDTYDDDDLNLLAQIGDQYADLFLAAKNLSDAIISSDILRVNTEIT
 KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSCKNGYAGYIDGGASQEEF
 YKFIKPITLEKMDGTEELLVKLNRDPLLKRQRTFDNGSIPHQIHLGELHAILRQEDFYPFL
 KDNREKIEKILTFRIPIYYVGPALARNSRFAAMTRKSEETITTPWNFEVVDKGASAQSIE
 RMTNFDKNLNPNEVKLPKHSSLYEYFTVYNELTKVVKVTEGMRKPAFLSGEQKKAIVDL
 LFKTNRKTVKQLKEDYFKKIICFDSEVEISGVEDRPNASLGTYHDLLKIIKDKDFLDNEE
 NEDILEDIVLTLTFEDREMIERLKTYAHLFFDKVMKQLKRRRTGWRSLRKLINGIR
 DKQSGKTIIDFLKSDGFANRFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP
 AIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE
 LGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKD
 DSIDNKVLTRSDKNRGSNDVPSEEVVKMKNYWRQLLNALKLITQRKFDSLTKAERGG
 LSELDAKFGIKRQLVETRQITKVAQILDLSRMTKYDENDKLIREVKVITLKSCLVSDPR
 KDFQFYKVREINNNYHHAHDAYLNAVVTALIKKPKLESEFVYGDYKVDVRKMIAKS
 EOEIGKATAKYFFYSNIMNFFKETITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
 VLSMPQVNIVKTEVQTGGFSKESILPKRNSDKLIAKRDWDPKKYGGFSPTVAYSVL
 VVAKVEKGSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI1KLPKYSIF
 LENGRKMLASAGELOQGNELALPSKYVNFPLYASHYEKLKGSPEDNEQQLFVBQHK
 HYLDEIEEQISEFSKRVILADANLDKVLSAVKNHRDKPIREQAENIIHLFTLTNLGAPAAFK
 YFDTTIDRKRYTSTKEVLDAATLHQSIITGLYETRIDLSQLGGDSGGSMTNLSDIIEKETGK
 QLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPVWALVIQD
 GENKIKMLSGGSPKKRKV

huAPOBEC3G-XTEN-dCas9-UGI (bacteria)

(SEQ ID NO: 300)

MDPTPTFNFNNEPWVGRHETYLCEVERMHNDTWVLLNQRRGFLCNQAPHKKHGF
 EGRHAELCFLDVIPFWKLDQDYRVTCTFTSWSPCFSCAQEMAFISKNKHVSCLCIFTAR
 IYDQGRQEGLRTLAEAGAKISIMTYSEFKHCDTVDHGCPQFWPDGLDEHSQDL
 SGRLRAILQSGSETPGTSESATPESDKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG
 NTDRHSIKKNLIGALLFDGETAEATRLKRTARRYYTRRKNRICYLQEIFSNEMAKVDDS
 FPHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKPTIYHLRKLVDSSTDADRLIYL
 ALAHMIKFRGHPLIEGDLNPDPNSDVKLFIQLVQTYNQLEENPINASGVDAKIALSARL
 SKSRRLENLIAJLPGEEKNGLFGNLIALSGLTPNPKNSFDLAEDAKLQLSKDTYDDDD
 NLLAQIGDQYADLFLAAKNLSDAIISSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPKEKEIFFDQSCKNGYAGYIDGGASQEEFYKFIKPITLEKMDGTEELLVKLNR
 DLLRKQRTFDNGSIPHQIHLGELHAILRQEDFYPFLKDNREKIEKILTFRIPIYYVGPALAR
 GNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPEVKLPKHSSLYE
 YFTVYNELTKVVKVTEGMRKPAFLSGEQKKAIVDLLFKTNRKTVKQLKEDYFKKIECF
 DSVEISGVEDRPNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTFEDREMEERLK
 TYAHLFDDKVMKQLKRRRTGWRSLRKLINGIRDQSGKTIIDFLKSDGFANRFM
 LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRH
 KPNIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSGQILKEHPVENTQLQNEKLYLY
 YLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGSNDVP
 EEVVKMKNYWRQLLNALKLITQRKFDSLTKAERGGLELDKAGFIKRQLVETRQITKH
 VAQILDLSRMTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRREINNNYHHAHDAYL
 NAVVGTALIKKPKLESEFVYGDYKVDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLMSMPQVNIVKTEVQTGGFSKE
 SILPKRNSDKLIAKRDWDPKKYGGFSPTVAYSVLVVAKEVKGSKKLKSVKELLGIT

-continued

IMERSSFEKNP1DFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRLMASAGELQGNELA
 LPSKYVNFYLYASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANL
 DVLSAYNKHRDKPIREQAENI1HLFTLTNLGAPAAFKYFDTIDRKRYTSTKEVLDAI
 HQSITGLYETRIDLSQLGGDSGGSTMNLSDII1EKETGQVLVIQESILMLPEEEVIGNKPE
 SDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV

huAPOBEC3G-XTEN-nCas9-UGI-NLS (mammalian construct)

(SEQ ID NO: 301)

MDPPTFTFNFPNNEPWVRGRHETYL CYEVERMHNDTWVLLNQRGFLCNQAPHKGFL
 EGRHAELCFLDVIPFWKLDLQDYRVTCTSWSPCFSCAQEMAKFISKNKHVSCLCIFTAR
 IYDQGRQEGRLTLABAGAKISIMTYSEFKHCWDTFVDHQGCPQFWPDGLDEHSQDL
 SGRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG
 NTRHS1KKNLIGALLFDGETAEATRLKRTARRRYTRRKNRICYLQEIIFSNEMAKVDDDS
 FFRLEESFLVEEDKKHERHP1FGNIDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYL
 LAHMIKFRGHFLIEGLNPDNSDVKLFIQLVQTYNQLEENPINASGVDAKIALSARL
 SKSRRLENLIAQLPGEKKNGLFGNLLIASLGLTPNPKNSNFDAEADAKLQLSKDTYDDLD
 NLLAQ1GDQYADLFLAAKNLSDAILLSDILRVNTETIKAPLSSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKE1FFDQSNSKNGYAGYIDGGASQEEFYKFKPILEKMDGTEELLVKNRE
 DLLRKQRTFDNGS1PHQIHLGELHAILRRQEDFYPLKDNRKREKIEKILTFRIPYVGPLAR
 GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKLPNEKVLPKHSSLYE
 YFTVYNELTKVYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKTVQKLDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDEDIVLTLTFEDREMIEERLK
 TYAHLFDKVMQLKRRRTGWGRSLRKLINGIRDQSGKTILDPLKSDGFANRNMQ
 LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA1KKGILQTVKVVDELVKVMGRH
 KPNIVIEMARENQTTOKGQKNSRERMKRIEGI1KELGSQ1LKEHPVENTQLNQEKLYLY
 YLQNQGRDMYVQDQELDINRLSDYDHDVHQPSFLKDDSIDNKVLTRSDKNRGKSDNVPS
 EEVVKKMKNYWRQLLNALKLITQRFKDNLTAKERGGLSELDAKQF1KQLVETRQITKH
 VAQILDLSRMNTYKDENDKLIREVKVITLKSKLVDPRKDFQFYKVREINNNYHHADAYL
 NAVVGTALIKKYPKLESFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFK
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATRKVLSPMQVNIVKKTEVQTTGFSKE
 SILPKRNSDKL1ARKWDWPKYYGGFDSPVAYSVLVAKEVKGSKKLKSVKELLGIT
 IMERSSFEKNP1DFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRLMASAGELQGNELA
 LPSKYVNFYLYASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANL
 DVLSAYNKHRDKPIREQAENI1HLFTLTNLGAPAAFKYFDTIDRKRYTSTKEVLDAI
 HQSITGLYETRIDLSQLGGDSGGSTMNLSDII1EKETGQVLVIQESILMLPEEEVIGNKPE
 DILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV

huAPOBEC3G (D316R_D317R) -XTEN-nCas9-UGI-NLS (mammalian construct)

(SEQ ID NO: 302)

MDPPTFTFNFPNNEPWVRGRHETYL CYEVERMHNDTWVLLNQRGFLCNQAPHKGFL
 EGRHAELCFLDVIPFWKLDLQDYRVTCTSWSPCFSCAQEMAKFISKNKHVSCLCIFTAR
 IYRQGRQEGRLTLABAGAKISIMTYSEFKHCWDTFVDHQGCPQFWPDGLDEHSQDL
 GRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG
 TDRHS1KKNLIGALLFDGETAEATRLKRTARRRYTRRKNRICYLQEIIFSNEMAKVDDDS
 FFRLEESFLVEEDKKHERHP1FGNIDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYL
 LAHMIKFRGHFLIEGLNPDNSDVKLFIQLVQTYNQLEENPINASGVDAKIALSARL
 SKSRRLENLIAQLPGEKKNGLFGNLLIASLGLTPNPKNSNFDAEADAKLQLSKDTYDDLD
 NLLAQ1GDQYADLFLAAKNLSDAILLSDILRVNTETIKAPLSSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKE1FFDQSNSKNGYAGYIDGGASQEEFYKFKPILEKMDGTEELLVKNRE
 DLLRKQRTFDNGS1PHQIHLGELHAILRRQEDFYPLKDNRKREKIEKILTFRIPYVGPLAR
 GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKLPNEKVLPKHSSLYE
 YFTVYNELTKVYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKTVQKLDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDEDIVLTLTFEDREMIEERLK
 TYAHLFDKVMQLKRRRTGWGRSLRKLINGIRDQSGKTILDPLKSDGFANRNMQ
 LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA1KKGILQTVKVVDELVKVMGRH
 KPNIVIEMARENQTTOKGQKNSRERMKRIEGI1KELGSQ1LKEHPVENTQLNQEKLYLY
 YLQNQGRDMYVQDQELDINRLSDYDHDVHQPSFLKDDSIDNKVLTRSDKNRGKSDNVPS
 EEVVKKMKNYWRQLLNALKLITQRFKDNLTAKERGGLSELDAKQF1KQLVETRQITKH
 VAQILDLSRMNTYKDENDKLIREVKVITLKSKLVDPRKDFQFYKVREINNNYHHADAYL
 NAVVGTALIKKYPKLESFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFK
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATRKVLSPMQVNIVKKTEVQTTGFSKE
 SILPKRNSDKL1ARKWDWPKYYGGFDSPVAYSVLVAKEVKGSKKLKSVKELLGIT
 IMERSSFEKNP1DFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRLMASAGELQGNELA
 LPSKYVNFYLYASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANL
 DVLSAYNKHRDKPIREQAENI1HLFTLTNLGAPAAFKYFDTIDRKRYTSTKEVLDAI
 HQSITGLYETRIDLSQLGGDSGGSTMNLSDII1EKETGQVLVIQESILMLPEEEVIGNKPE
 DILVHTA YDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV

Base Editor 4 (BE4; APOBEC1-linker(32 aa)-Cas9n(D10A)-linker(9 aa)-
 UGI-linker(9 aa)-UGI)

(SEQ ID NO: 1084)

MSETGPVAVDPLRRRIEPHEFEVFFDPRELRKETCLLYEINWGRHSIWRHTSQNTNKHVENVFIEKPTT
 ERYFCPNTRCSITWFLSWSPCGECSCRAITEFLSRPHVTLIFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTE
 QESGYCWRFVNYSPSNEAHWPRYPHLWVRLVLYELCYIILGLPPCLNILRRKQPLTFFTIALQSCHYQRL
 PPHILWATGLKSGGGSSGGSGSETPGTSESATPESGGGGSSGSDKYSIGLAIGTNSVGWAVITDEYKVPSSKK
 FVFLGNTDRHS1KKNLIGALLFDGETAEATRLKRTARRRYTRRKNRICYLQEIIFSNEMAKVDDSFHRLEE
 SFLVEEDKKHERHP1FGNIDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGLDN
 PDNSDVKLFIQLVQTYNQLEENPINASGVDAKIALSARLSKSRRLNLLAQLPGEKKNGLFGNLLIASLGL
 TPNFKSNFDLAEDAKLQLSKDTYDDLDNLNLLAQ1GDQYADLFLAAKNLSDAILLSDILRVNTETIKAPLSSAS
 MIKRYDEHHQDLTLLKALVRQQLPKYKE1FFDQSNSKNGYAGYIDGGASQEEFYKFKPILEKMDGTEELLV
 KLNREDLLRKQRTFDNGS1PHQIHLGELHAILRRQEDFYPLKDNRKREKIEKILTFRIPYVGPLARGNSRFAW

-continued

```

MTRKSEETITPWNFEEVVDKGASAQSFIERMNFDFKNLPNEVKVLPKHSLLYEEFTVYNELTKVKYVTEGM
RKP AFLSGEQQKKAIVD LFLKTRNKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYH DLLKI IKDKDF
LDNEENEDILEDIVLTTLFEDREMEERLKTYAHLFDDKVMQLKRRRTGWRGLSRKLINGIRDKQSGK
TILD FLKSDGFANRNF MQLIH DSDSLTFKEDIQKAQVGQGDSLHEHIANLAGSPA KKG I LQT V KV VDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIBEGIKELGSOILKEHPVENTQLQNEKLLYYYLONGR
DMYV DQELDINRLS DYDV DHI VPQ SFLK DSD I DNKV LTRSD KNRG KSDNP SEEV VKKMKN YWR QLLNA
KL ITQRKF DNL TKAER GGL SELD KAGF I K RQ L VETR QI TKHVA QI L D S R MNT K YD E N D K L I R E V K V I TL KSK
LVSDFRKD FQFYKV REIN NYHHAH DAYLN A VVGT ALIK KYPKLES EF VYGDYK VD VRK MIAK SE QEIGK
ATAKYFFYNSNIMFFKTEITLANGEIRKPLIETNGETGEI WWDKGRDFATVRKVLSMPVNIVVKTEVQTG
GFSKESI LPKRNDSK L I ARKKDWPKYVGGFDSP TV ASV L VVAKVEKGKS KKL KSV KELL GITIMERSSPE
KNP ID FLEAKGYKEVKKDL I I KLPK YSLF ELENGRKRMLASAGELQKG NELALPSKYVNF LYLA SHYEKLK
GS P DNEQKQQLFVEQHKHYLDEIIEQISEFSKRV ILADANLDKV L SAYNKHR DKP I R EQAENI I HLF TL TNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLHQ SIT GLYETRIDL S QLGGD SGSSGGG STNLS DII E KETG KQ
LVIQESI LMLP EEEV EIGNK PESD I LVHTAY DESTDEN VMLL TS D APEY KPWAL V IQDSNGEN KI KMLSGG
SGGSSGG STNLS DII E KETG KQ LVIQESI LMLP EEEV EIGNK PESD I LVHTAY DESTDEN VMLL TS D APEY KPWAL V IQDSNGEN KI KMLSGG
WALVIQDSNGEN KI KMLSGGSPKKRK

```

15

Example 2: Anti-Cancer Vaccination Using CRISPR-Cas9 Genome/Base-Editing Technologies

Described herein are new methods to stimulate the immune system to treat tumors and prevent metastatic lesions. By turning the genome and proteome of the malignant cells into a personalized endogenous anti-cancer vaccine *in vivo*.

Provided herein is a new immuno-oncology methodology to raise robust T-cell and B-cell mediated immune responses against tumor-specific proteins, which are otherwise tolerated as “self” by the immune system of a cancer patient¹⁻³ (FIG. 1, Tables 1-3). This methodology is uniquely suited for programmable CRISPR-Cas9 genome- and base-editing tools,⁴⁻¹⁰ exploited to alter the translated sequences of tumor specific genes¹¹ to produce highly immunogenic heteroclitic and cryptic peptide epitopes *in situ* (Tables 5 and 7, FIG. 1). Heteroclitic epitopes are altered versions of endogenous peptide sequences engineered to elicit potent immune reactions through the MHC-I and MHC-II antigen presentation pathways,¹² which also produce cross-reactive responses towards the parent wild-type peptide sequences (FIG. 2A)^{2, 13,14}. For example, the peptide epitope EAAGIGILTV (SEQ ID NO: 388) from the melanocyte differentiation and melanoma marker MART-1²⁶⁻³⁶ is weakly immunogenic, whereas vaccination with a similar peptide engineered with a hydrophobic residue^{15,16} on the MHC-anchor position MART-1 (27L) ELAGIGILTV (SEQ ID NO: 1085) promotes robust T-cell immune responses against melanoma (Table 5).^{17,18} Cryptic epitopes arise from non-translated genomic sequences through processes that are elevated in cancer cells, such as aberrant mRNA splicing, alternative open-reading frames (ORFs), and deglycosidation of proteins.¹⁹ For example, LAGE-1 immunogenic antigens are expressed from ORF-2^{20,21} of the gene NY-ESO-122-24 (Table 7). Introduction of these immunogenic protein sequences using genome/base editing is designed to break “self”-tolerance to cancer-specific antigens,¹⁻³ which is known drive the infiltration of immune cells into the tumor promoting the recognition of malignant cells as foreign through 3,827,234 multiple mechanisms (FIG. 3).^{25,26} Anti-cancer vaccination using genome/base-editing is rendered cancer-specific by targeting genes that are preferentially or exclusively expressed by tumor cells to prevent autoimmunity side effects.^{2,27} Anti-cancer vaccination strategies could be particularly useful for the treatment of melanoma (Table 1),²⁸⁻³⁰ as well as colorectal tumors, stomach cancer, and other highly mutagenized cancers that accumulate non-synonymous hitchhiker mutations with high frequencies (FIG. 4).³¹ In such cases, spreading of the adaptive immune response towards translated “neo-epitopes”^{32,33} that are

unique to the cell lineage facilitate remission³⁴ and prevent metastatic lesions (abscopal effect) 35-39 (FIG. 3). Importantly, the recently FDA-approved checkpoint inhibitors (anti-PD1 and anti-CTLA4 antibodies), which lower the threshold for T-cell stimulation,^{27,40} have shown promise for co-administration with anti-cancer vaccines,⁴¹⁻⁴⁶ and could enhance the clinical effectiveness of immunization against a broader assortment of cancer types.^{47,48}

The heteroclitic and cryptic epitopes programmed by genome/base-editing may be personalized to match each patient’s malignancy and immune system, or alternatively a guide-RNA cocktail can be developed to engage the most frequent HLA allele supertypes that broadly cover the human population (FIG. 2B). Recent advances for the delivery of genome editing tools are enabling for anti-cancer vaccination. These methods include intracellular delivery using electroporation,⁴⁹⁻⁵⁰ viral vectors,^{51,52} cell penetrating peptides,^{53,54} liposomes,^{55,56} polymers,⁵⁷ membrane deformation,⁵⁸ and nanoparticles⁵⁹; and the types of cargo include RNA transcripts, DNA expression vectors, or Cas9 protein-guide RNA complexes purified,⁶⁰ or within cationic lipid vesicles.^{61,62} Therefore, the vaccination treatments could be potentially performed *in vivo* directly on tumor cells,⁶³ the tissues that originated the tumor,⁶⁴ or alternatively *ex vivo* for re-injection of irradiated whole-cell vaccines.⁶⁵⁻⁶⁷

Also provided herein are numerous specific examples of genomic target sites in tumor-associated genes (Tables 1-3), and the guide-RNAs designed to program the alteration of these translated sequences (Tables 5 and 7), in order to replicate or closely mimic known epitopes that have literature or pre-clinical precedent. The genome editing reactions were designed for one of the CRISPR/Cas9 tools: (i) “base editors” that catalyze chemical reactions on nucleobases (e.g. cytidine deaminase-Cas9 fusion); or (ii) engineered nucleases with DNA cutting activity (e.g. WT Cas9,⁵⁻⁷ Cas9 nickases⁸ or Fok1-nuclease-dCas9 fusions^{9,10}). Examples of other potentially useful genome-editing reactions to alter cancer-specific genes to produce heteroclitic/cryptic epitopes are shown in Tables 5 and 7. By extension, Cas9 tools and Homology-Directed Repair (HDR) pathways may also be exploited to introduce heteroclitic epitopes through DNA templates by lowering the rate of indels using several techniques.⁶⁸⁻⁷⁰ Finally, to expand the repertoire of heteroclitic and cryptic epitopes in an unbiased high-throughput manner, the aforementioned tools could be used to screen libraries of guide-RNAs targeting all PAM sites across a tumor-associated genes of interest, which can be replicated using the genome/base-editing reactions shown in Tables 5 and 7.^{71,72}

1. Melero, I. et al. Therapeutic vaccines for cancer: an overview of clinical trials. *Nature reviews. Clinical oncology* 11, 509-524, (2014).
2. Pardoll, D. M. Inducing autoimmune disease to treat cancer. *Proceedings of the National Academy of Sciences of the United States of America* 96, 5340-5342, (1999).
3. Buonaguro, L., Petrizzo, A., Tornesello, M. L. & Buonaguro, F. M. Translating tumor antigens into cancer vaccines. *Clinical and vaccine immunology: CVI* 18, 23-34, (2011).
4. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* advance online publication, (2016).
5. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823, (2013).
6. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821, (2012).
7. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* 339, 823-826, (2013).
8. Ran, F. A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380-1389, (2013).
9. Guilinger, J. P., Thompson, D. B. & Liu, D. R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nature biotechnology* 32, 577-582, (2014).
10. Tsai, S. Q. et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature biotechnology* 32, 569-576, (2014).
11. Vigneron, N., Stroobant, V., Van den Eynde, B. J. & van der Bruggen, P. Database of T cell-defined human tumor antigens: the 2013 update. *Cancer immunity* 13, 15, (2013).
12. Borbulevych, O. Y., Baxter, T. K., Yu, Z., Restifo, N. P. & Baker, B. M. Increased immunogenicity of an anchor-modified tumor-associated antigen is due to the enhanced stability of the peptide/MHC complex: implications for vaccine design. *J Immunol* 174, 4812-4820, (2005).
13. Bakker, A. B. et al. Analogs of CTL epitopes with improved MHC class-I binding capacity elicit anti-melanoma CTL recognizing the wild-type epitope. *International journal of cancer* 70, 302-309, (1997).
14. Purcell, A. W., McCluskey, J. & Rossjohn, J. More than one reason to rethink the use of peptides in vaccine design. *Nat Rev Drug Discov* 6, 404-414, (2007).
15. Chowell, D. et al. TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. *Proceedings of the National Academy of Sciences of the United States of America* 112, E1754-1762, (2015).
16. Ruppert, J. et al. Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell* 74, 929-937, (1993).
17. Madura, F. et al. Structural basis for ineffective T-cell responses to MHC anchor residue-improved "heteroclitic" peptides. *European journal of immunology* 45, 584-591, (2015).
18. Rivoltini, L. et al. A superagonist variant of peptide MART1/Melan A27-35 elicits anti-melanoma CD8+ T cells with enhanced functional characteristics: implication for more effective immunotherapy. *Cancer research* 59, 301-306, (1999).

19. Andersen, R. S. et al. High frequency of T cells specific for cryptic epitopes in melanoma patients. *Oncimmunology* 2, e25374, (2013).
20. Rimoldi, D. et al. Efficient simultaneous presentation of NY-ESO-1/LAGE-1 primary and nonprimary open reading frame-derived CTL epitopes in melanoma. *J Immunol* 165, 7253-7261, (2000).
21. Mandic, M. et al. The alternative open reading frame of LAGE-1 gives rise to multiple promiscuous HLA-DR-restricted epitopes recognized by T-helper 1-type tumor-reactive CD4+ T cells. *Cancer research* 63, 6506-6515, (2003).
22. Campos-Perez, J. et al. DNA fusion vaccine designs to induce tumor-lytic CD8+ T-cell attack via the immunodominant cysteine-containing epitope of NY-ESO 1. *International journal of cancer* 133, 1400-1407, (2013).
23. Webb, A. I. et al. Functional and structural characteristics of NY-ESO-1-related HLA A2-restricted epitopes and the design of a novel immunogenic analogue. *The Journal of biological chemistry* 279, 23438-23446, (2004).
24. Chen, J. L. et al. Identification of NY-ESO-1 peptide analogues capable of improved stimulation of tumor-reactive CTL. *J Immunol* 165, 948-955, (2000).
25. Lally, K. M. et al. Unmasking cryptic epitopes after loss of immunodominant tumor antigen expression through epitope spreading. *International journal of cancer* 93, 841-847, (2001).
26. Fridman, W. H. et al. The ultimate goal of curative anti-cancer therapies: inducing an adaptive anti-tumor immune response. *Frontiers in immunology* 2, 66, (2011).
27. Attia, P. et al. Autoimmunity correlates with tumor regression in patients with metastatic melanoma treated with anti-cytotoxic T-lymphocyte antigen-4. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 23, 6043-6053, (2005).
28. Rosenberg, S. A. et al. Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nature medicine* 4, 321-327, (1998).
29. Slingluff, C. L., Jr. et al. Clinical and immunologic results of a randomized phase II trial of vaccination using four melanoma peptides either administered in granulocyte-macrophage colony-stimulating factor in adjuvant or pulsed on dendritic cells. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 21, 4016-4026, (2003).
30. Slingluff, C. L., Jr. et al. Immunologic and clinical outcomes of vaccination with a multiepitope melanoma peptide vaccine plus low-dose interleukin-2 administered either concurrently or on a delayed schedule. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 22, 4474-4485, (2004).
31. Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214-218, (2013).
32. Duan, F. et al. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anti-cancer immunogenicity. *J Exp Med* 211, 2231-2248, (2014).
33. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* 348, 69-74, (2015).
34. Gubin, M. M. et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* 515, 577-581, (2014).
35. Kreiter, S. et al. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* 520, 692-696, (2015).

36. Chandra, R. A. et al. A systematic evaluation of abscopal responses following radiotherapy in patients with metastatic melanoma treated with ipilimumab. *Oncimmunology* 4, e1046028, (2015).
37. Ma, Y. et al. Chemotherapy and radiotherapy: cryptic anticancer vaccines. *Seminars in immunology* 22, 113-124, (2010).
38. Karbach, J. et al. Long-term complete remission following radiosurgery and immunotherapy in a melanoma patient with brain metastasis: immunologic correlates. *Cancer immunology research* 2, 404-409, (2014).
39. Demaria, S. et al. Ionizing radiation inhibition of distant untreated tumors (abscopal effect) is immune mediated. *International journal of radiation oncology, biology, physics* 58, 862-870, (2004).
40. Wong, R. M. et al. Programmed death-1 blockade enhances expansion and functional capacity of human melanoma antigen-specific CTLs. *International immunology* 19, 1223-1234, (2007).
41. Fu, J. et al. Preclinical evidence that PD1 blockade cooperates with cancer vaccine TEGVAX to elicit regression of established tumors. *Cancer research* 74, 4042-4052, (2014).
42. Gibney, G. T. et al. Safety, correlative markers, and clinical results of adjuvant nivolumab in combination with vaccine in resected high-risk metastatic melanoma. *Clinical cancer research: an official journal of the American Association for Cancer Research* 21, 712-720, (2015).
43. Soares, K. C. et al. PD-1/PD-L1 blockade together with vaccine therapy facilitates effector T-cell infiltration into pancreatic tumors. *J Immunother* 38, 1-11, (2015).
44. Le, D. T. et al. Evaluation of ipilimumab in combination with allogeneic pancreatic tumor cells transfected with a GM-CSF gene in previously treated pancreatic cancer. *J Immunother* 36, 382-389, (2013).
45. Sierro, S. R. et al. Combination of lentivector immunization and low-dose chemotherapy or PD-1/PD-L1 blocking primes self-reactive T cells and induces anti-tumor immunity. *European journal of immunology* 41, 2217-2228, (2011).
46. Hodi, F. S. et al. Biologic activity of cytotoxic T lymphocyte-associated antigen 4 antibody blockade in previously vaccinated metastatic melanoma and ovarian carcinoma patients. *Proceedings of the National Academy of Sciences of the United States of America* 100, 4712-4717, (2003).
47. Morse, M. A. & Lyerly, H. K. Checkpoint blockade in combination with cancer vaccines. *Vaccine* 33, 7377-7385, (2015).
48. Vanneman, M. & Dranoff, G. Combining immunotherapy and targeted therapies in cancer treatment. *Nature reviews. Cancer* 12, 237-251, (2012).
49. Bakondi, B. et al. In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa. *Molecular therapy: the journal of the American Society of Gene Therapy* 24, 556-563, (2016).
50. Chen, S., Lee, B., Lee, A. Y., Modzelewski, A. J. & He, L. Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. *The Journal of biological chemistry*, (2016).
51. Maggio, I. et al. Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. *Scientific reports* 4, 5105, (2014).

52. Wang, W. et al. CCR5 gene disruption via lentiviral vectors expressing Cas9 and single guided RNA renders cells resistant to HIV-1 infection. *PloS one* 9, e115987, (2014).
53. Ramakrishna, S. et al. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome research* 24, 1020-1027, (2014).
54. Liu, J., Gaj, T., Patterson, J. T., Sirk, S. J. & Barbas, C. F., 3rd. Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. *PloS one* 9, e85755, (2014).
55. Ye, L. et al. Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection. *Proceedings of the National Academy of Sciences of the United States of America* 111, 9591-9596, (2014).
56. Allen, T. M. & Cullis, P. R. Liposomal drug delivery systems: from concept to clinical applications. *Advanced drug delivery reviews* 65, 36-48, (2013).
57. *Efficient Delivery of Sigma CRISPRs via a Non-Liposomal Polymeric Transfection Reagent, TransIT®-CRISPR*, <www.sigmaldrich.com/technical-documents/articles/biology/transit-crispr-transfection-reagent.html> (2016).
58. Han, X. et al. CRISPR-Cas9 delivery to hard-to-transfect cells via membrane deformation. *Science advances* 1, e1500454, (2015).
59. Cox, D. B., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nature medicine* 21, 121-131, (2015).
60. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J. S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome research* 24, 1012-1019, (2014).
61. Zuris, J. A. et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nature biotechnology* 33, 73-80, (2015).
62. Wang, M. et al. Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. *Proceedings of the National Academy of Sciences of the United States of America*, (2016).
63. Villarreal, D. O. et al. Alarmin IL-33 acts as an immunoadjuvant to enhance antigen-specific tumor immunity. *Cancer research* 74, 1789-1800, (2014).
64. Ginsberg, B. A. et al. Immunologic response to xenogeneic gp100 DNA in melanoma patients: comparison of particle-mediated epidermal delivery with intramuscular injection. *Clinical cancer research: an official journal of the American Association for Cancer Research* 16, 4057-4065, (2010).
65. Geary, S. M., Lemke, C. D., Lubaroff, D. M. & Salem, A. K. Proposed mechanisms of action for prostate cancer vaccines. *Nature reviews. Urology* 10, 149-160, (2013).
66. Dranoff, G. et al. Vaccination with irradiated tumor cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity. *Proceedings of the National Academy of Sciences of the United States of America* 90, 3539-3543, (1993).
67. Chiang, C. L., Benencia, F. & Coukos, G. Whole tumor antigen vaccines. *Seminars in immunology* 22, 132-143, (2010).
68. Chu, V. T. et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature biotechnology* 33, 543-548, (2015).

69. Yu, C. et al. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell stem cell* 16, 142-147, (2015).
70. Paquet, D. et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125-129, (2016).
71. Chen, S. et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* 160, 1246-1260, (2015).
72. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nature reviews. Genetics* 16, 299-311, (2015).
73. Overwijk, W. W. & Restifo, N. P. B16 as a mouse model for human melanoma. *Current protocols in immunology* / edited by John E. Coligan . . . [et al.] Chapter 20, Unit 20 21, (2001).
74. Saenger, Y. M. et al. Improved tumor immunity using anti-tyrosinase related protein-1 monoclonal antibody combined with DNA vaccines in murine melanoma. *Cancer research* 68, 9884-9891, (2008).
75. Yu, Z. et al. Poor immunogenicity of a self/tumor antigen derives from peptide-MHC-I instability and is independent of tolerance. *The Journal of clinical investigation* 114, 551-559, (2004).
76. Robbins, P. F. et al. The intronic region of an incompletely spliced gp100 gene transcript encodes an epitope recognized by melanoma-reactive tumor-infiltrating lymphocytes. *J Immunol* 159, 303-308, (1997).
77. Pinilla, C. et al. Combinatorial peptide libraries as an alternative approach to the identification of ligands for tumor-reactive cytolytic T lymphocytes. *Cancer research* 61, 5153-5160, (2001).
78. Wang, R. F., Parkhurst, M. R., Kawakami, Y., Robbins, P. F. & Rosenberg, S. A. Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen. *J Exp Med* 183, 1131-1140, (1996).
79. Lupetti, R. et al. Translation of a retained intron in tyrosinase-related protein (TRP) 2 mRNA generates a new cytotoxic T lymphocyte (CTL)-defined and shared human melanoma antigen not expressed in normal cells of the melanocytic lineage. *J Exp Med* 188, 1005-1016, (1998).
80. Visseren, M. J. et al. Affinity, specificity and T-cell-receptor diversity of melanoma-specific CTL generated in vitro against a single tyrosinase epitope. *International journal of cancer* 72, 1122-1128, (1997).
81. Valmori, D. et al. Analysis of the cytolytic T lymphocyte response of melanoma patients to the naturally HLA-A*0201-associated tyrosinase peptide 368-376. *Cancer research* 59, 4050-4055, (1999).
82. Rubio-Godoy, V. et al. Toward synthetic combinatorial peptide libraries in positional scanning format (PS-SCL)-based identification of CD8+ Tumor-reactive T-Cell Ligands: a comparative analysis of PS-SCL recognition by a single tumor-reactive CD8+ cytolytic T-lymphocyte clone. *Cancer research* 62, 2058-2063, (2002).
83. Skipper, J. C. et al. An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. *J Exp Med* 183, 527-534, (1996).

84. Bernatchez, C. et al. Altered decamer and nonamer from an HLA-A0201-restricted epitope of Survivin differentially stimulate T-cell responses in different individuals. *Vaccine* 29, 3021-3030, (2011).
85. Hirohashi, Y. et al. An HLA-A24-restricted cytotoxic T lymphocyte epitope of a tumor-associated protein, survivin. *Clinical cancer research: an official journal of the American Association for Cancer Research* 8, 1731-1739, (2002).
86. Gross, D. A. et al. High vaccination efficiency of low-affinity epitopes in antitumor immunotherapy. *The Journal of clinical investigation* 113, 425-433, (2004).
87. Scardino, A. et al. HER-2/neu and hTERT cryptic epitopes as novel targets for broad spectrum tumor immunotherapy. *J Immunol* 168, 5900-5906, (2002).
88. Aurisicchio, L. et al. A novel minigene scaffold for therapeutic cancer vaccines. *Oncoimmunology* 3, e27529, (2014).
89. Bae, J., Martinson, J. A. & Klingemann, H. G. Identification of novel CD33 antigen-specific peptides for the generation of cytotoxic T lymphocytes against acute myeloid leukemia. *Cellular immunology* 227, 38-50, (2004).
90. Bae, J., Martinson, J. A. & Klingemann, H. G. Heteroclitic CD33 peptide with enhanced anti-acute myeloid leukemic immunogenicity. *Clinical cancer research: an official journal of the American Association for Cancer Research* 10, 7043-7052, (2004).
91. Smith, H. A., Rekoske, B. T. & McNeel, D. G. DNA vaccines encoding altered peptide ligands for SSX2 enhance epitope-specific CD8+ T-cell immune responses. *Vaccine* 32, 1707-1715, (2014).
92. Oka, Y., Tsuboi, A., Oji, Y., Kawase, I. & Sugiyama, H. WT1 peptide vaccine for the treatment of cancer. *Current opinion in immunology* 20, 211-220, (2008).
93. Pinilla-Ibarz, J. et al. Improved human T-cell responses against synthetic HLA-0201 analog peptides derived from the WT1 oncoprotein. *Leukemia* 20, 2025-2033, (2006).
94. Di Stasi, A., Jimenez, A. M., Minagawa, K., Al-Obaidi, M. & Rezvani, K. Review of the Results of WT1 Peptide Vaccination Strategies for Myelodysplastic Syndromes and Acute Myeloid Leukemia from Nine Different Studies. *Frontiers in immunology* 6, 36, (2015).
95. Tsuboi, A. et al. Enhanced induction of human WT1-specific cytotoxic T lymphocytes with a 9-mer WT1 peptide modified at HLA-A*2402-binding residues. *Cancer immunology, immunotherapy: CII* 51, 614-620, (2002).
96. Trojan, A. et al. Generation of cytotoxic T lymphocytes against native and altered peptides of human leukocyte antigen-A*0201 restricted epitopes from the human epithelial cell adhesion molecule. *Cancer research* 61, 4761-4765, (2001).
97. Volpe, G. et al. Alternative BCR/ABL splice variants in Philadelphia chromosome-positive leukemias result in novel tumor-specific fusion proteins that may represent potential targets for immunotherapy approaches. *Cancer research* 67, 5300-5307, (2007).
98. Casnici, C. et al. Out of frame peptides from BCR/ABL alternative splicing are immunogenic in HLA A2.1 transgenic mice. *Cancer letters* 276, 61-67, (2009).

189

99. Casnici, C. et al. Immunologic evaluation of peptides derived from BCR/ABL-out-of-frame fusion protein in HLA A2.1 transgenic mice. *J Immunother* 35, 321-328, (2012).
100. Tangri, S. et al. Structural features of peptide analogs of human histocompatibility leukocyte antigen class I epitopes that are more potent and immunogenic than wild-type peptide. *J Exp Med* 194, 833-846, (2001).
101. Barve, M. et al. Induction of immune responses and clinical efficacy in a phase II trial of IDM-210I, a 10-epitope cytotoxic T-lymphocyte vaccine, in metastatic non-small-cell lung cancer. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 26, 4418-4425, (2008).
102. Graff-Dubois, S. et al. Generation of CTL recognizing an HLA-A*0201-restricted epitope shared by MAGE-A1, -A2, -A3, -A4, -A6, -A10, and -A12 tumor antigens: implication in a broad-spectrum tumor immunotherapy. *J Immunol* 169, 575-580, (2002).
103. Tourdot, S. et al. A general strategy to enhance immunogenicity of low-affinity HLA-A2. 1-associated peptides: implication in the identification of cryptic tumor epitopes. *European journal of immunology* 30, 3411-3421, (2000).
104. Tsang, K. Y., Palena, C., Gulley, J., Arlen, P. & Schlom, J. A human cytotoxic T-lymphocyte epitope and its agonist epitope from the nonvariable number of tandem repeat sequence of MUC-1. *Clinical cancer research: an official journal of the American Association for Cancer Research* 10, 2139-2149, (2004).
105. Geynisman, D. M. et al. A randomized pilot phase I study of modified carcinoembryonic antigen (CEA) peptide (CAP1-6D)/montanide/GM-CSF-vaccine in patients with pancreatic adenocarcinoma. *Journal for immunotherapy of cancer* 1, 8, (2013).
106. Kleinstiver, B. P. et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481-485, (2015).
107. Guevara-Patino, J. A. et al. Optimization of a self antigen for presentation of multiple epitopes in cancer immunity. *The Journal of clinical investigation* 116, 1382-1390, (2006).
108. Duan, F. et al. Immune rejection of mouse tumors expressing mutated self. *Cancer research* 69, 3545-3553, (2009).
109. Myers, C. E. et al. Variation in cytotoxic T-lymphocyte responses to peptides derived from tyrosinase-related protein-2. *Human immunology* 69, 24-31, (2008).
110. Castle, J. C. et al. Exploiting the mutanome for tumor vaccination. *Cancer research* 72, 1081-1091, (2012).
111. Cho, H. I. & Celis, E. Optimized peptide vaccines eliciting extensive CD8 T-cell responses with therapeutic antitumor effects. *Cancer research* 69, 9012-9019, (2009).

EQUIVALENTS AND SCOPE

In the claims articles such as "a," "an," and "the" may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The disclosure includes embodiments in which

190

exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The disclosure includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.

Furthermore, the disclosure encompasses all variations, combinations, and permutations in which one or more limitations, elements, clauses, and descriptive terms from one or more of the listed claims is introduced into another claim. For example, any claim that is dependent on another claim can be modified to include one or more limitations found in any other claim that is dependent on the same base claim. Where elements are presented as lists, e.g., in Markush group format, each subgroup of the elements is also disclosed, and any element(s) can be removed from the group. It should be understood that, in general, where the disclosure, or aspects of the disclosure, is/are referred to as comprising particular elements and/or features, certain embodiments of the disclosure or aspects of the disclosure consist, or consist essentially of, such elements and/or features. For purposes of simplicity, those embodiments have not been specifically set forth in haec verba herein.

It is also noted that the terms "comprising" and "containing" are intended to be open and permits the inclusion of additional elements or steps. Where ranges are given, endpoints are included. Furthermore, unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or sub-range within the stated ranges in different embodiments of the disclosure, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

This application refers to various issued patents, published patent applications, journal articles, and other publications, all of which are incorporated herein by reference. If there is a conflict between any of the incorporated references and the instant specification, the specification shall control. In addition, any particular embodiment of the present disclosure that falls within the prior art may be explicitly excluded from any one or more of the claims. Because such embodiments are deemed to be known to one of ordinary skill in the art, they may be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the disclosure can be excluded from any claim, for any reason, whether or not related to the existence of prior art.

Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation many equivalents to the specific embodiments described herein. The scope of the present embodiments described herein is not intended to be limited to the above Description, but rather is as set forth in the appended claims. Those of ordinary skill in the art will appreciate that various changes and modifications to this description may be made without departing from the spirit or scope of the present disclosure, as defined in the following claims.

SEQUENCE LISTING

The patent contains a lengthy sequence listing. A copy of the sequence listing is available in electronic form from the USPTO web site (<https://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US12390514B2>). An electronic copy of the sequence listing will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

What is claimed is:

1. A method for producing a heteroclitic epitope in a subject in need thereof, the method comprising administering to the subject a therapeutically effective amount of a composition comprising:
 - (i) a fusion protein comprising (a) a guide nucleotide sequence-programmable DNA-binding protein domain, and (b) a cytosine deaminase domain; and
 - (ii) a guide nucleotide sequence, wherein the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 724-751, 870-877, 888-905, and 907-985;

wherein the guide nucleotide sequence of (ii) targets the fusion protein of (i) to a polynucleotide encoding a tumor-specific antigen in a tumor cell; and

wherein the fusion protein changes a target cytosine (C) base to a thymine (T) base via deamination.
2. The method of claim 1, wherein the polynucleotide encoding the tumor-specific antigen is located in the genome of the tumor cell.
3. The method of claim 1, wherein the guide nucleotide sequence-programmable DNA-binding protein domain is selected from the group consisting of nuclease inactive Cas9 (dCas9) domains, Cas9 nickase (nCas9) domains, nuclease inactive Cpf1 domains, and nuclease inactive Argonaute domains.
4. The method of claim 3, wherein the guide nucleotide sequence-programmable DNA-binding protein domain is a nuclease inactive Cas9 (dCas9) domain or a Cas9 nickase (nCas9) domain.
5. The method of claim 1, wherein the cytosine deaminase domain is selected from the group consisting of APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, APOBEC4, and activation-induced deaminase (AID).
6. The method of claim 1, wherein the fusion protein of (i) further comprises a uracil glycosylase inhibitor (UGI) domain.
7. The method of claim 6, wherein the fusion protein comprises the structure: NH₂-[cytosine deaminase domain]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA-binding protein domain]-[optional linker sequence]-[UGI domain]-COOH; NH₂-[UGI domain]-[optional linker sequence]-[cytosine deaminase domain]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA-binding protein domain]-COOH; or NH₂-[cytosine deaminase domain]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA-binding protein domain]-COOH; or NH₂-[cytosine deaminase domain]-[optional linker sequence]-[guide nucleotide sequence-programmable DNA-binding protein domain]-COOH.
8. The method of claim 7, wherein the optional linker sequence comprises (GGGS)_n, (SEQ ID NO: 337), (GGGGS)_n (SEQ ID NO: 308), (G)_n (SEQ ID NO: 783), (EAAAK)_n (SEQ ID NO: 309), (GGS)_n (SEQ ID NO: 784), SGSETPGTSESATPES (SEQ ID NO: 310), (XP)_n (SEQ ID NO: 785), or a combination of any of these, wherein n is independently an integer between 1 and 30, and wherein X is any amino acid.
9. The method of claim 1, wherein the fusion protein comprises the amino acid sequence of any one of SEQ ID NOS: 293-302, 1071, and 1084.
10. The method of claim 1, wherein the tumor-specific antigen is selected from the group consisting of: gp100; MART-1; hTERT; TyRP1; HER2; CEA-CAM; tyrosinase (TYR); CD33; MAGE-A3; MAGE-A4; NY-ESO-1; SSX-2; survivin; EpCAM; and MUC1.
11. The method of claim 1, wherein the target C base is in a target codon in a coding region of the polynucleotide encoding the tumor-specific antigen, and wherein the target codon is any one of the following target codons: CTT (Leu/L), CTC (Leu/L), ATG (Met/M), GTT (Val/V), GTA (Val/V), GTC (Val/V), GTG (Val/V), TCT (Ser/S), TCC (Ser/S), TCA (Ser/S), TCG (Ser/S), AGT (Ser/S), AGC (Ser/S), CCT (Pro/P), CCC (Pro/P), CCA (Pro/P), CCG (Pro/P), ACT (Thr/T), ACC (Thr/T), ACA (Thr/T), ACG (Thr/T), GCT (Ala/A), GCC (Ala/A), GCA (Ala/A), GCG (Ala/A), CAT (His/H), CAC (His/H), GAT (Asp/D), GAC (Asp/D), GAA (Glu/E), GAG (Glu/E), TGT (Cys/C), TGC (Cys/C), CGT (Arg/R), CGC (Arg/R), AGA (Arg/R), AGG (Arg/R), CGG (Arg/R), GGT (Gly/G), GGC (Gly/G), GGA (Gly/G), GGG (Gly/G), CAG (Gln/Q), TGG (Trp/W), CGA (Arg/R), CAA (Gln/Q), TGG (Trp/W), and CGA (Arg/R).
12. The method of claim 11, wherein the target codon is converted to a modified codon selected from any one of the following modified codons: ATA (Ile/I), ATT (Ile/I), ATC (Ile/I), ATG (Met/M), TTT (Phe/F), TTC (Phe/F), TTA (Leu/L), TTG (Leu/L), AAT (Asp/N), AAC (Asp/N), TCT (Ser/S), TCC (Ser/S), TCA (Ser/S), TCG (Ser/S), CTT (Leu/L), CTC (Leu/L), CTA (Leu/L), CTG (Leu/L), GTT (Val/V), GTC (Val/V), GTA (Val/V), GTG (Val/V), ACT (Thr/T), ACC (Thr/T), ACA (Thr/T), ACG (Thr/T), TAT (Tyr/Y), TAC (Tyr/Y), AAA (Lys/K), AAG (Lys/K), TGT (Cys/C), TGC (Cys/C), CAG (Gln/Q), TGG (Trp/W), GAT (Asp/D), GAC (Asp/D), GAA (Glu/E), GAG (Glu/E), AGT (Ser/S), AGC (Ser/S), AGA (Arg/R), AGG (Arg/R), TAG (amber), TGA (opal), and TAA (ochre).
13. The method of claim 1, wherein the target C base is located in a non-coding region of the polynucleotide encoding the tumor-specific antigen.
14. The method of claim 1, wherein the heteroclitic epitope is at least 5-fold more immunogenic than a native epitope from the tumor-specific antigen.
15. The method of claim 1, wherein the heteroclitic epitope is displayed on the surface of the tumor cell via the MHC class I antigen presentation pathway.
16. The method of claim 1, wherein the method is carried out in vivo.

193

17. The method of claim **1**, wherein the tumor-specific antigen is gp100.

18. The method of claim **17**, wherein the deamination of the target C base:

- (a) in codon T210 of gp100 results in a T210I mutation (SEQ ID NO: 786);
- (b) in codon A288 of gp100 results in a A288V mutation (SEQ ID NO: 818); or
- (c) in codon T155 of gp100 results in a T155I mutation (SEQ ID NO: 787).

19. The method of claim **18**, wherein;

- (a) a heteroclitic epitope comprising the amino acid sequence of IIDQVPFSV (SEQ ID NO: 786) is generated, and wherein the I at position 2 corresponds to the T210I mutation;
- (b) a heteroclitic epitope comprising the amino acid sequence of YLEPGPVTY (SEQ ID NO: 818) is generated, and wherein the V at position 7 corresponds to the A288V mutation; or

194

(c) a heteroclitic epitope comprising the amino acid sequence of KIWGQYWQV (SEQ ID NO: 787) is generated, and wherein the I at position 2 corresponds to the T155I mutation.

5 20. The method of claim **19**, wherein the guide nucleotide sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 724, 725, 870-877, 888, and 889.

10 21. The method of claim **1**, the method further comprising administering to the subject a therapeutically effective amount of an immune checkpoint inhibitor.

22. The method of claim **1**, wherein the target C base is in a target codon in a coding region of the polynucleotide encoding the tumor-specific antigen.

15 23. The method of claim **1**, wherein the heteroclitic epitope is displayed on the surface of an antigen presenting cell (APC) via the MHC class II antigen presentation pathway.

* * * * *