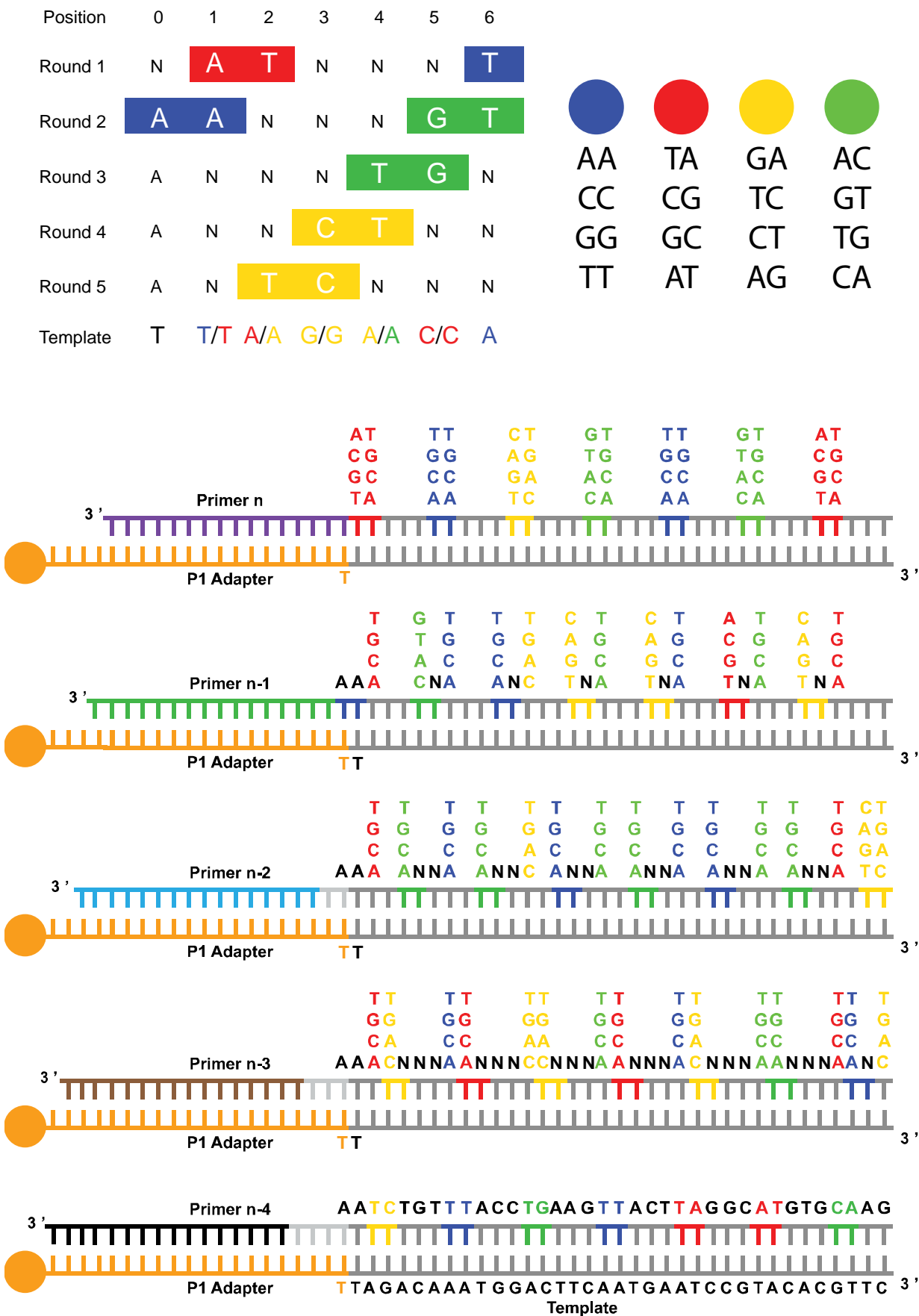


Supplemental Data Figure 1.



Supplemental Figure 1. SOLiD Color Space Coding and Sequence Deconvolution. Top. Left) Interrogation of a template with sequence TAGACA (positions 1-6 with adapter sequence T at position 0). Right) Color space coding is reshown as reference. Bottom. Sequencing by ligation and deconvolution of template sequence. Extension of primer (n) interrogates template positions 1 and 2, 6 and 7, 11 and 12, and so forth. No sequence information can be determined during the first round since each fluor represents four two base possible combinations. The extension of primer (n-1) allows for the deconvolution of the base at position 1 with the knowledge of the P1 adapter base "T" at position 0. The extension of primers (n-2) and (n-3) provide additional positional color space information prior to deconvolution of the template performed after round 5 (primer n-4). The "N" bases show the positions along the template that have been sequenced twice prior to deconvolution.

Supplementary Data Table 1 Next Generation Sequencing Software Summary

Program Name	Alignment to Reference	De Novo Assembly	SNP Indel Discovery	ChIP/ Bisulphite -Seq	RNA-Seq	Alternate Base Caller	Alignment Viewer	Statistical Analysis	Primary Operating Environments	References	Authors	Web Site
ALLPATHS		X								(1)	Jonathan Butler et al.	
Alta-cyclic						X			Linux, SGE cluster environment	(2)	Yaniv Erlich et al.	hannonlab.cshl.edu/Alta-Cyclic/main.html
Bowtie	X								Linux, Windows, Mac		Ben Langmead and Cole Trapnell	bowtie-bio.sourceforge.net
BS-Seq				X					POSIX OS	(3)	S.J. Cokus et al	epigenomics.mcdb.ucla.edu/BS-Seq/download.html
ChIPSeq				X					Linux, Mac	(4)	D.S. Johnson et al.	woldlab.caltech.edu/html/chipseq_peak_finder
CLCbio Genomics Workbench	X	X	X				X		Windows		CLCbio	www.clcbio.com
EagleView							X		Linux, Windows, Mac	(5)	W. Huang, G. Marth	bioinformatics.bc.edu/marthlab/EagleView
Edena		X							Linux, Windows	(6)		www.genomic.ch/edena.php
ELAND	X								Linux		Anthony J. Cox	bioinfo.cgrb.oregonstate.edu/docs/solexa
Erangle				X	X				Linux, Mac	(7)	Ali Mortazavi et al.	woldlab.caltech.edu/rnaseq
Euler-SR		X								(8, 9)	Mark J. Chaisson and Pavel A. Pevzner	
Exonerate	X								Linux	(10)	Guy St, C Slater and Ewan Birney	www.ebi.ac.uk/~guy/exonerate
FindPeaks				X					Java	(11)	Anthony Fejes	vancouvershortr.sourceforge.net
GMAP	X								Unix	(12)	Thomas Wu and Colin Watanabe	
JMP Genomics							X	X	Linux, Win, Mac		SAS Institue Inc.	www.jmp.com/software/genomics
MAQ	X								Linux, Mac	(13)	Heng Li	maq.sourceforge.net/index.shtml
MIRA2	X	X							Linux	(14)	Bastien Chevreux	chevreux.org/projects_mira.html
MOSAIK	X										Michael Strömberg and Gabor Marth	
MUMmer	X								POSIX OS	(15)	Stefan Kurtz, Adam Phillippy, Arthur L Delcher, Michael Smoot, Martin Shumway, Corina Antonescu and Steven L. Salzberg	http://mummer.sourceforge.net
NextGENe	X	X	X				X		Windows		SoftGenetics	www.softgenetics.com
NOVOALIGN	X								Linux, Mac OS X		Novocraft Technologies	www.novocraft.com/products.html
PBShort			X						Linux		Gabor Marth	bioinformatics.bc.edu/marthlab/PbShort
PyroBayes			X						Linux	(16)	A.R. Quinlan, D.A. Stewart, M.P. Strömberg, G.T. Marth	http://bioinformatics.bc.edu/marthlab/PyroBayes

Program Name	Alignment to Reference	De Novo Assembly	SNP Indel Discovery	ChIP/Bisulphite -Seq	RNA-Seq	Alternate Base Caller	Alignment Viewer	Statistical Analysis	Primary Operating Environments	References	Authors	Web Site
QuEST				X					Linux	(17)	Anton Valouev	mendel.stanford.edu/sidowlab/downloads/quest
RMAP	X								POSIX OS	(18)	Andrew D. Smith and Zhenyu Xuan	http://rulai.cshl.edu/rmap
Roche 454 analysis tools	X	X	X				X		Linux		Roche	www.454.com/products-solutions/analysis-tools/index.asp
Rolexa						X			R	(19)	Jacques Rougemont	svitsrv25.epfl.ch/R-doc/library/Rolexa/html/00Index.html
SAM							X		Linux		Rene Warren, Yaron Butterfield, Asim Siddiqui and Steven Jones	www.bcgsc.ca/platform/bioinfo/software/sam
SeqMan Ngen with SeqMan Pro	X	X	X				X		Windows	(20)	DNASStar	www.dnastar.com
SeqMap	X								Linux, Windows, Mac	(21)	H. Jiang and W.H. Wong	biogibbs.stanford.edu/~jiangh/SeqMap
SHARCGS		X							Perl	(22)	Juliane C. Dohm, Claudio Lottaz, and Heinz Himmelbauer	sharcgs.molgen.mpg.de
SHRAP		X								(23)	Andreas Sundquist et al.	
SHRiMP	X								Linux, Mac		Michael Brudno and Stephen Rumble	http://compbio.cs.toronto.edu/shrimp
SISSRs				X					Perl		Raja Jothi	dir.nhlbi.nih.gov/papers/lmi/epigenomes/sissrs
Slider	X								Java	(24)	N. Malhis, Y. Butterfield, M. Ester, S.J. Jones	http://www.bcgsc.ca/platform/bioinfo/software/slider
SOAPaligner/soap2	X								Linux	(25)	T.W. Lam, Alan Tam, Simon Wong, Edward Wu and S.M. Yiu, Ruiqiang Li	soap.genomics.org.cn/soapaligner.html
SOAPdenovo		X							Linux		T.W. Lam, Alan Tam, Simon Wong, Edward Wu and S.M. Yiu, Ruiqiang Li	soap.genomics.org.cn/#
SSAHA2	X								Linux	(26)	Zemin Ning, Anthony Cox and James Mullikin	www.sanger.ac.uk/Software/analysis/SSAHA
ssahaSNP			X						Linux, Solaris, Mac		Adam Spargo and Zemin Ning	www.sanger.ac.uk/Software/analysis/ssahaSNP
SSAKE		X							Linux	(27)	René Warren, Granger Sutton, Steven Jones and Robert Holt	http://www.bcgsc.ca/platform/bioinfo/software/ssake
SXOligoSearch	X								Web Based		Synamatix	http://synasite.mgrc.com.my:8080/sxog/NewSXOligoSearch.php
VCAKE		X							Perl	(28)	W. Jeck	http://sourceforge.net/projects/vcake
Velvet		X							Linux, Mac	(29)	Daniel Zerbino and Ewan Birney	http://www.ebi.ac.uk/~zerbino/velvet/
XMatchView							X		Linux, Windows		Rene Warren and Steven Jones	http://www.bcgsc.ca/platform/bioinfo/software/xmatchview

References

1. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, et al. ALLPATHS: de novo assembly of whole-genome shotgun microreads. *Genome Res* 2008;18:810-20.
2. Erlich Y, Mitra PP, delaBastide M, McCombie WR, Hannon GJ. Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nature methods* 2008;5:679-82.
3. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 2008;452:215-9.
4. Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-wide mapping of in vivo protein-DNA interactions. *Science* 2007;316:1497-502.
5. Huang W, Marth G. EagleView: A genome assembly viewer for next-generation sequencing technologies. *Genome Res* 2008;18:1538-43.
6. Hernandez D, Francois P, Farinelli L, Osteras M, Schrenzel J. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res* 2008;18:802-9.
7. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* 2008;5:621-8.
8. Chaisson MJ, Brinza D, Pevzner PA. De novo fragment assembly with short mate-paired reads: Does the read length matter? *Genome Res* 2008.
9. Chaisson MJ, Pevzner PA. Short read fragment assembly of bacterial genomes. *Genome Res* 2008;18:324-30.
10. Slater GS, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC bioinformatics* 2005;6:31.
11. Fejes AP, Robertson G, Bilenky M, Varhol R, Bainbridge M, Jones SJ. FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. *Bioinformatics (Oxford, England)* 2008;24:1729-30.
12. Wu TD, Watanabe CK. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics (Oxford, England)* 2005;21:1859-75.
13. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 2008.
14. Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Muller WE, Wetter T, Suhai S. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* 2004;14:1147-59.
15. Delcher AL, Salzberg SL, Phillippy AM. Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics* 2003;Chapter 10:Unit 10 3.
16. Quinlan AR, Stewart DA, Stromberg MP, Marth GT. Pyrobayes: an improved base caller for SNP discovery in pyrosequences. *Nature methods* 2008;5:179-81.
17. Valouev A, Johnson DS, Sundquist A, Medina C, Anton E, Batzoglu S, et al. Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nature methods* 2008.

18. Smith AD, Xuan Z, Zhang MQ. Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC bioinformatics* 2008;9:128.
19. Rougemont J, Amzallag A, Iseli C, Farinelli L, Xenarios I, Naef F. Probabilistic base calling of Solexa sequencing data. *BMC bioinformatics* 2008;9:431.
20. Durfee T, Nelson R, Baldwin S, Plunkett G, 3rd, Burland V, Mau B, et al. The complete genome sequence of *Escherichia coli* DH10B: insights into the biology of a laboratory workhorse. *J Bacteriol* 2008;190:2597-606.
21. Jiang H, Wong WH. SeqMap : mapping massive amount of oligonucleotides to the genome. *Bioinformatics* (Oxford, England) 2008.
22. Dohm JC, Lottaz C, Borodina T, Himmelbauer H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic acids research* 2008;36:e105.
23. Sundquist A, Ronaghi M, Tang H, Pevzner P, Batzoglou S. Whole-genome sequencing and assembly with high-throughput, short-read technologies. *PLoS ONE* 2007;2:e484.
24. Malhis N, Butterfield YS, Ester M, Jones SJ. Slider--maximum use of probability information for alignment of short sequence reads and SNP detection. *Bioinformatics* (Oxford, England) 2009;25:6-13.
25. Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. *Bioinformatics* (Oxford, England) 2008;24:713-4.
26. Ning Z, Cox AJ, Mullikin JC. SSAHA: a fast search method for large DNA databases. *Genome Res* 2001;11:1725-9.
27. Warren RL, Sutton GG, Jones SJ, Holt RA. Assembling millions of short DNA sequences using SSAKE. *Bioinformatics* (Oxford, England) 2007;23:500-1.
28. Jeck WR, Reinhardt JA, Baltrus DA, Hickenbotham MT, Magrini V, Mardis ER, et al. Extending assembly of short DNA sequences to handle error. *Bioinformatics* (Oxford, England) 2007;23:2942-4.
29. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008;18:821-9.