

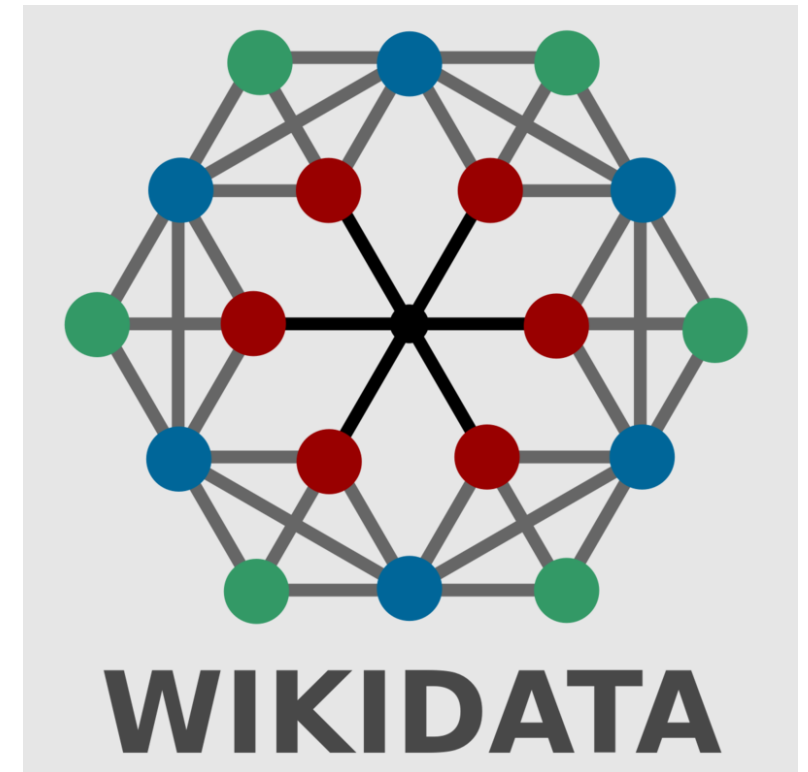
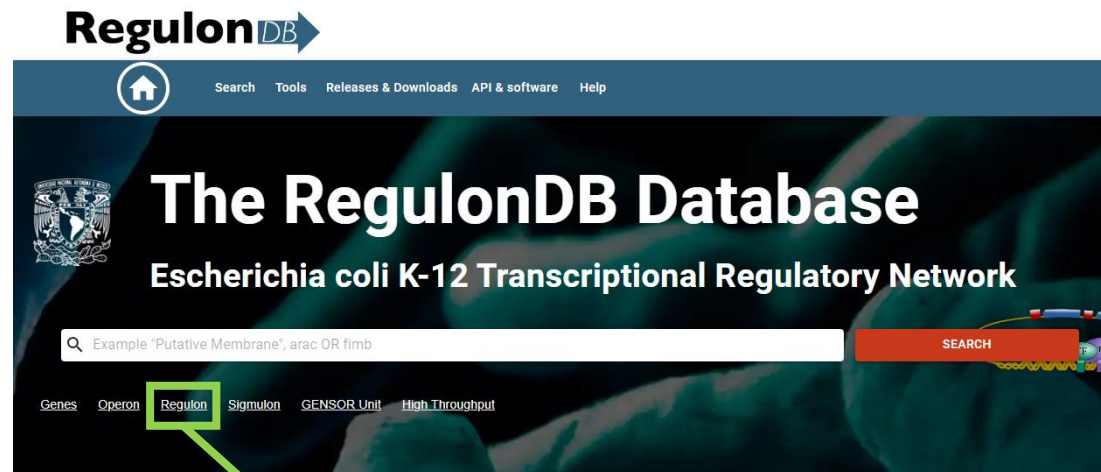
Molekulare Daten von Bakterien in einen Knowledge Graph überführen

Gruppe 4

Kerem & Tim

31.07.2024

Ziel



Regulprotein

Regulon

DicF	repressor	ftsZ
DicF	repressor	xyIR
DicF	repressor	pykA
DicF	repressor	manX

Gene

Schritt 1 – Daten sammeln

1. Regulon Kennzeichnung

RegulonDB

Search Tools Releases & Downloads API & software

Regulons

Regulon	Regulated			
Name	Genes	Operons	Transcription Unit	
Filter... (287)	Min (1) Max (57)	Min Max (24)	Min Max (31)	
	308	141	169	
DicF	4	0	0	
DsrA	10	4	4	
GcvB	60	16	17	
McaS	6	2	3	

DicF, DsrA, GcvB, McaS, ...

2. Link zum DB Eintrag

regulondb.ccg.unam.mx/regulon/RDBECOLIPDC00328

RegulonDB

Search Tools Releases & Downloads API & software Help

Qin prophage; small regulatory RNA DicF

Synonyms: DicF
Type: RNA-Mediated-Translation-Regulation

Regulon DicF

Encoded By: Gene: [dicF](#)

Note: The DicF small regulatory RNA acts to inhibit cell division [Bouche F., et al. 1989⁶](#). DicF causes inhibition of the constriction and the separation of the replicated chromosomes into the daughter nucleoids [Tatart F., et al. 1992¹](#).

DicF interferes with translation of *ftsZ* [Tatart F., et al. 1992¹](#), by base-pairing with the *ftsZ* mRNA [Balasubramanian D., et al. 2016³](#). Additional direct targets include *xyiR* and *pykA*; different residues within DicF are used for base-pairing with the different targets [Balasubramanian D., et al. 2016³](#). Like SgrS, DicF was found to regulate the translation of *manX* indirectly, by facilitating the interaction of Hfq with a site near the ribosome binding site, which then inhibits initiation of translation [Azam MS., et al. 2018¹⁵](#).

Under anaerobic conditions, the enolase component of the [FRAME: CPLX0-2381] together with Hfq stabilizes DicF, leading to a predominantly filamentous morphology of the cells [Murashko ON., et al. 2017²³](#).

The DicF-mediated delay in nucleoid segregation is suppressed by FtsZ overproduction [Tatart F., et al. 1992¹](#). The DicF-mediated decrease in FtsZ abundance and activity is suppressed by an *rpoB* mutation, and this effect is partially counteracted by an *rpoS* mutation [Cam K., et al. 1995²¹](#).

DicF was shown to bind to Hfq [Zhang A., et al. 2003⁵](#), [Chihara K., et al. 2022²²](#), with a K_d of about 16 nM [Ojciniczak M., et al. 2011¹³](#). The DicF RNA is generated via processing of a larger transcript by RNase III and RNase E [Faulstich M., et al. 1990²⁰](#), [Murashko ON., et al. 2017²³](#), [Ragunathan PT., et al. 2021¹⁷](#). DicF predominantly accumulates under oxygen-limited conditions [Murashko ON., et al. 2017²³](#). It has been observed that DicF production occurs when *E. coli* is cultivated aerobically during the stationary phase [Ragunathan PT., et al. 2023¹⁷](#). Additionally, its expression is induced by urea and high temperatures [Ragunathan PT., et al. 2023¹⁷](#).

Genes similar to *dicF* occur in multiple prophage-like elements in a variety of genomes [Faulstich M., et al. 1994¹⁸](#).

Overexpression of DicF leads to decreased expression of RpoS [Mandin P., et al. 2010⁹](#), slower cell growth and decreased swimming and swarming motility and biofilm formation [Bak G., et al. 2015¹⁴](#).

Reviews: [Dethas N., et al. 1995¹⁵](#), [Wassarman KM., et al. 1999¹⁰](#), [Storz G., et al. 2004⁷](#), [Gottesman S., et al. 2004¹²](#), [Altuvia S., et al. 2018¹⁹](#), [Hor J., et al. 2020⁶](#).

Citations
(1) Tatart F., Albiges-P., Carter A., Makler F., Bouche JP. 1992. Involvement of FtsZ in coupling of nucleoid separation with septation. Evidence: FEVJEXP-3GI
(2) Wundtli S., von Pelster F., Mayer O., Casarini F., Schneider B. 2008. Isolation of small RNA-binding proteins from *E. coli*: evidence for frequent interaction of RNAs with RNA polymerase.
(3) Balasubramanian D., Ragunathan PT., Fan J., Vanderpool CK. 2016. A Prophage-Encoded Small RNA Controls Metabolism and Cell Division in *Escherichia coli*. Evidence: FEVJEXP-IMP
(4) van Helvoort M., Kool J., Woldringh CL. 1999. Chloramphenicol causes fusion of separated nucleoids in *Escherichia coli* K-12 cells and filaments.

regulondb.ccg.unam.mx/regulon/RDBECOLIPDC00328

Schritt 1 – Daten sammeln

```
chrome_options = webdriver.ChromeOptions()
#chrome_options.add_argument("--headless")
chrome_options.add_argument("--no-sandbox")
chrome_options.add_argument('--window-size=1920,1080')
chrome_options.add_experimental_option("prefs", {
    "download.prompt_for_download": False,
    "download.directory_upgrade": True,
    "safebrowsing.enabled": True
})
driver = webdriver.Chrome(options=chrome_options)
actions = ActionChains(driver)

def click_button_wait(XPATH):
    WebDriverWait(driver, 10).until(EC.element_to_be_clickable((By.XPATH, XPATH))).click()

def find_next_page_button():
    driver.execute_script("window.scrollTo(0, document.body.scrollHeight);") # scroll down
    button_xpath = "/html/body/div[1]/div/div[2]/div/div[2]/div/div[3]/div/table/tbody/tr[19]/td/div/div/div/button[3]"
    button = WebDriverWait(driver, 10).until(EC.element_to_be_clickable((By.XPATH, button_xpath)))
    button.click()

# seite aufrufen
driver.get("https://regulondb.ccg.unam.mx/regulon")

# Cookies Rejecten
click_button_wait('/html/body/div[1]/div/div[4]/div/div[3]/button[2]')

is_last_page = False
regulon_links = {}
try:
    while True:
        # Get Table Body
        table_body = WebDriverWait(driver, 10).until(EC.visibility_of_element_located((By.XPATH, '/html/body/div[1]/div/div[2]/div/div[2]/div/div[3]/div/table/tbody')))
        # jedes tr (zeile der tabelle) sammeln
        elements = table_body.find_elements(By.XPATH, ".*[contains(@id, 'tr')]")


        for i in range(0, len(elements)):
            try:
                link = elements[i].find_elements(By.XPATH, ".*[0].find_element(By.XPATH, "div/div/a").get_attribute("href")
                name = elements[i].find_elements(By.XPATH, ".*[0].find_element(By.XPATH, "div/div/a").get_attribute("name")
                regulon_links[name] = link
            except:
                pass

        find_next_page_button()
except:
    print("finished")
```



```
{'DicF': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC00328',
'DsrA': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC00358',
'GcvB': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03275',
'McaS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03411',
'GadY': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03413',
'MicF': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03508',
'OxyS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03909',
'ChiX': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04122',
'SgrS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04132',
'IstR-1': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04134',
'RseX': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04142',
'RydC': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04143',
'OhsC': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04144',
'SymR': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04148',
'MgrR': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04149',
'FnrS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04150',
'ArrS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04151',
'SdsN': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04154',
'AgrB': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04160',
```

Schritt 2 – Daten abfragen

RegulonDB 

Search Tools Releases & Downloads API & software Help

Qin prophage; small regulatory RNA DicF

Synonyms: DicF
Type: RNA-Mediated-Translation-Regulation

Regulator DicF

Encoded By
Gene: [dicF](#)

Note:
The DicF small regulatory RNA acts to inhibit cell division [Bouché F. et al. 1999¹](#). DicF causes inhibition of the formation of the constriction and the separation of the replicated chromosomes into the daughter nucleoids [Tatart F. et al. 1992¹](#).

DicF interferes with translation of *ftsZ* [Tatart F. et al. 1992¹](#), by base-pairing with the *ftsZ* mRNA [Balasubramanian D. et al. 2016²](#). Additional direct targets include *xyiR* and *pykA*, different residues within DicF are used for base-pairing with the different targets [Balasubramanian D. et al. 2016²](#). Like SgrS, DicF was found to regulate the translation of *manX* indirectly, by facilitating the interaction of Hfq with a site near the ribosome binding site, which then inhibits initiation of translation [Acam MS. et al. 2018³](#).

Under anaerobic conditions, the enolase component of the [FRAME: CPLX0-2381] together with Hfq stabilizes DicF, leading to a predominantly filamentous morphology of the cells [Murashko ON. et al. 2017²³](#).

The DicF-mediated delay in nucleoid segregation is suppressed by FtsZ overproduction [Tatart F. et al. 1992¹](#). The DicF-mediated decrease in FtsZ abundance and activity is suppressed by an *rpoD* mutation, and this effect is partially counteracted by an *rpoS* mutation [Cam K. et al. 1999²¹](#).

DicF was shown to bind to Hfq [Zheng A. et al. 2003²²](#), with a K_D of about 16 nM [Chenican M. et al. 2011¹³](#). The DicF RNA is generated via processing of a larger transcript by RNase III and RNase E [Faulstich M. et al. 1992²](#), [Murashko ON. et al. 2017²³](#), [Baouathan PT. et al. 2023¹⁷](#). DicF predominantly accumulates under oxygen-limited conditions [Murashko ON. et al. 2017²³](#). It has been observed that DicF production occurs when *E. coli* is cultivated aerobically during the stationary phase [Baouathan PT. et al. 2023¹⁷](#). Additionally, its expression is induced by urea and high temperatures [Baouathan PT. et al. 2023¹⁷](#).

Genes similar to *dicF* occur in multiple prophage-like elements in a variety of genomes [Faulstich M. et al. 1994¹⁰](#).

Overexpression of DicF leads to decreased expression of RpoS [Mandin P. et al. 2019²](#), slower cell growth and decreased swimming and swarming motility and biofilm formation [Bak G. et al. 2015¹⁴](#).

Reviews: [Delmas N. et al. 1999¹⁵](#), [Wasserman KM. et al. 1999¹⁶](#), [Storz G. et al. 2004¹²](#), [Gottesman S. et al. 2004¹²](#), [Ahuvia S. et al. 2018¹⁸](#), [Hor J. et al. 2020⁶](#). | [Show less](#)

Citations
(1) Tatart F, Alléport P, Contre A, Mulder F, Bouché JP. 1992. Involvement of FtsZ in coupling of nucleoid separation with septation. Evidence. *EVJ/IEVP-303*
(2) Wundrich N, von Peltzmann F, Järegård O, Casagrande B, Schroeder B. 2018. Isolation of small RNA-binding proteins from *E. coli*: evidence for frequent interaction of RNAs with RNA polymerase.
(3) Balasubramanian D, Zengwu C, Fan J. 2016. A Prophage-Encoded Small RNA Controls Morphology and Cell Division in *Escherichia coli*. Evidence. *EVJ/IEVP-303*
(4) van Heltvoort JM, Knoll J, Woldrich C. 1995. Chloramphenicol causes fusion of separated nucleoids in *Escherichia coli* K-12 cells and filaments.

Regulon

Regulatory Network



Triplets

Regulatory Interactions						
Active Conformation Name	Function	Regulated Entity Name	Regulated Entity Type	Distance to First Gene	Distance to Promoter	Regulated Genes
Filter... (1)	Filter... (1)	Filter... (4)	Filter... (1)	Min (-15) Max (42)	Filter... (1)	Filter... (4)
DicF	repressor	ftsZ	gene	-15		ftsZ
DicF	repressor	xyiR	gene	13		xyiR
DicF	repressor	pykA	gene	-10		pykA
DicF	repressor	manX	gene	42		manX

RBS LeftEndPosition
Min (105277)
Max (373498)

Schritt 2 – Daten abfragen

The image displays the RegulonDB website interface on the left and a browser's developer tools on the right, illustrating the process of querying data.

RegulonDB Interface:

- Search Bar:** Contains the text "Qin prophage; small regulatory RNA DicF".
- Regulator DicF:** A section detailing the gene *dicE*. It includes a note: "The *DicF* small regulatory RNA acts to inhibit cell division" (Bouche *et al.*, 1989). It also mentions that *DicF* causes inhibition of the formation of the constriction and the separation of the replicated chromosomes into the daughter nucleoids (Tetart *et al.*, 1992).
- Citations:** A list of references, including: (1) Tetart F, Albisot R, Conter A, Mulder F, Bouche JP. 1992. Involvement of *FtsZ* in coupling of nucleoid separation with septation. Evidence: [EV1]EXP-IGI; (2) Windichler N, von Pelchrim F, Mayer O, Csaszar E, Schroeder R. 2008. Isolation of small RNA-binding proteins from *E. coli*: evidence for frequent interaction of RNAs with RNA polymerase; (3) Balasubramanian D, Raghunathan PT, Fei J, Vanderpool CK. 2016. A Prophage-Encoded Small RNA Controls Metabolism and Cell Division in *Escherichia coli*. Evidence: [EV2]EXP-IMP; (4) van Helvoort JM, Kool J, Woldriphe CL. 1996. Chloramphenicol causes fusion of separated nucleoids in *Escherichia coli* K-12 cells and filaments.
- Regulon Network:** A section for visualizing the regulatory network, with a "Layout" dropdown set to "dagre".

Browser Developer Tools:

- Network Tab:** Shows a list of requests. The request to `https://regulondb.ccg.unam.mx/graphql` is highlighted with a yellow box.
- Request Details:** The selected request shows a status of 200 OK. The request headers include: `Accept: */*`, `Accept-Encoding: gzip, deflate, br, zstd`, `Accept-Language: de,en-US;q=0.9,en;q=0.8`, `Connection: keep-alive`, `Content-Length: 8838`, `Content-Type: application/json`, `Cookie: _ga=GA1.1.1106156118.1722190013; cookiePolicy=false; _ga_0ZFSS5E5YD=GS1.1.1722190013.1.1.1722191291.0.0.0`, `Host: regulondb.ccg.unam.mx`.

Schritt 2 – Daten abfragen

```
url = "https://regulondb.ccg.unam.mx/graphql"
```

```
request1["variables"]["advancedSearch"] = "RDBECOLIPDC00328[_id]"  
response = json.loads(requests.post(url = url, json = request1, verify = False).content)
```

GraphQL Endpoint

```
C:\Anaconda\lib\site-packages\urllib3\connectionpool.py:1061: InsecureRequestWarning: Unverified HTTPS request is being made  
to host 'regulondb.ccg.unam.mx'. Adding certificate verification is strongly advised. See: https://urllib3.readthedocs.io/en/  
1.26.x/advanced-usage.html#ssl-warnings  
warnings.warn()
```

```
response
```

```
{'data': {'getRegulonBy': {'data': [{'_id': 'RDBECOLIPDC00328',  
'alignmentMatrix': {'align': None,  
'consensus': None,  
'matrix': None,  
'urlMatrixQualityResult': None,  
'urlPWLLogo': None,  
'__typename': 'AlignmentMatrix'}},  
'allCitations': [{'evidence': {'_id': 'RDBECOLIEVC00062',  
'additiveEvidenceCodeRule': None,  
'code': 'EXP-IGI',  
'name': 'Inferred from genetic interaction',  
'type': 'W',  
'__typename': 'Evidence'},  
'publication': {'_id': 'RDBECOLIPRC06277',  
'authors': ['Tetart F',  
'Albigot R',  
'Conter A',  
'Mulder E',  
'Bouche JP'],  
'citation': 'Tetart F, Albigot R, Conter A, Mulder E, Bouche JP, 1992, Involvement of FtsZ in coupling of nucleoid s  
eparation with septation.',  
'pmid': '1552861',  
'title': 'Involvement of FtsZ in coupling of nucleoid separation with septation.'}]
```

```
{'DicF': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC00328',  
'DsrA': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC00358',  
'GcvB': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03275',  
'McaS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03411',  
'GadY': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03413',  
'MicF': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03508',  
'OxyS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC03909',  
'ChiX': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04122',  
'SgrS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04132',  
'IstR-1': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04134',  
'RseX': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04142',  
'RydC': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04143',  
'OhsC': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04144',  
'SymR': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04148',  
'MgrR': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04149',  
'FnrS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04150',  
'ArrS': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04151',  
'SdsN': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04154',  
'AgrB': 'https://regulondb.ccg.unam.mx/regulon/RDBECOLIPDC04160',
```

Schritt 3 – Daten Preparieren

sRNA encoded by dsrA (Q50418258)

Microbial non-coding RNA 'antisense RNA' found in Escherichia coli str. K-12 substr. MG1655. Note: silencer of rcsA gene, interacts with rpoD. [translation: component of acid resistance regulatory circuit](#)
ECK1952 | ISO95 | JWR0030 | dsrA | b1954

~ In more languages
Configure

Language	Label	Description	Also known as
English	sRNA encoded by dsrA	Microbial non-coding RNA 'antisense RNA' found in Escherichia coli str. K-12 substr. MG1655. Note: silencer of rcsA gene, interacts with rpoD translation: component of acid resistance regulatory circuit	ECK1952 ISO95 JWR0030 dsrA b1954
German	No label defined	No description defined	
French	No label defined	No description defined	
Bavarian	No label defined	No description defined	

Statements

Instance of	RNA	edit
	0 references	+ add reference
	non-coding RNA	edit
	0 references	+ add reference
		+ add value

antisense inhibitor of	DNA-binding transcriptional dual regulator H-NS b1237	edit
	0 references	+ add reference
	RNA polymerase, sigma S (sigma 38) factor b2741	edit
	1 reference	
	acetolactate synthase/acetohydroxybutanoate synthase, catalytic subunit b0077	edit
	1 reference	
	DNA-binding transcriptional dual regulator ArgR b3237	edit
	1 reference	
	D-ribose pyranase b3748	edit
	0 references	+ add reference
		+ add value

found in taxon	Escherichia coli str. K-12 substr. MG1655	edit
	0 references	+ add reference
		+ add value

strand orientation	reverse strand	edit
	RefSeq genome ID NC_000913.3	
	0 references	+ add reference
		+ add value

genomic start	2025226	edit
	RefSeq genome ID NC_000913.3	
	0 references	+ add reference
		+ add value

genomic end	2025313	edit
	RefSeq genome ID NC_000913.3	
	0 references	+ add reference
		+ add value

UPLOAD REGULONS

1. kurze beschreibung des items
2. konstant: instance of [P31] (RNA [Q11053] / non-coding RNA [Q427087])
3. konstant: found in taxon [P703] (Escherichia coli str. K-12 substr. MG1655 [Q21102938])
4. statements für gene: inhibitor / repressor / dual
5. strand orientation [P2548] (reverse [Q22809711] / forward [Q22809680])
6. genomic start [P644] / genomic end [P645]

```
1 list(regulon_data.values())[0]
```

```
{'beschreibung': 'The DicF small regulatory RNA acts to inhibit cell division.',  
'left_position': 1649382,  
'right_position': 1649434,  
'strand_orientation': 'Q22809711'}
```

```
1 list(triplets.values())[0]
```

```
[{'subjekt': 'DicF', 'predikat': 'repressor', 'objekt': 'ftsZ'},  
{ 'subjekt': 'DicF', 'predikat': 'repressor', 'objekt': 'xylR'},  
{ 'subjekt': 'DicF', 'predikat': 'repressor', 'objekt': 'pykA'},  
{ 'subjekt': 'DicF', 'predikat': 'repressor', 'objekt': 'manX'}]
```

```
1 print(gene_ids["ftsZ"])  
2 print(gene_ids["xylR"])  
3 print(gene_ids["pykA"])  
4 print(gene_ids["manX"])
```

```
Q127734188  
Q127735541  
Q127735563  
Q127735585
```


Schritt 4 – Daten hochladen (PyWikiBot)

Genes erzeugen

```
1 i = 0
2 for key, value in list(gene_descriptions.items())[6:]:
3     label = key
4     description = value
5
6     # Create Item
7     data = {
8         'labels': {'en': {'language': 'en', 'value': label}},
9         'descriptions': {'en': {'language': 'en', 'value': str(description)}}
10    }
11    new_item = pywikibot.ItemPage(repo)
12    new_item.editEntity(data, summary=f'Creating new item: {label}')
13
14    # Create Statement
15    claim = pywikibot.Claim(repo, "P31") # instance of
16    target = pywikibot.ItemPage(repo, "Q7187") # gene
17    claim.setTarget(target)
18    new_item.addClaim(claim, summary='Adding education information')
19
20    gene_ids[key] = new_item.getID()
21    print(key, i, "/", len(gene_descriptions))
22    i += 1
```

Sleeping for 9.4 seconds, 2024-07-24 15:43:27

citR 2686 / 2697

Sleeping for 9.3 seconds, 2024-07-24 15:43:38
Sleeping for 9.6 seconds, 2024-07-24 15:43:47

glsB 2687 / 2697

Sleeping for 9.4 seconds, 2024-07-24 15:43:57
Sleeping for 9.3 seconds, 2024-07-24 15:44:08

yneG 2688 / 2697

Sleeping for 9.2 seconds, 2024-07-24 15:44:18
Sleeping for 9.4 seconds, 2024-07-24 15:44:27

sfmC 2689 / 2697

Sleeping for 9.4 seconds, 2024-07-24 15:44:37
Sleeping for 9.5 seconds, 2024-07-24 15:44:47

RegulonDB			Sei
Genes			
Name			
Filter... (4747)			
3'ETS-leuZ	G0-16636	small regulatory RNA 3'ETS ^{leuZ}	
C0067	G0-8896		
C0293	G0-8897	e14 prophage; small regulatory RNA C0293	
C0299	G0-8898		
C0362	G0-8902		
C0465	G0-8903		
C0614	G0-8907		
C0664	G0-8911		
C0719	G0-8912		
G0-10697	G0-10697		
G0-10698	G0-10698		
G0-10700	G0-10700		
G0-10702	G0-10702		
G0-10703	G0-10703		

Schritt 4 – Daten hochladen (PyWikiBot)

```
def generate_claims(regulon, info):  
    label = regulon  
    if info["beschreibung"] != None:  
        description = info["beschreibung"]  
    else:  
        description = "description: " + regulon  
  
    if len(str(description)) >= 250:  
        description = str(description[:249])  
    # Create Item  
    data = {  
        'labels': {'en': {'language': 'en', 'value': label}},  
        'descriptions': {'en': {'language': 'en', 'value': str(description)}}  
    }  
    new_item = pywikibot.ItemPage(repo)  
    new_item.editEntity(data, summary=f'Creating new item: {label}')
```

Item wird erzeugt

```
claims = []  
# claims für instance of  
instance_of_rna = {'mainsnak': {'snaktype': 'value', 'property': "P31", 'datavalue': {'value': {'entity-type': 'item', 'numeric-id': "11053"}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
claims.append(instance_of_rna)  
instance_of_non_coding_rna = {'mainsnak': {'snaktype': 'value', 'property': "P31", 'datavalue': {'value': {'entity-type': 'item', 'numeric-id': "427087"}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
claims.append(instance_of_non_coding_rna)  
  
# claims für funktionen der gene (repressor, activator, dual)  
for i in range(0, len(triplets[regulon])):  
    value = triplets[regulon][i]  
    if value["predikat"] == "repressor":  
        repressor = {'mainsnak': {'snaktype': 'value', 'property': "P3776", 'datavalue': {'value': {'entity-type': 'item', 'numeric-id': str(gene_ids[value["objekt"]][1:])}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
        claims.append(repressor.copy())  
    elif value["predikat"] == "activator":  
        activator_of = {'mainsnak': {'snaktype': 'value', 'property': "P3771", 'datavalue': {'value': {'entity-type': 'item', 'numeric-id': str(gene_ids[value["objekt"]][1:])}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
        claims.append(activator_of.copy())  
    elif value["predikat"] == "dual":  
        dual = {'mainsnak': {'snaktype': 'value', 'property': "P1322", 'datavalue': {'value': {'entity-type': 'item', 'numeric-id': str(gene_ids[value["objekt"]][1:])}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
        claims.append(dual.copy())  
  
# claim für found in taxon  
found_in_taxon = {'mainsnak': {'snaktype': 'value', 'property': "P703", 'datavalue': {'value': {'entity-type': 'item', 'numeric-id': "21102938"}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
claims.append(found_in_taxon)  
  
# claims für strand orientation  
if info["left_position"] != None:  
    genomic_start = {'mainsnak': {'snaktype': 'value', 'property': "P644", 'datavalue': {'value': str(info["left_position"]), 'type': 'string'}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
    claims.append(genomic_start)  
if info["right_position"] != None:  
    genomic_end = {'mainsnak': {'snaktype': 'value', 'property': "P645", 'datavalue': {'value': str(info["right_position"]), 'type': 'string'}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
    claims.append(genomic_end)  
if info["strand_orientation"] != None and info["left_position"] != None and info["right_position"] != None:  
    switch = {"22809711" : "22809680", "22809680" : "22809711"}  
    so = switch[str(info["strand_orientation"])[1:]]  
    strand_orientation = {'mainsnak': {'snaktype': 'value', 'property': "P2548", 'datavalue': {'value': {'entity-type': 'item', 'numeric-id': so}, 'type': 'wikibase-entityid'}, 'type': 'statement', 'rank': 'normal'}}  
    claims.append(strand_orientation)  
  
# upload  
data = {'claims': claims}  
new_item.editEntity(data, summary='Adding multiple claims in one request')  
return new_item.getID()
```

Ergebnis

OxyS (Q127789260)

OxyS is a small RNA that plays a regulatory role in the oxidative stress response and has an antimutator effect.

[In more languages](#)

[Configure](#)

Language	Label	Description	Also known as
English	OxyS	OxyS is a small RNA that plays a regulatory role in the oxidative stress response and has an antimutator effect.	
German	No label defined	No description defined	
French	No label defined	No description defined	
Bavarian	No label defined	No description defined	

Statements

instance of

RNA

0 references

non-coding RNA

0 references

inhibitor of

rpoS

0 references

fhlA

0 references

[Wikipedia](#) (0 entries) [edit](#)

[Wikibooks](#) (0 entries) [edit](#)

[Wikinews](#) (0 entries) [edit](#)

[Wikiquote](#) (0 entries) [edit](#)

[Wikisource](#) (0 entries) [edit](#)

[Wikiversity](#) (0 entries) [edit](#)

[Wikivoyage](#) (0 entries) [edit](#)

[Wiktionary](#) (0 entries) [edit](#)

[Multilingual sites](#) (0 entries) [edit](#)

1 regulon_qids

```
{'OxyS': 'Q127789260',  
'ChiX': 'Q127789264',  
'SgrS': 'Q127789265',  
'IstR-1': 'Q127789266',  
'RseX': 'Q127789268',  
'RydC': 'Q127789270',  
'OhsC': 'Q127789272',  
'SymR': 'Q127789274',  
'MgrR': 'Q127789276',  
'FnrS': 'Q127789277',  
'ArrS': 'Q127789278',  
'SdsN': 'Q127789281',  
'AgrB': 'Q127789283',  
'RalA': 'Q127789285',  
'MicL-S': 'Q127789288',  
'CpxQ': 'Q127789289',  
'SdhX': 'Q127789292',  
'C0293': 'Q127789294',
```