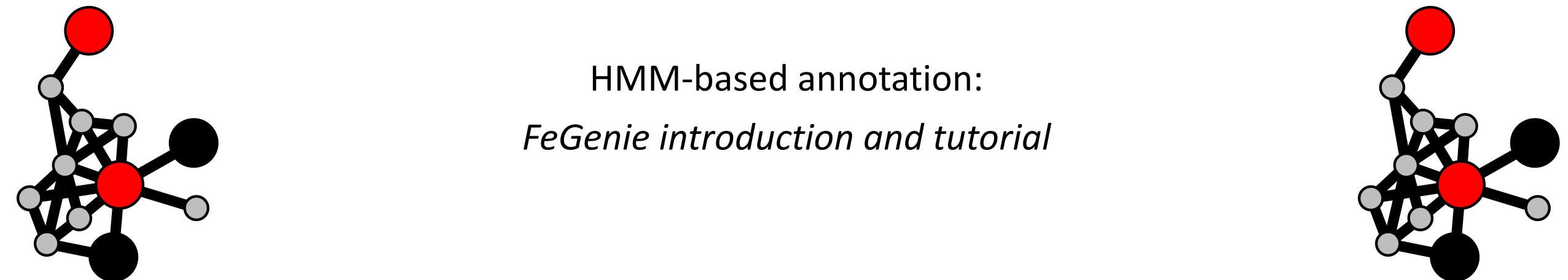
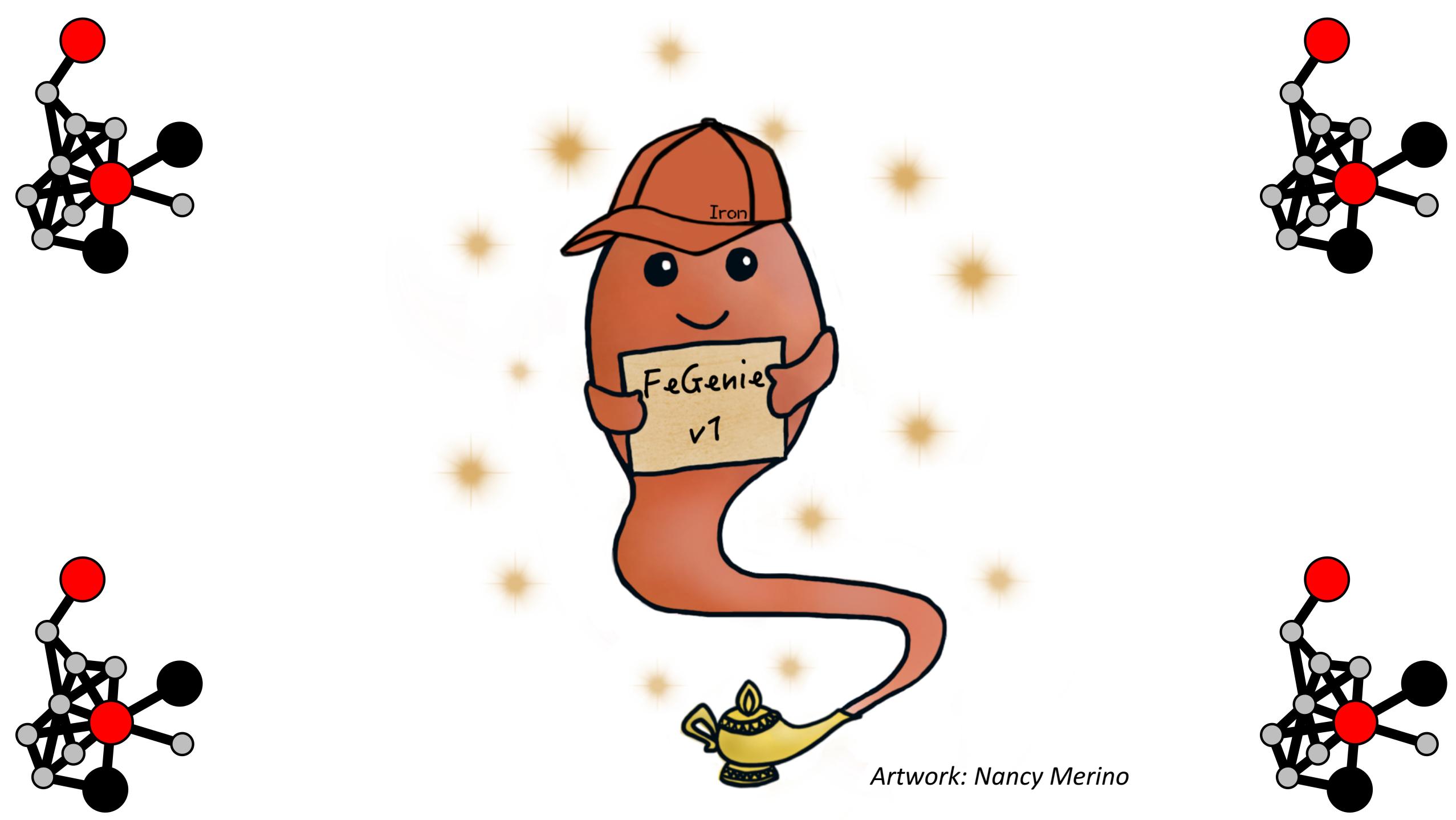




Functional annotation lesson



HMM-based annotation:
FeGenie introduction and tutorial



Artwork: Nancy Merino

What is FeGenie

1. A collection of HMMs, based on a comprehensive set of genetic markers related to iron acquisition/scavenging, transport, efflux, storage, as well as iron redox (dissimilatory reduction and oxidation)
2. A bioinformatics software that uses these HMMs to profile genomic datasets (in FASTA format)
3. Uses rules for filtering out potential false positives (operon structure, genetic co-occurrence, bit scores)

Category	Function	Protein Families
Iron acquisition	Iron(II)/(III) transport	Efe <u>UOB</u> ¹ , FbpABC ² , SfuABC ³ , YfuABC ⁴ , FeoAB(C) ⁵ , FutA1 ⁶ , FutA2 ⁶ , FutB ⁶ , FutC ⁶ , YfeABCD ⁷
	Heme oxygenase	ChuS ⁸ , ChuZ ⁹ , HemO ^{10,11} , PigA ^{10,11} , Hem <u>RSTUV</u> ¹² , HmoB ¹³ , HmuO ¹⁴ , HugZ ¹⁵ , HupZ ¹⁶ , Isd-LmHde ¹⁷ , IsdG ¹⁸ , IsdI ¹⁹ , MhuD ²⁰ , PhuS ²¹ (in PhuRSTUVW)
	Heme transport	HasRADE(B)E ²² , HmuR <u>STUV</u> ²² , HmuY ²³ , HmuY' ²³ , HutZ ²⁴ , Hxu <u>CBA</u> ²⁵ , IsdX1 ²⁶ , IsdX2 ²⁶ , Phu <u>RSTUVW</u> ²¹ , Rv0203 ²⁷
	Transferrin/Lactoferrin	Tbp <u>AB</u> (Lbp <u>AB</u>) ²⁸ , Sst <u>ABCD</u> ²⁹
	Siderophore synthesis	Acs <u>ABCDEF</u> ³⁰ , AmoA ³¹ , AngR ³² , Asb <u>ABCDEF</u> ³³ , DhbACEBF ³⁴ , entD- feP <u>A</u> -fes-entF-fepEC <u>GDB</u> -entCEBA-ybdA ³⁵ , IroD in Iro <u>NBCDE</u> ³⁶ , lucABCD ^{37,38} , IutA ^{37,38} , MbtIABCDEF ³⁵ , LbtA ³⁹ (in LbtUABC), PchABCDREFGH ³⁵ , PvdQAPMNOFEDJIHLGS ⁴⁰ , PvsABCDE ⁴¹ , VenB ⁴² , Vab genes in VabR-fur-vabGA-fur-VabCEBSFH-fur-fvtA-vabD ⁴³ , Vib genes in VibB-vibEC-vibA-vibH-viuP <u>DGC</u> -vibD and viuAB-vibF ⁴⁴⁻⁴⁶ , RhbABCDEF-rhrA-rhtA ⁴⁷
	Siderophore transport	BesA ⁴⁸ , Cbr <u>ABCD</u> ⁴⁹ , TonB-ExbB-ExbD ⁵⁰ , Fat <u>ABCD</u> ⁵¹ , FecI <u>RABCDE</u> ⁵² , FeuABC-yusV ⁵³ , Fhu <u>ACDB</u> ⁵⁴⁻⁵⁶ , FhuF ⁵⁴⁻⁵⁶ , Fpt <u>ABCX</u> ⁵⁷ , FpuA ⁵⁸ , FpuC ⁵⁸ , FpuD ⁵⁸ , FpvI R-FpvA -FpvGHJKCDEF ⁵⁹ , FvtA in VabR-fur-vabGA-fur-VabCEBSFH-fur-fvtA-vabD ⁴³ , Hat <u>CD</u> ³⁷ , Iro <u>NBCDE³⁶, LbtUABC³⁹, PirA⁶⁰, PiuA⁶⁰, Pvu<u>ABCDE</u>⁴¹, Viu genes in VibB-vibEC-vibA-vibH-viuP<u>DGC</u>-vibD and viuAB-vibF⁴⁴⁻⁴⁶, YifZ-yfhA⁶¹, YfiY⁶¹, YqjH⁶², ybdA and Fep genes in entD-feP<u>A</u>-fes-entF-fepEC<u>GDB</u>-entCEBA-ybdA³⁵</u>
Iron Gene regulation	Transcriptional regulation	DtxR ⁶³ , FecR (in FecI <u>RABCDE</u>) ⁵² , FeoC in FeoAB(C) ⁵ , Fur ⁶⁴ , IdeR ⁶⁵ , YqjI ⁶² , RhrA in RhbABCDEF-rhrA-rhtA ⁴⁷
Iron oxidation and reduction	Iron oxidation	Cyc1 ^{66,67} , Cyc2 ^{66,67,68} , FoxABC ⁶⁹ , FoxEYZ ⁷⁰ , Sulfocyanin ⁷¹ , PioA <u>BC</u> ⁷²
	Probable iron oxidation and possible iron reduction	MtoAB ⁷³ , Cyc2 (cluster 3)
	Dissimilatory iron reduction	CymA ⁷⁴ , Mtr <u>CAB</u> ⁷⁵ , OmcF ⁷⁶ , OmcS ⁷⁶ , OmcZ ⁷⁶ , FmnA-dmkA-fmnB-pplA-ndh2-eetAB-dmkB ⁷⁷ , DFE_0448-0451, DFE_0461-0465 ⁷⁸
Iron storage	Probable iron reduction	MtrCB, MtrAB, MtoAB-MtrC
	Iron storage	Bfr ⁷⁹ , DpsA ⁸⁰ , Ftn ⁸¹
Magnetosome-related	Magnetosome formation	MamABEKLMOPQI ^{82,83} (Note: These genes are found in all known magnetotactic microorganisms, except for mamL which is found in magnetite-producing magnetotactic microorganisms ⁸¹)

Bolded and underlined HMMs are derived from Pfam or TIGRFAMs databases. Other HMMs were created by using select sequences. See **Supplementary Table S1** for more information, including the corresponding Pfam or TIGRFAMs families and the sequences used to create the HMMs. ¹Miethke et al., 2013, ²Adhikari et al., 1996, ³Angerer et al., 1990, ⁴Gong et al., 2001, ⁵Lau et al., 2016, ⁶Katoh et al., 2001, ⁷Bearden et al., 1998, ⁸Suits et al., 2006, ⁹Zhang et al., 2011, ¹⁰Friedman et al., 2003, ¹¹Friedman et al., 2004, ¹²Schneider et al., 2006, ¹³Park et al., 2012, ¹⁴Matsui et al., 2005, ¹⁵Hu et al., 2011, ¹⁶Sachla et al., 2016, ¹⁷Duong et al., 2014, ¹⁸Reniere et al., 2010, ¹⁹Skaar et al., 2004, ²⁰Graves et al., 2014, ²¹Ochsner et al., 2000, ²²Tong and Guo, 2009, ²³Wojtowicz et al., 2009, ²⁴Liu X. et al., 2012, ²⁵Morton et al., 2007, ²⁶Honsa et al., 2014, ²⁷Tullius et al., 2011, ²⁸Gray-Owen et al., 1995, ²⁹Morrissey et al., 2000, ³⁰Carroll and Moore, 2018, ³¹Barghouthi et al., 1991, ³²Wertheimer et al., 1999, ³³Oves-Costales et al., 2007, ³⁴May et al., 2001, ³⁵Crosa and Walsh, 2002, ³⁶Hantke et al., 2003, ³⁷Suzuki et al., 2006, ³⁸Martinez et al., 1994, ³⁹Cianciotto, 2015, ⁴⁰Lamont and Martin, 2003, ⁴¹Tanabe et al., 2003, ⁴²Tan et al., 2014, ⁴³Balado et al., 2008, ⁴⁴Wyckoff et al., 2001, ⁴⁵Keating et al., 2000, ⁴⁶Wyckoff et al., 1999, ⁴⁷Lynch et al., 2001, ⁴⁸Miethke et al., 2006, ⁴⁹Mahé et al., 1995, ⁵⁰Garcia-Herrero et al., 2007, ⁵¹Lemos et al., 2010, ⁵²Braun, 2003, ⁵³Peuckert et al., 2011, ⁵⁴Köster and Braun, 1989, ⁵⁵Coulton et al., 1987, ⁵⁶Braun et al., 2002, ⁵⁷Youard et al., 2011, ⁵⁸Dixon et al., 2012, ⁵⁹Brillet et al., 2012, ⁶⁰Moynie et al., 2017, ⁶¹Ollinger et al., 2006, ⁶²Wang et al., 2011, ⁶³Guedon and Helmann, 2003, ⁶⁴Escolar et al., 1998, ⁶⁵Rodriguez et al., 2002, ⁶⁶Castelle et al., 2008, ⁶⁷Barco et al., 2015, ⁶⁸McAllister et al., 2019, ⁶⁹Bathe and Norris, 2007, ⁷⁰Croal et al., 2007, ⁷¹Ilbert and Bonnefoy, 2013, ⁷²Liu J. et al., 2012, ⁷³Jiao and Newman, 2007, ⁷⁴Castelle et al., 2015, ⁷⁵Pitts et al., 2003, ⁷⁶Santos et al., 2015, ⁷⁷Light et al., 2018, ⁷⁸Deng et al., 2018, ⁷⁹Grossman et al., 1992, ⁸⁰Grant et al., 1998, ⁸¹Andrews, 1998, ⁸²Uebe and Schuler, 2016, ⁸³Kolinko et al., 2016.

Bolded and underlined genes represent HMMs taken from Pfam and TIGRFAMs

Others were designed specifically for FeGenie

Getting HMMs from Pfam



HOME | SEARCH | BROWSE | FTP | HELP | ABOUT

Pfam
keyword search Go

Family: 7tm_1 (PF00001)

Summary

Domain organisation

Clan

Alignments

HMM logo

Trees

Curation & model

Species

Interactions

Structures

Jump to... ⓘ

enter ID/acc Go

Summary: 7 transmembrane receptor (rhodopsin family)

Pfam includes annotations and additional family information from a range of different sources. These sources can be accessed via the tabs below.

Wikipedia: Rhodopsin-like receptors Pfam InterPro

This is the Wikipedia entry entitled "Rhodopsin-like receptors". [More...](#)

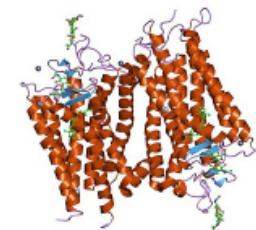
Rhodopsin-like receptors [Edit Wikipedia article](#)

Rhodopsin-like receptors are a family of proteins that comprise the largest group of G protein-coupled receptors.^[2]

Contents [hide]

- 1 Scope
- 2 Function
- 3 Classes
 - 3.1 Subfamily A1
 - 3.2 Subfamily A2
 - 3.3 Subfamily A3
 - 3.4 Subfamily A4
 - 3.5 Subfamily A5
 - 3.6 Subfamily A6
 - 3.7 Subfamily A7
 - 3.8 Subfamily A8
 - 3.9 Subfamily A9
 - 3.10 Subfamily A10
 - 3.11 Subfamily A11
 - 3.12 Subfamily A12
 - 3.13 Subfamily A13
 - 3.14 Subfamily A14
 - 3.15 Subfamily A15
 - 3.16 Subfamily A16

Rhodopsin-like receptors



Structure of rhodopsin: A G protein-coupled receptor.^[1]

Identifiers	
Symbol	7tm_1
Pfam	PF00001 ⓘ
Pfam clan	GPCR_A ⓘ
InterPro	IPR000276 ⓘ
PROSITE	PDOC00211 ⓘ
SCOPe	1frz ⓘ / 1ldeam ⓘ

Getting HMMs from Pfam

Family: **7tm_1 (PF00001)**



653 architectures



56771 sequences



12 interactions



425 species



344 structures

Summary

Domain organisation

Clan

Alignments

HMM logo

Trees

Curation & model

Species

Interactions

Structures

Jump to...

enter ID/acc

Go

Curation and family details

This section shows the detailed information about the Pfam family. You can see the definitions of many of the terms in this section in the [glossary](#) and a fuller explanation of the scoring system that we use in the [scores](#) section of the help pages.

Curation

Seed source:	Prosite
Previous IDs:	none
Type:	Family
Sequence Ontology:	SO:0100021
Author:	Sonnhammer ELL
Number in seed:	64
Number in full:	56771
Average length of the domain:	255.40 aa
Average identity of full alignment:	18 %
Average coverage of the sequence by the domain:	67.02 %

HMM information

HMM build commands: *build method: hmmbuild -o /dev/null --hand HMM SEED
search method: hmmsearch -Z 45638612 -E 1000 --cpu 4 HMM pfamseq*

Model details:	Parameter	Sequence	Domain
	Gathering cut-off	30.5	30.5
	Trusted cut-off	30.5	30.5
	Noise cut-off	30.4	30.4

Model length: 263

Family (HMM) version: 21

Download: [download](#) the raw HMM for this family

Getting HMMs from Pfam

```
HMMER3/f [3.1b2 | February 2015]
NAME 7tm_1
ACC  PF00001.21
DESC 7 transmembrane receptor (rhodopsin family)
LENG 263
ALPH amino
RF yes
MM no
CONS yes
CS yes
MAP yes
DATE Sun Aug 5 12:31:46 2018
NSEQ 64
EFFN 3.902344
GA 30.50 30.50;
TC 30.50 30.50;
NC 30.40 30.40;
DM hmmsuite --nanomHM.ann SEED.ann
SM hmmssearch -Z 45638612 -E 1000 --cpu 4 HMM pfamseq
STATS LOCAL MSV -11.1451 0.70242
STATS LOCAL VITERBI -11.8766 0.70242
STATS LOCAL FORWARD -5.2869 0.70242
HMM   A   C   D   E   F   G   H   I   K
L   M   N   P   Q   R   S   T   V   W
Y
m->m m->i m->d i->m i->i d->m d->d
COMPO 2.55504 3.82425 3.20574 3.00158 3.11128 2.96544 3.81642 2.70779 2.90079
2.32505 3.56717 3.09038 3.17297 3.27853 3.03161 2.57318 2.73378 2.51645 4.26736
```

```
arkadiygarber@Arkadiyys-MacBook-Pro:~$ hmmsearch -h
# hmmsearch :: search profile(s) against a sequence database
# HMMER 3.2.1 (June 2018); http://hmmer.org/
# Copyright (C) 2018 Howard Hughes Medical Institute.
# Freely distributed under the BSD open source license.
# -----
Usage: hmmsearch [options] <hmmfile> <seqdb>

Basic options:
-h : show brief help on version and usage

Options directing output:
-o <f>      : direct output to file <f>, not stdout
-A <f>      : save multiple alignment of all hits to file <f>
--tblout <f>  : save parseable table of per-sequence hits to file <f>
--domtblout <f> : save parseable table of per-domain hits to file <f>
--pfamtblout <f> : save table of hits and domains to file, in Pfam format <f>
--acc        : prefer accessions over names in output
--noali      : don't output alignments, so output is smaller
--notextw    : unlimit ASCII text output line width
--textw <n>    : set max width of ASCII text output lines [120] (n>=120)

Options controlling reporting thresholds:
-E <x>      : report sequences <= this E-value threshold in output [10.0] (x>0)
-T <x>      : report sequences >= this score threshold in output
--domE <x>    : report domains <= this E-value threshold in output [10.0] (x>0)
--domT <x>    : report domains >= this score cutoff in output

Options controlling inclusion (significance) thresholds:
--incE <x>    : consider sequences <= this E-value threshold as significant
--incT <x>    : consider sequences >= this score threshold as significant
--incdomE <x> : consider domains <= this E-value threshold as significant
--incdomT <x> : consider domains >= this score threshold as significant

Options controlling model-specific thresholding:
--cut_ga : use profile's GA gathering cutoffs to set all thresholding
--cut_nc : use profile's NC noise cutoffs to set all thresholding
--cut_tc : use profile's TC trusted cutoffs to set all thresholding
```

Getting HMMs from TIGRFAMS

J. Craig Venter[™]
INSTITUTE

TIGRFAMS

JCVI Home

TIGRFAMs Home

Genome Properties

→ TIGRFAMs Home

→ TIGRFAMs Terms

→ TIGRFAMs Complete Listing

→ TIGRFAMs FTP site

→ TIGRFAMs Resources

TIGRFAMs COMPLETE LISTING

Accession Name	Function	Isology Type	EC Number
TIGR00001 rpmI_bact	ribosomal protein bL35	equivalog	
TIGR00002 S16	ribosomal protein bS16	equivalog	
TIGR00003 TIGR00003	copper ion binding protein	subfamily_domain	
TIGR00004 TIGR00004	reactive intermediate/imine deaminase	subfamily	
TIGR00005 rluA_subfam	pseudouridine synthase, RluA family	subfamily	5.4.99.-
TIGR00006 TIGR00006	16S rRNA (cytosine(1402)-N(4))-methyltransferase 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide isomerase	equivalog	2.1.1.199
TIGR00007 TIGR00007		equivalog	5.3.1.16
TIGR00008 infA	translation initiation factor IF-1	equivalog	
TIGR00009 L28	ribosomal protein bL28	equivalog	
TIGR00010 TIGR00010	hydrolase, TatD family	subfamily	
TIGR00011 YbaK_EbsC	Cys-tRNA(Pro) deacylase	hypoth_equivalog	3.6.1.-
TIGR00012 L29	ribosomal protein uL29	equivalog	
TIGR00013 taut	4-oxalocrotonate tautomerase family enzyme	subfamily	5.3.2.-
TIGR00014 arsC	arsenate reductase (glutaredoxin)	equivalog	1.20.4.1
TIGR00016 ackA	acetate kinase	equivalog	2.7.2.1
TIGR00017 cmk	cytidylate kinase	equivalog	2.7.4.14
TIGR00018 panC	pantoate--beta-alanine ligase	equivalog	6.3.2.1
TIGR00019 prfA	peptide chain release factor 1	equivalog	
TIGR00020 prfB	peptide chain release factor 2	equivalog	
TIGR00021 rplA	ribose 5-phosphate isomerase A	equivalog	5.3.1.6
TIGR00022 TIGR00022	Yhch/YjgK/YiaL family protein	subfamily	
TIGR00023 TIGR00023	acyl-phosphate glycerol 3-phosphate acyltransferase	equivalog	2.3.1.-
TIGR00024 SbcD_rel_arch	putative phosphoesterase	hypoth_equivalog	3.1.-.-
TIGR00025 Mtu_efflux	ABC transporter efflux protein, DrrB family	paralog	
TIGR00026 hi_GC_TIGR00026	deazaflavin-dependent oxidoreductase, nitroreductase family	subfamily	
TIGR00027 methyl_TIGR00027	methyltransferase, TIGR00027 family	subfamily	2.1.1.-
TIGR00028 Mtu_PIN_fam	toxin-antitoxin system PIN domain toxin	subfamily	
TIGR00029 S20	ribosomal protein bS20	equivalog	
TIGR00030 S21p	ribosomal protein bS21	equivalog	
TIGR00031 UDP-GALP_mutase	UDP-galactopyranose mutase	equivalog	5.4.99.9
TIGR00032 argG	argininosuccinate synthase	equivalog	6.3.4.5
TIGR00033 aroC	chorismate synthase	equivalog	4.2.3.5
TIGR00034 aroFGH	3-deoxy-7-phosphoheptulonate synthase	equivalog	2.5.1.54

Getting HMMs from TIGRFAMS

JCVI Home | TIGRFAMs Home | Genome Properties

HMM SUMMARY PAGE: TIGR00001

TIGRFAMs Home	Accession	TIGR00001
TIGRFAMs Terms	Name	rpmI_bact
TIGRFAMs Complete Listing	Function	ribosomal protein bL35
TIGRFAMs FTP site	Gene Symbol	rnmI
TIGRFAMs Resources	Trusted Cutoff	51.20
TIGR00001 Seed Alignment	Domain Trusted Cutoff	51.20
	Noise Cutoff	23.65
	Domain Noise Cutoff	23.65
	Isology Type	equivalog
	HMM Length	63
	Mainrole Category	Protein synthesis
	Subrole Category	Ribosomal proteins: synthesis and modification
	Gene Ontology Term	GO:0000315 : organellar large ribosomal subunit cellular_component GO:0003735 : structural constituent of ribosome molecular_function GO:0006412 : translation biological_process GO:0022625 : cytosolic large ribosomal subunit cellular_component
	Author	Haft DH, Dodson RJ, Richardson DL
	Entry Date	Apr 20 1999 2:04PM
	Last Modified	Sep 4 2014 11:07AM
	Comment	This ribosomal protein is found in bacteria and organelles only. It is not closely related to any eukaryotic or archaeal ribosomal protein.
	References	SE TIGR DR HAMAP; MF_00514; 522 of 531
	Genome Property	GenProp0174 : ribosome, bacterial, large subunit (HMM) GenProp0799 : bacterial core gene set, exactly 1 per genome (HMM)

Getting HMMs from TIGRFAMS

[TIGRFAMs Home](#)

[TIGRFAMs Terms](#)

[TIGRFAMs Complete Listing](#)

[TIGRFAMs FTP site](#)

[TIGRFAMs Resources](#)

HMM SEED ALIGNMENT: TIGR00001

[Download Alignment](#)

```
>SP|P94976|RL35_MYCTU/2-64
PKAKTHSGASKRFRRTGT.GKIVRQKANRRHLLEHKPSTR.RRLDGRTVVAANDTKRVT.SLLNG
>SP|O88059|RL35_STRCO/2-64
PKNKSHSGASKRFKITGS.GKVLRERAGKRHLLEHKSSRVT.RRLTGNAEMAPGDAAKIK.KLLGK
>SP|P80341|RL35_THETH/1-63
PKMKTHGAKKRVKTAS.GKVVAMKTGKRHLNWQKSGKEI.RQ.KGRKFVLAPEAERIKLLLPHY
>SP|Q9RSW6|RL35_DEIRA/2-64
PKMKTHMAKRRIKTTGT.GKVMASFSGKRHQNTGKSGDEI.RG.KGKGFLAKAEWARMKLMPLR
>SP|Q9ZCV1|RL35_RICPR/2-65
PKLTKSAVKKRFKFTAT.GKVIASQAGKKHFMRRTKAQI.RNLRGTTILCPQDGHNICKYFLPY
>OMNI|NTL01BS2880/2-64
PKMKTHRGSACKRFKKTGS.GKLKRSHAYTSHLFANKSQKQK.RKLRSAVVSAGDFKRIK.QMLAN
>SP|P47439|RL35_MYCCE/1-59
..MKTKSAAVKRFKLTKS.GQIKRKHAYTSHLAPHKSTKQK.RHLRKQATVSNSELKRIG.ILI..
>GP|150150|gb|AAA25414.1/2-62
PKMKTRKSAKLRKRIKTTGT.GKTMREQAYRSHLSQNKTQK.RQARKSVQMHSVDVKRFK.ALI..
>SP|Q9X1S7|RL35_THEME/2-64
PKVKTNRSAAKRFRITKN.GKIMRNHAYRSHKTGKRRNAL.RALRKKDVVSSADKNRVL.RLLGK
>SP|P48959|RL35_SYN3/2-66
PKLTKRAAAKRFKPTGSKKIRRKAFKHNLLEHKSSEQHRRRLSNLALVHEADENVR.LMLPY
>SP|P14810|RK35_CYAPA/2-64
YKLTKRAAAKRYKAVGN.KKISRRKAFRSHLLQKKSTNRK.RQLSQVVIASPGDTKIIY.LMLPY
>SP|P49567|RK35_ODOSI/2-64
PKLTKRAALKRYKKTAT.GKFLLRRHAYKGHLLMKSQTQK.RKLSQIICVSNNDSKPIK.LMLPY
>SP|P51270|RK35 PORPU/2-64
PKLTKSAIAKRFKVSS.GKLRLRKASHSHLLQKKSSKR.RHLSSTCQVDSRDAKNIS.INLPY
>SP|P23326|RK35_SPIOL/90-152
YKMKTHRGSACKRFKVTGK.GKIVRERRAGKQHLLAKNTKRK.NRLSKLIQVDRSDYDNVI.GALPY
>SP|O78496|RK35_GUITH/2-64
PKLTKRKSARKRFYISAK.GKFVRKRAFKSHILEKKTSKRK.RNLKRKMIVFKGENLALK.TMLPY
>OMNI|NTL01EC01685/2-64
PKIKTVRGAAKRFKKTGK.GGFKHKHANLRHILTKATKRK.RHLRPKAMVSKGDGLGLVI.ACIPY
>SP|P52830|RL35_PSESM/2-63
PKMKTRKSAKRFKLTAN.G.IKHKHAFKSHILTKMSTKRK.RQLRGSSLHPSDVAKVE.RMLRL
>SP|O68845|RL35_VIBCH/2-63
PKMKNNGAAKRFKKTAG.G.IKYKHATKRHILTKRTTKNK.RQLRPNAILPKCELAAVA.RMLPY
>SP|Q9JQN7|RL35_NEIMA/2-64
PKMKTRKSAKRFKVGLN.GGVKRAHAFKRHILTKTCKNN.RQLRGTSMVNDRLASVA.KMLPY
>SP|O83821|RL35_TREPA/2-65
AKMKTRKSAAAKRFKVTGK.GKVKFKMMNLRHILTKKAPKRK.RKLHAGFLSKVELVVKRKLLPY
>SP|O51207|RL35_BORBU/3-65
NKMKTTRKSAKRYSFPTVN.GKVVKYKKQNLRHILTKKSSKRK.RNLRKSGNLSCFEVKRIK.TLLPY
>SP|O66982|RL35_AQUAE/4-66
VNMKTRKSAKRFKVTAK.GKIKRKWSGGAHYNTKSSKRK.RHLRKHTYVKDNMLKHKV.ALLKE
>SP|P56057|RL35_HELPY/2-63
PKMKTRNGASKRFKVY.K.NLIKRGSAFKSHILTKKSPKRK.ANLNAPKHHVHTNAHSVM.SLLCR
>SP|O84841|RL35_CHLTR/2-64
PKMKSNSKVAARFKLTGS.GQLKRTRPGKRHKLSKRSSQQK.RNLSKQPLVDQGQVGMYK.RMMLV
```

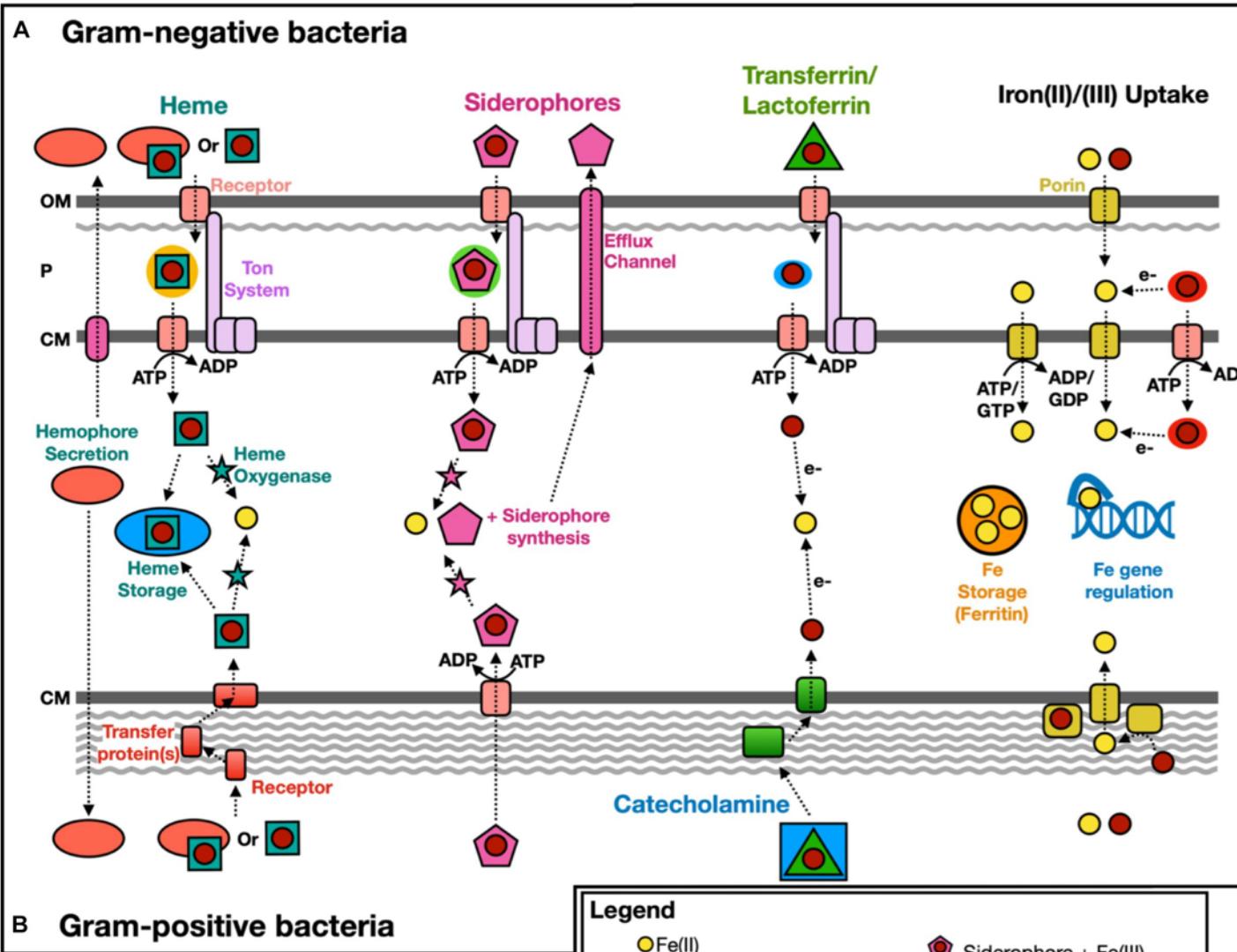
Category	Function	Protein Families
Iron acquisition	Iron(II)/(III) transport	Efe <u>UOB</u> ¹ , FbpABC ² , SfuABC ³ , YfuABC ⁴ , FeoAB(C) ⁵ , FutA1 ⁶ , FutA2 ⁶ , FutB ⁶ , FutC ⁶ , YfeABCD ⁷
	Heme oxygenase	ChuS ⁸ , ChuZ ⁹ , HemO ^{10,11} , PigA ^{10,11} , Hem <u>RSTUV</u> ¹² , HmoB ¹³ , HmuO ¹⁴ , HugZ ¹⁵ , HupZ ¹⁶ , Isd-LmHde ¹⁷ , IsdG ¹⁸ , IsdI ¹⁹ , MhuD ²⁰ , PhuS ²¹ (in PhuRSTUVW)
	Heme transport	HasRADE(B)E ²² , HmuR <u>STUV</u> ²² , HmuY ²³ , HmuY' ²³ , HutZ ²⁴ , Hxu <u>CBA</u> ²⁵ , IsdX1 ²⁶ , IsdX2 ²⁶ , Phu <u>RSTUVW</u> ²¹ , Rv0203 ²⁷
	Transferrin/Lactoferrin	Tbp <u>AB</u> (Lbp <u>AB</u>) ²⁸ , Sst <u>ABCD</u> ²⁹
	Siderophore synthesis	Acs <u>ABCDEF</u> ³⁰ , AmoA ³¹ , AngR ³² , Asb <u>ABCDEF</u> ³³ , DhbACEBF ³⁴ , entD- feP <u>A</u> -fes-entF-fepEC <u>GDB</u> -entCEBA-ybdA ³⁵ , IroD in Iro <u>NBCDE</u> ³⁶ , lucABCD ^{37,38} , IutA ^{37,38} , MbtIABCDEF ³⁵ , LbtA ³⁹ (in LbtUABC), PchABCDREFGH ³⁵ , PvdQAPMNOFEDJIHLGS ⁴⁰ , PvsABCDE ⁴¹ , VenB ⁴² , Vab genes in VabR-fur-vabGA-fur-VabCEBSFH-fur-fvtA-vabD ⁴³ , Vib genes in VibB-vibEC-vibA-vibH-viuP <u>DGC</u> -vibD and viuAB-vibF ⁴⁴⁻⁴⁶ , RhbABCDEF-rhrA-rhtA ⁴⁷
	Siderophore transport	BesA ⁴⁸ , Cbr <u>ABCD</u> ⁴⁹ , TonB-ExbB-ExbD ⁵⁰ , Fat <u>ABCD</u> ⁵¹ , FecI <u>RABCDE</u> ⁵² , FeuABC-yusV ⁵³ , Fhu <u>ACDB</u> ⁵⁴⁻⁵⁶ , FhuF ⁵⁴⁻⁵⁶ , Fpt <u>ABCX</u> ⁵⁷ , FpuA ⁵⁸ , FpuC ⁵⁸ , FpuD ⁵⁸ , FpvI <u>R-FpvA</u> -FpvGHJKCDEF ⁵⁹ , FvtA in VabR-fur-vabGA-fur-VabCEBSFH-fur-fvtA-vabD ⁴³ , Hat <u>CD</u> ³⁷ , Iro <u>NBCDE³⁶, LbtUABC³⁹, PirA⁶⁰, PiuA⁶⁰, Pvu<u>ABCDE</u>⁴¹, Viu genes in VibB-vibEC-vibA-vibH-viuP<u>DGC</u>-vibD and viuAB-vibF⁴⁴⁻⁴⁶, YifZ-yfhA⁶¹, YfiY⁶¹, YqjH⁶², ybdA and Fep genes in entD-feP<u>A</u>-fes-entF-fepEC<u>GDB</u>-entCEBA-ybdA³⁵</u>
Iron Gene regulation	Transcriptional regulation	DtxR ⁶³ , FecR (in FecI <u>RABCDE</u>) ⁵² , FeoC in FeoAB(C) ⁵ , Fur ⁶⁴ , IdeR ⁶⁵ , YqjI ⁶² , RhrA in RhbABCDEF-rhrA-rhtA ⁴⁷
Iron oxidation and reduction	Iron oxidation	Cyc1 ^{66,67} , Cyc2 ^{66,67,68} , FoxABC ⁶⁹ , FoxEYZ ⁷⁰ , Sulfocyanin ⁷¹ , PioA <u>BC</u> ⁷²
	Probable iron oxidation and possible iron reduction	MtoAB ⁷³ , Cyc2 (cluster 3)
	Dissimilatory iron reduction	CymA ⁷⁴ , Mtr <u>CAB</u> ⁷⁵ , OmcF ⁷⁶ , OmcS ⁷⁶ , OmcZ ⁷⁶ , FmnA-dmkA-fmnB-pplA-ndh2-eetAB-dmkB ⁷⁷ , DFE_0448-0451, DFE_0461-0465 ⁷⁸
Iron storage	Probable iron reduction	MtrCB, MtrAB, MtoAB-MtrC
	Iron storage	Bfr ⁷⁹ , DpsA ⁸⁰ , Ftn ⁸¹
Magnetosome-related	Magnetosome formation	MamABEKLMOPI ^{82,83} (Note: These genes are found in all known magnetotactic microorganisms, except for mamL which is found in magnetite-producing magnetotactic microorganisms ⁸¹)

Bolded and underlined HMMs are derived from Pfam or TIGRFAMs databases. Other HMMs were created by using select sequences. See **Supplementary Table S1** for more information, including the corresponding Pfam or TIGRFAMs families and the sequences used to create the HMMs. ¹Miethke et al., 2013, ²Adhikari et al., 1996, ³Angerer et al., 1990, ⁴Gong et al., 2001, ⁵Lau et al., 2016, ⁶Katoh et al., 2001, ⁷Bearden et al., 1998, ⁸Suits et al., 2006, ⁹Zhang et al., 2011, ¹⁰Friedman et al., 2003, ¹¹Friedman et al., 2004, ¹²Schneider et al., 2006, ¹³Park et al., 2012, ¹⁴Matsui et al., 2005, ¹⁵Hu et al., 2011, ¹⁶Sachla et al., 2016, ¹⁷Duong et al., 2014, ¹⁸Reniere et al., 2010, ¹⁹Skaar et al., 2004, ²⁰Graves et al., 2014, ²¹Ochsner et al., 2000, ²²Tong and Guo, 2009, ²³Wojtowicz et al., 2009, ²⁴Liu X. et al., 2012, ²⁵Morton et al., 2007, ²⁶Honsa et al., 2014, ²⁷Tullius et al., 2011, ²⁸Gray-Owen et al., 1995, ²⁹Morrissey et al., 2000, ³⁰Carroll and Moore, 2018, ³¹Barghouthi et al., 1991, ³²Wertheimer et al., 1999, ³³Oves-Costales et al., 2007, ³⁴May et al., 2001, ³⁵Crosa and Walsh, 2002, ³⁶Hantke et al., 2003, ³⁷Suzuki et al., 2006, ³⁸Martinez et al., 1994, ³⁹Cianciotto, 2015, ⁴⁰Lamont and Martin, 2003, ⁴¹Tanabe et al., 2003, ⁴²Tan et al., 2014, ⁴³Balado et al., 2008, ⁴⁴Wyckoff et al., 2001, ⁴⁵Keating et al., 2000, ⁴⁶Wyckoff et al., 1999, ⁴⁷Lynch et al., 2001, ⁴⁸Miethke et al., 2006, ⁴⁹Mahé et al., 1995, ⁵⁰Garcia-Herrero et al., 2007, ⁵¹Lemos et al., 2010, ⁵²Braun, 2003, ⁵³Peuckert et al., 2011, ⁵⁴Köster and Braun, 1989, ⁵⁵Coulton et al., 1987, ⁵⁶Braun et al., 2002, ⁵⁷Youard et al., 2011, ⁵⁸Dixon et al., 2012, ⁵⁹Brillet et al., 2012, ⁶⁰Moynie et al., 2017, ⁶¹Ollinger et al., 2006, ⁶²Wang et al., 2011, ⁶³Guedon and Helmann, 2003, ⁶⁴Escolar et al., 1998, ⁶⁵Rodriguez et al., 2002, ⁶⁶Castelle et al., 2008, ⁶⁷Barco et al., 2015, ⁶⁸McAllister et al., 2019, ⁶⁹Bathe and Norris, 2007, ⁷⁰Croal et al., 2007, ⁷¹Ilbert and Bonnefoy, 2013, ⁷²Liu J. et al., 2012, ⁷³Jiao and Newman, 2007, ⁷⁴Castelle et al., 2015, ⁷⁵Pitts et al., 2003, ⁷⁶Santos et al., 2015, ⁷⁷Light et al., 2018, ⁷⁸Deng et al., 2018, ⁷⁹Grossman et al., 1992, ⁸⁰Grant et al., 1998, ⁸¹Andrews, 1998, ⁸²Uebe and Schuler, 2016, ⁸³Kolinko et al., 2016.

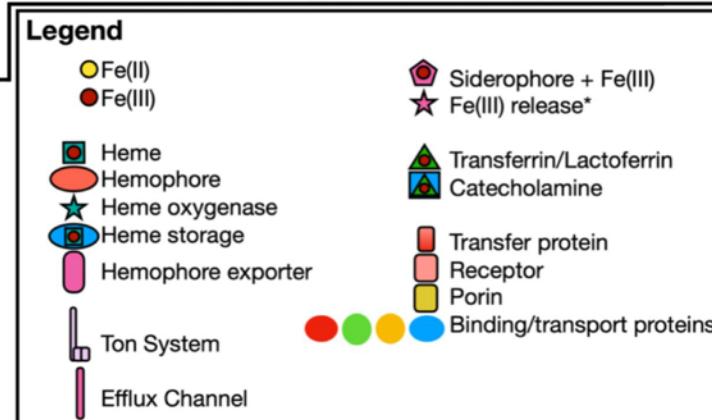
Bolded and underlined genes represent HMMs taken from Pfam and TIGRFAMs

Others were designed specifically for FeGenie

A Gram-negative bacteria

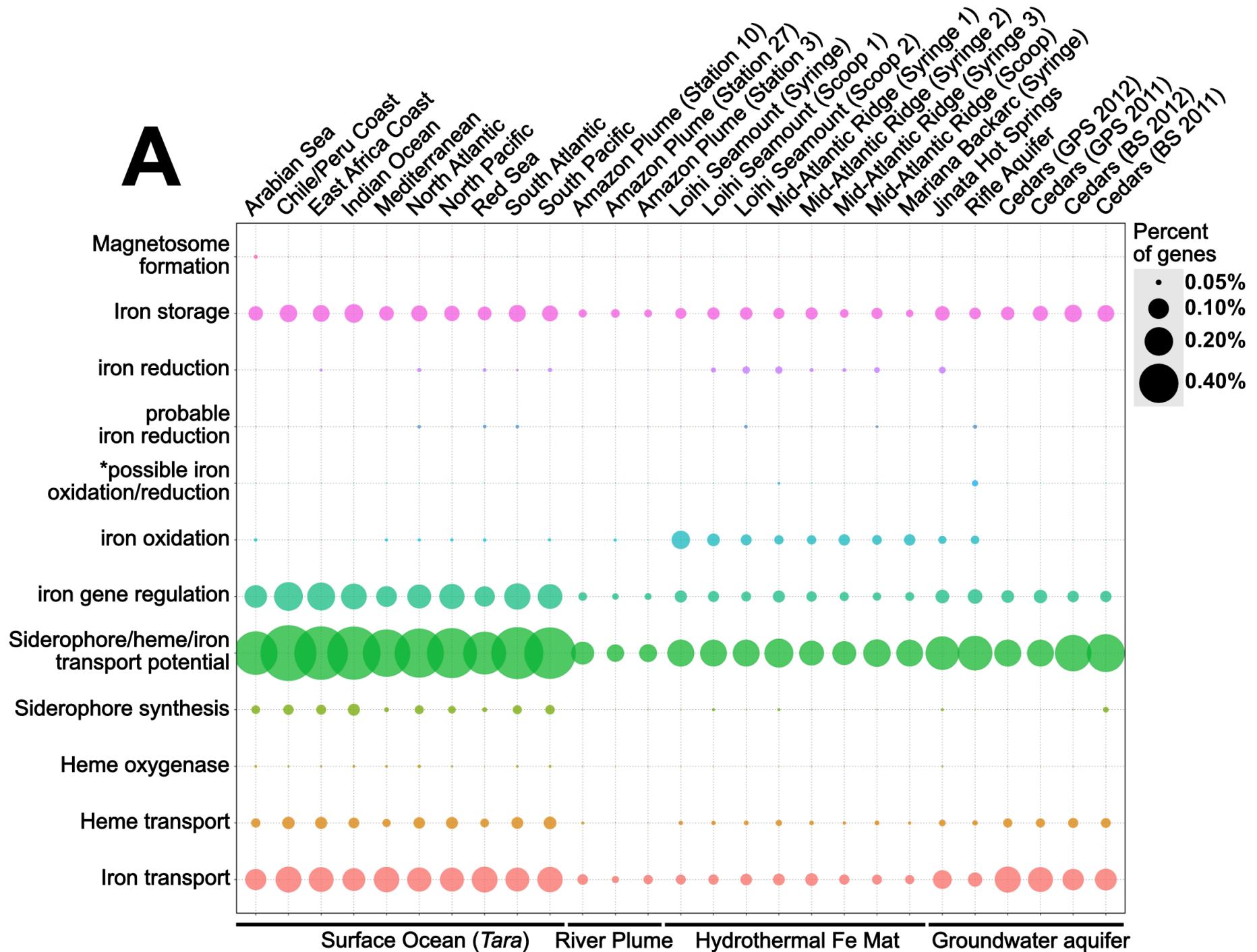


B Gram-positive bacteria

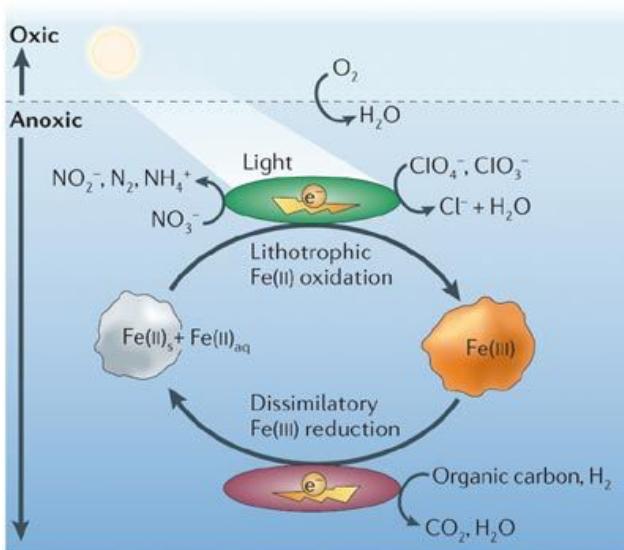


Summary of iron acquisition, efflux, and storage strategies in bacteria

Iron acquisition/scavenging strategies very important in iron-limiting environments like the open ocean

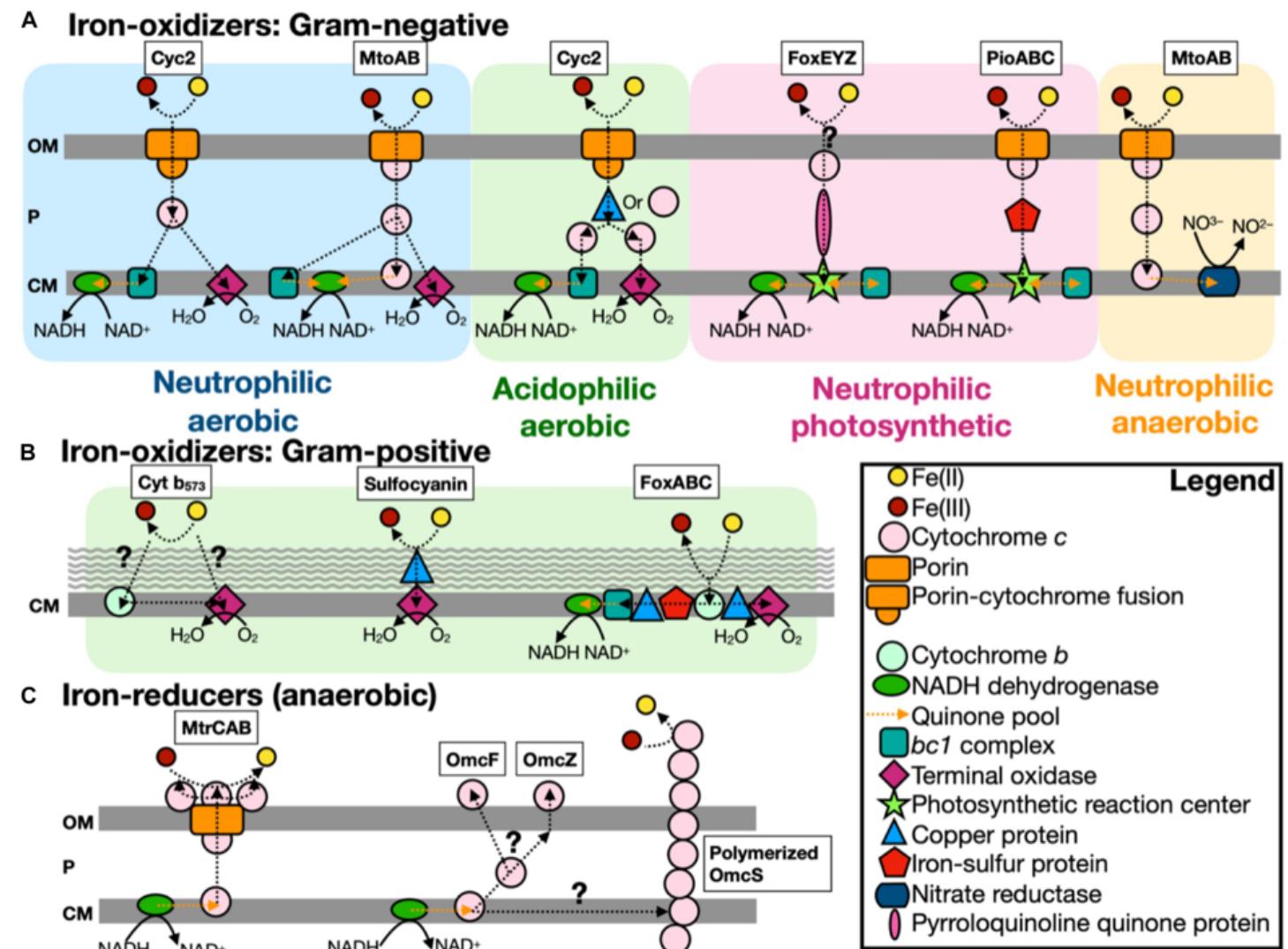
A

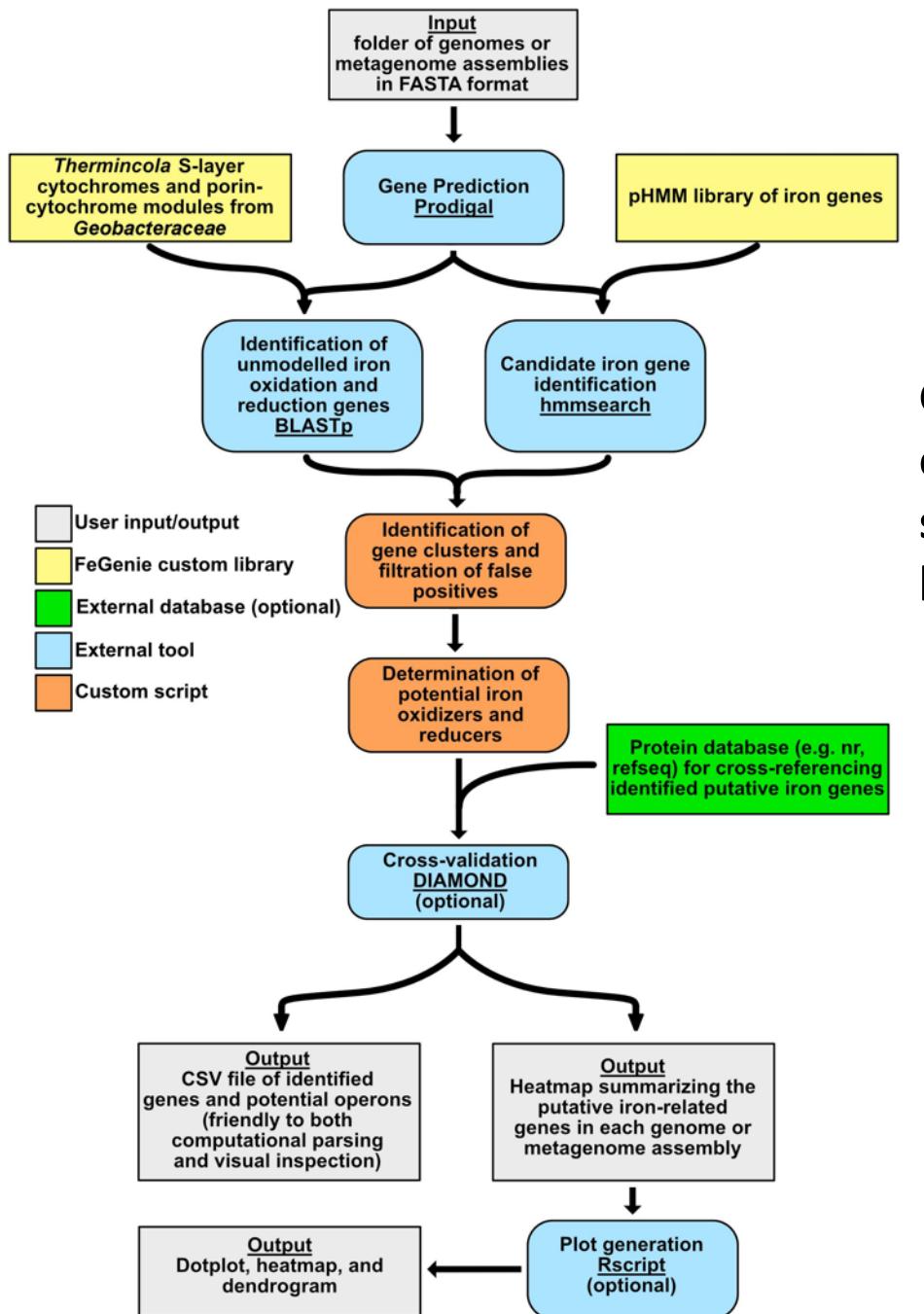
Iron oxidation and reduction – porin-cytochrome modules are key players in iron-related redox reactions in gram-negative bacteria



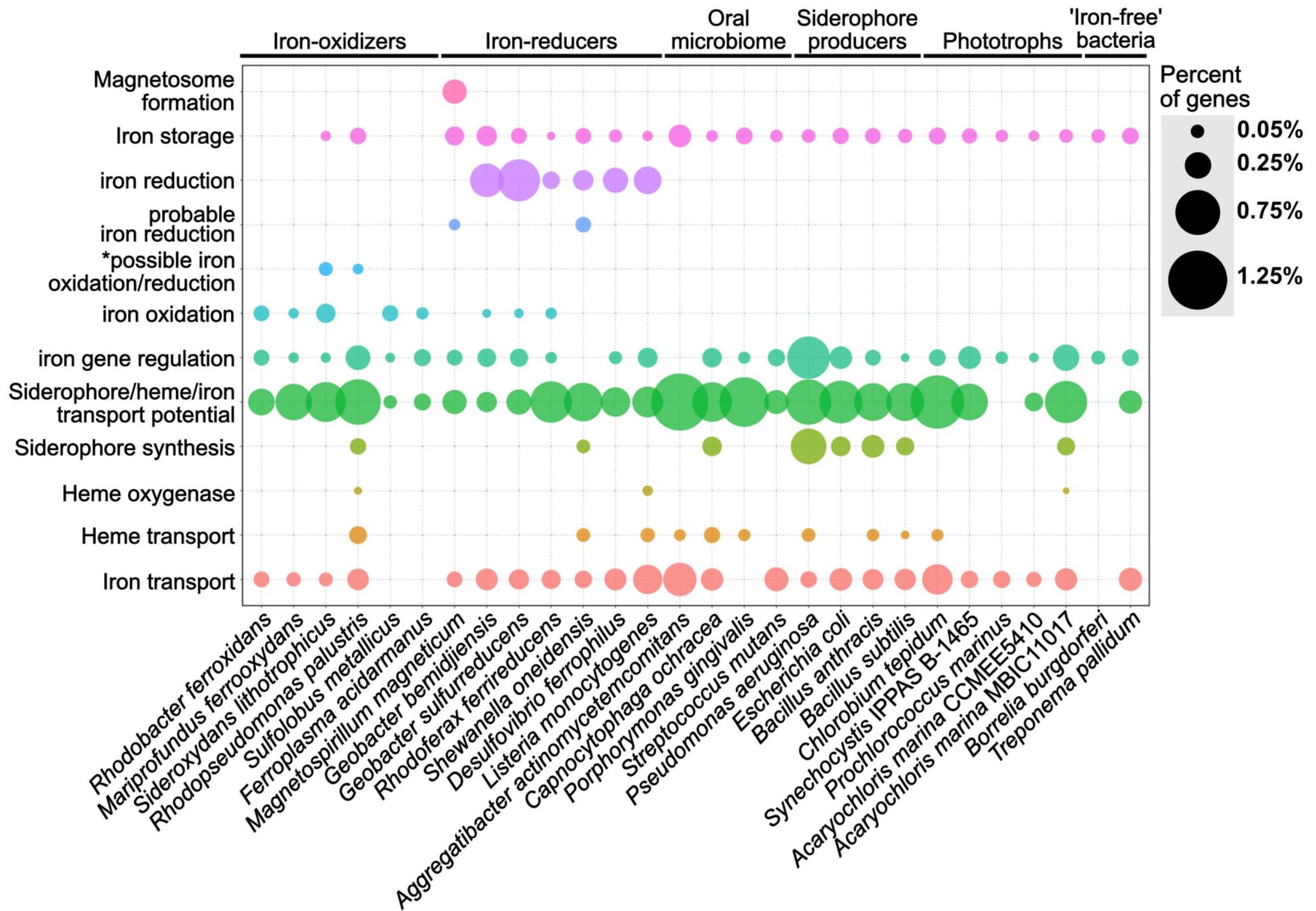
Copyright © 2006 Nature Publishing Group
Nature Reviews | Microbiology

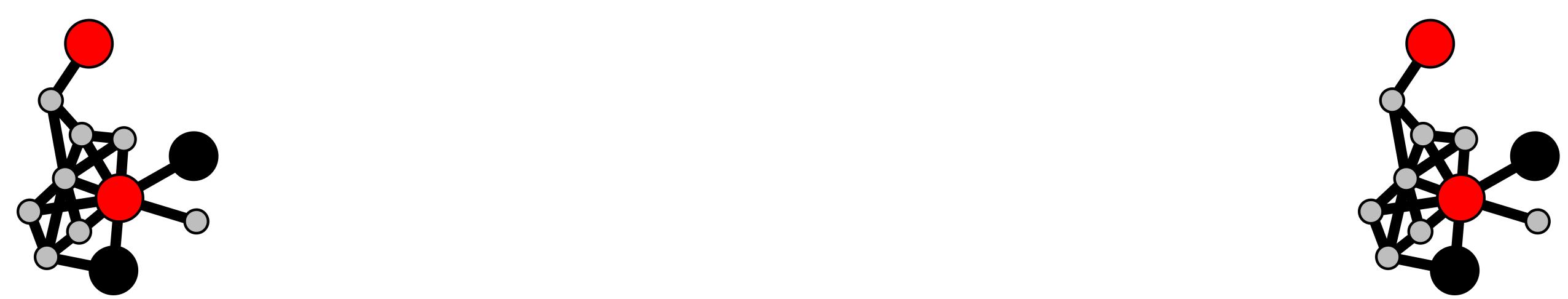
S-layer and cytoplasmic membrane associated proteins are used by gram-positive bacteria and Archaea





General workflow involves a combination of custom scripts and other bioinformatics software for ORF-prediction and HMM-based homology searching.





Jupyter Binder tutorial

