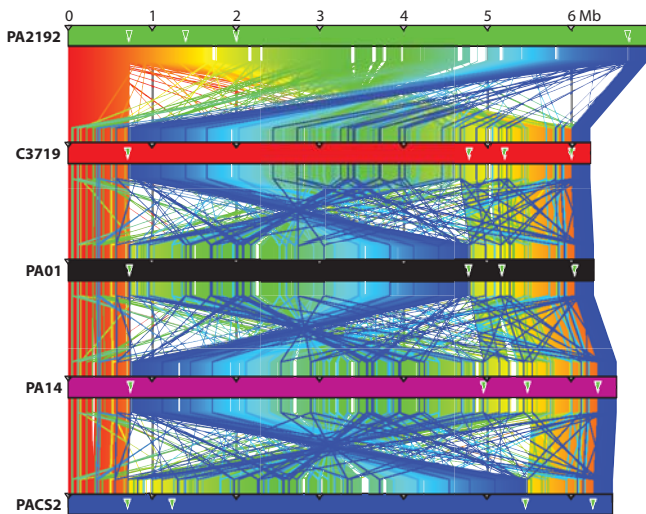
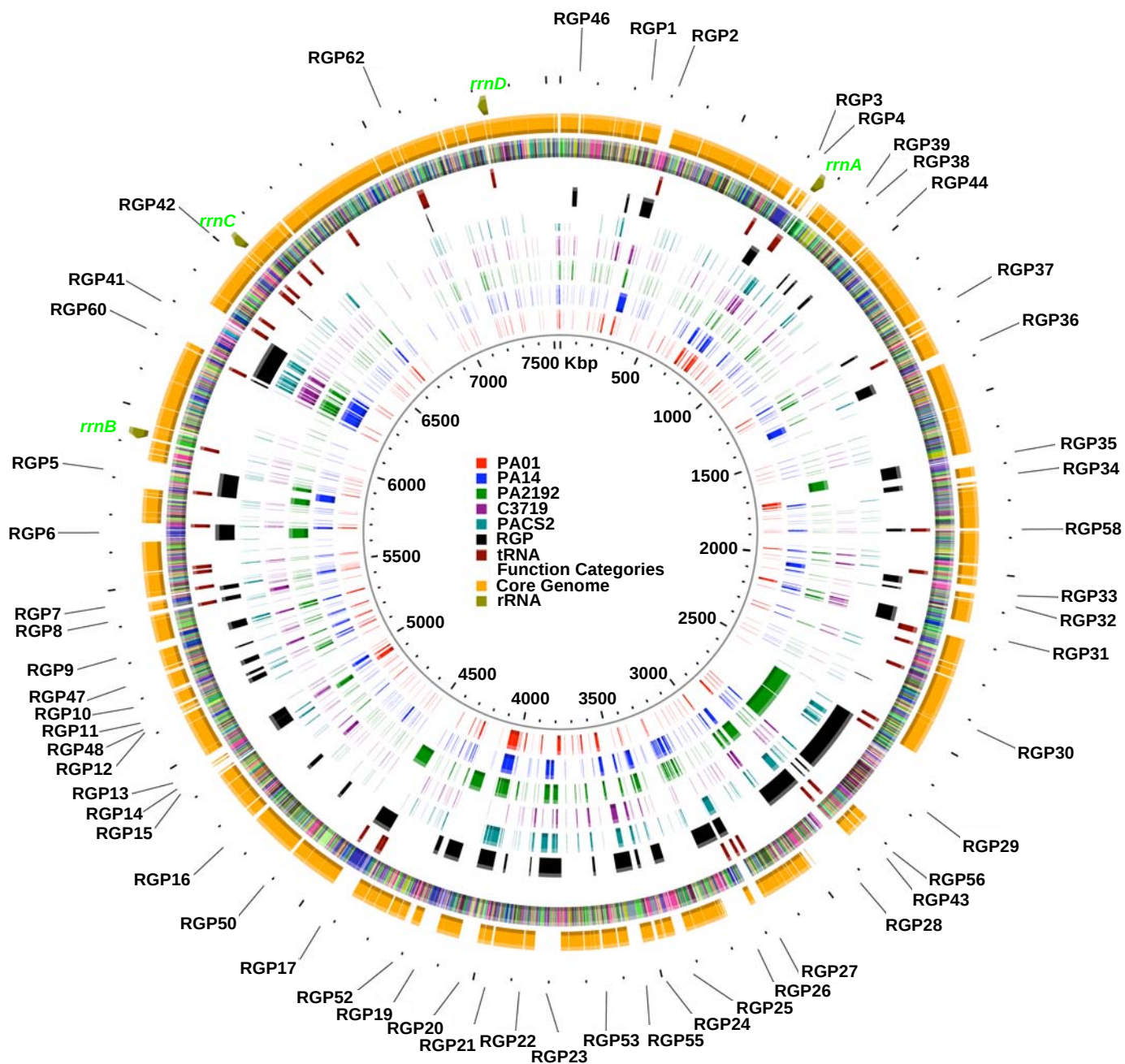
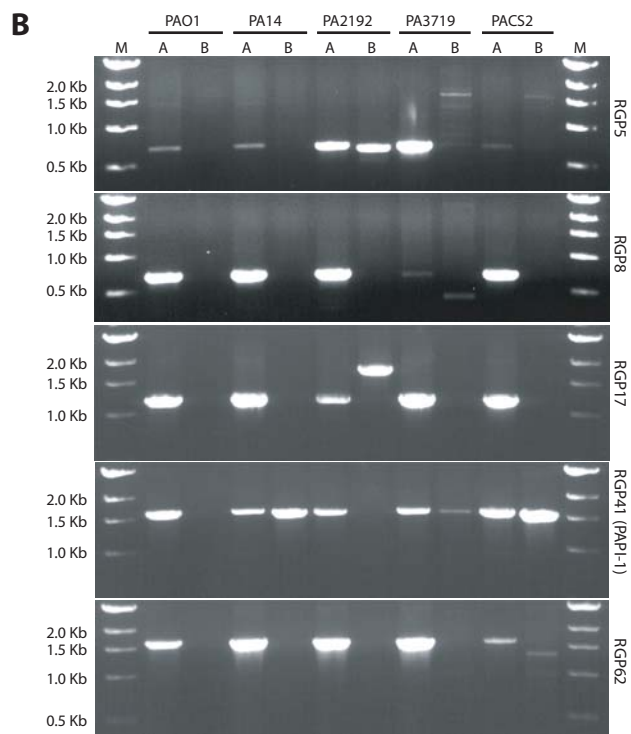
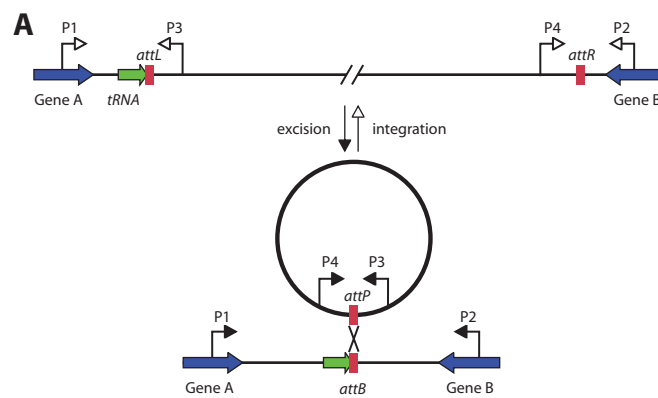
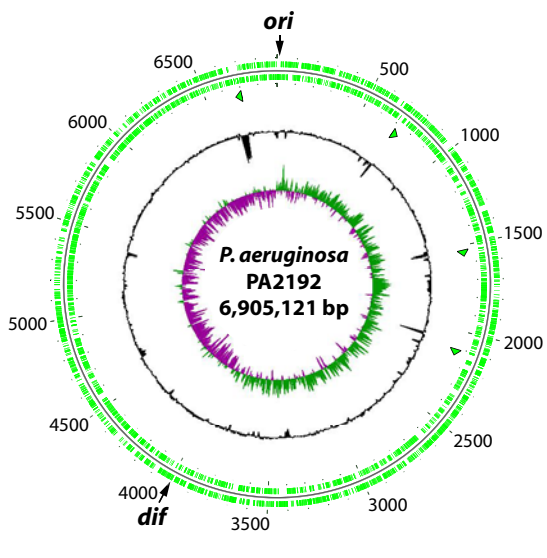
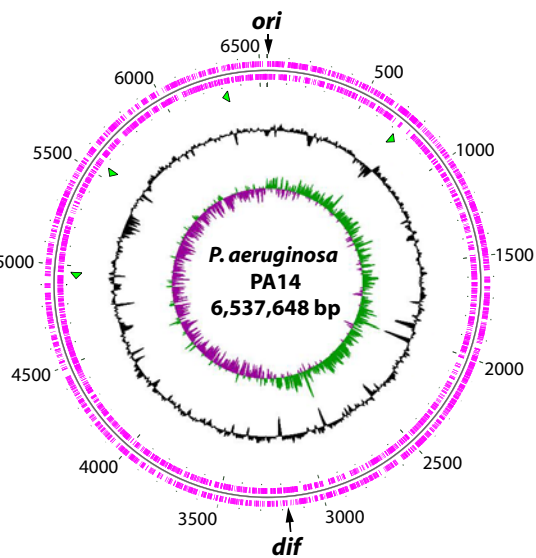
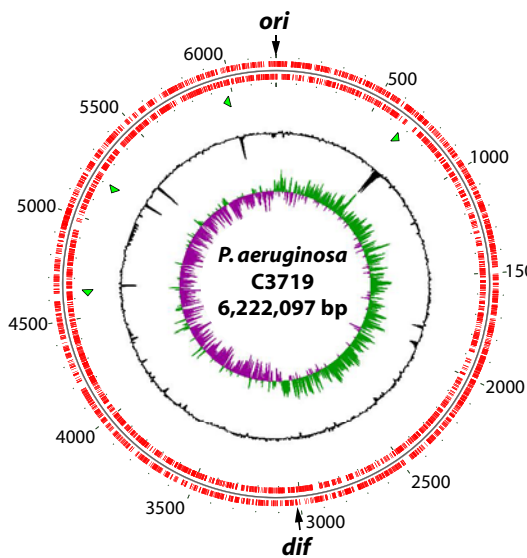
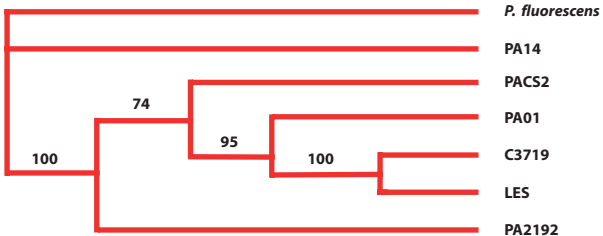


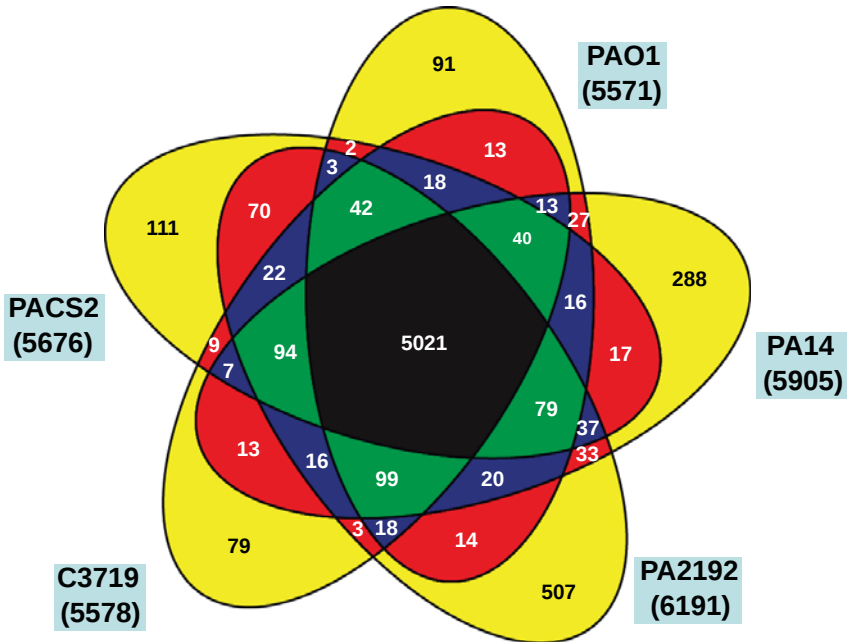


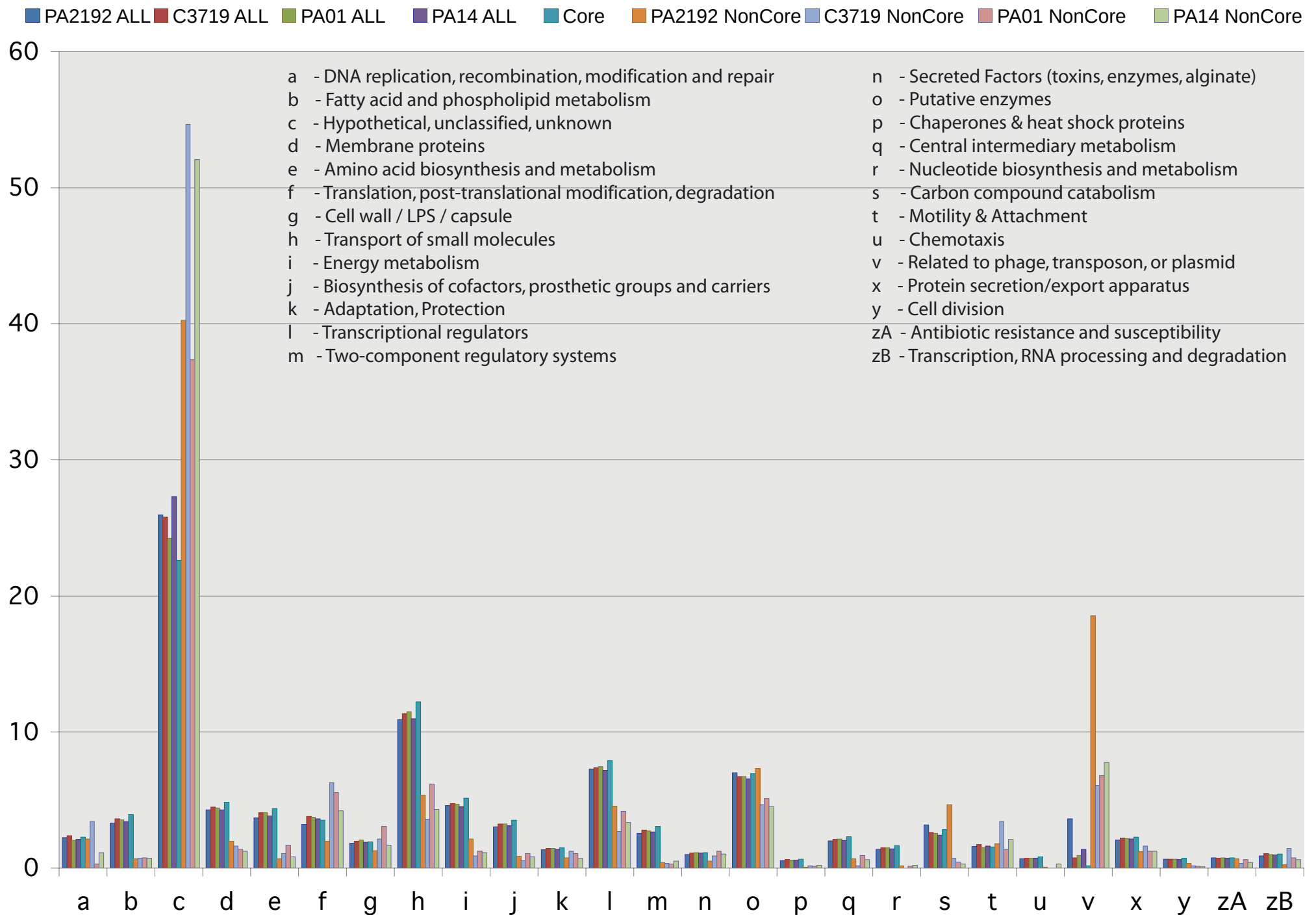


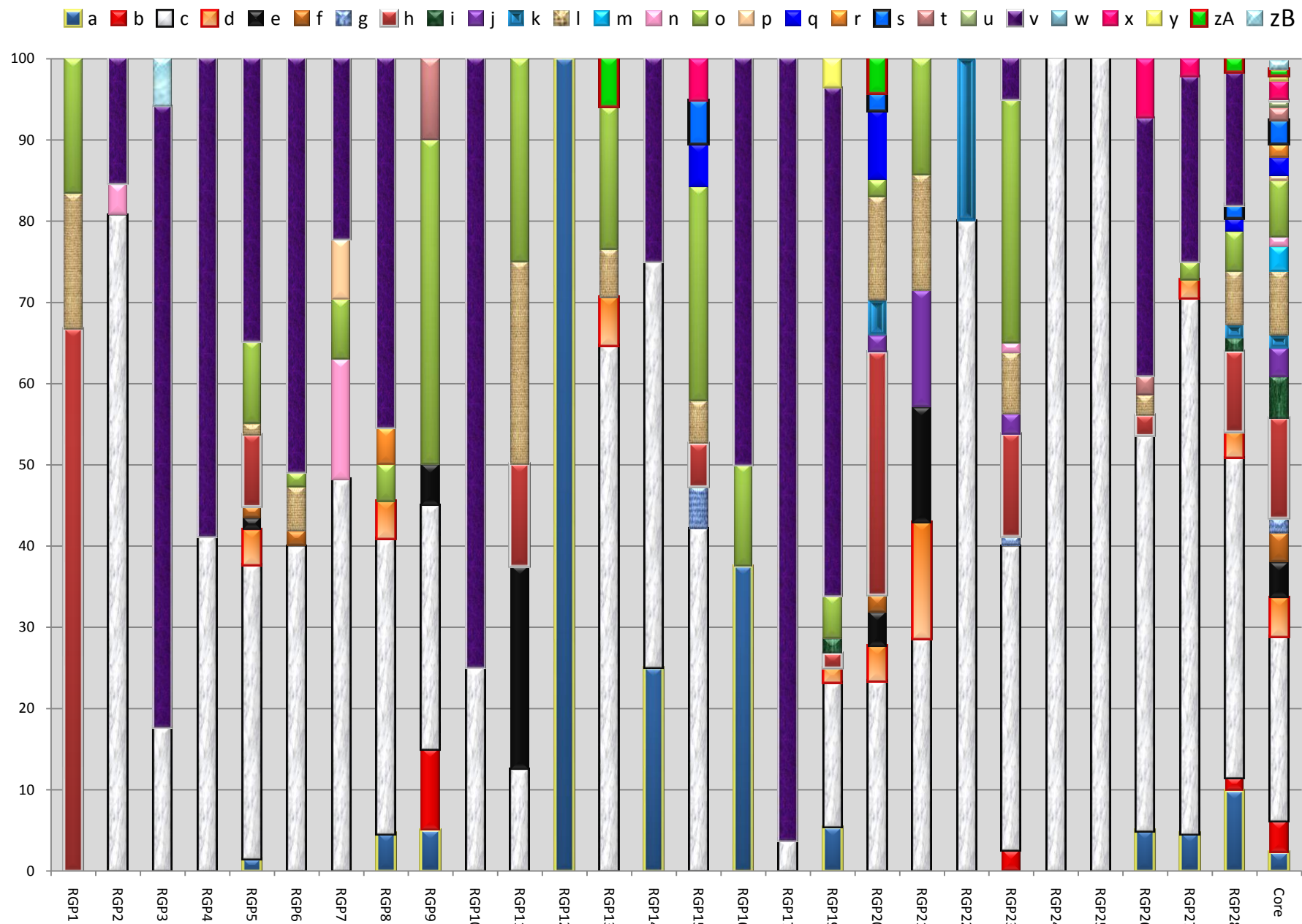


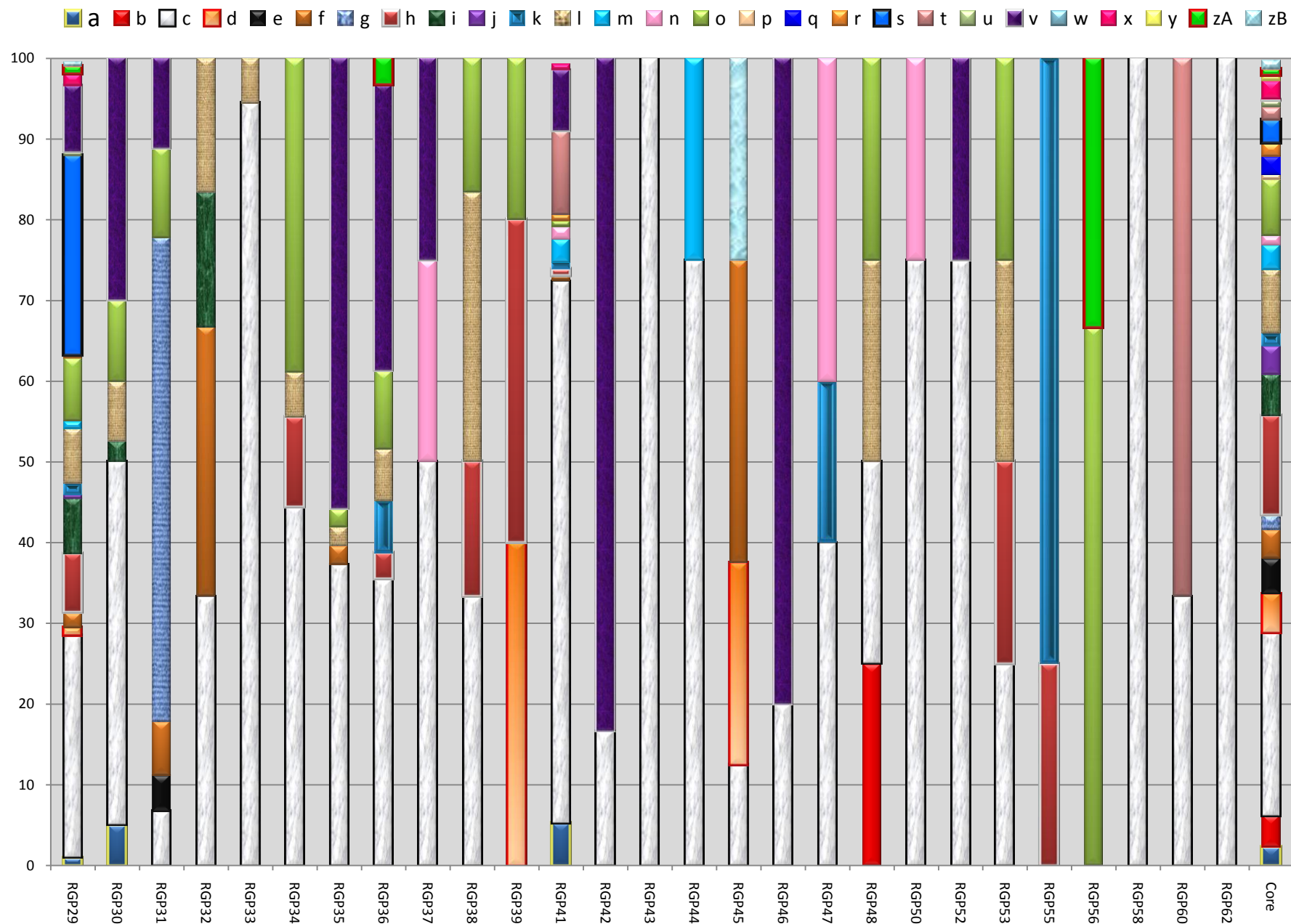


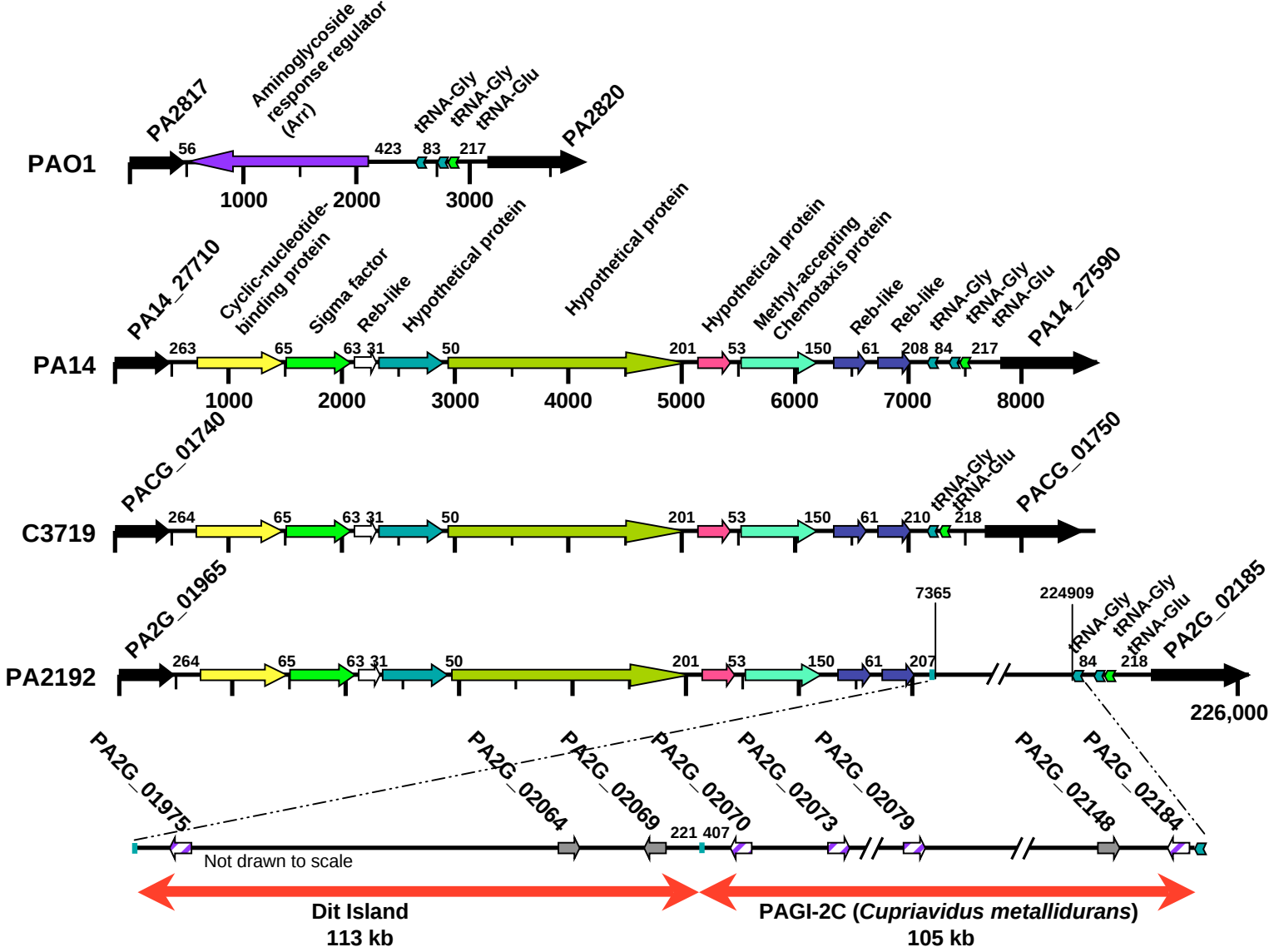












PAO1

GCGGCACTGAATCGCCTCGCGCCGGA<<Insertion1>>ACGAGCAGCCTGGCGCGCAGAAACGAAA

C3719

GCGGCACTAAGCCCGCCACGGCCACC<<Insertion1>>CAGGGCGGTGCGGATCGCAGAAACGAAA

PA14
PACS2

GCGGCACTGAGCCCGCCACGGCCACC<<Insertion1>>CAGGGCGGTGCGGATCGCAGAAACGAAA

GCGGCACTGAGCCCGCCACGGCCACC<<Insertion1>>CAGGGCGGTGCGGATCGCAGAAACGAAA

AAAGGGCCCGAAGGCCCTTTTCTCTTCGACCCAGGCCGAAACCTGGATCATGTTTGGAGCGGGAAACG

PA2192

GGACAAACCTTTTCCGCGTATCTCCTTGATTCAACAAAG<<DiTIsland>>TCCCCAGGGCTAAAC

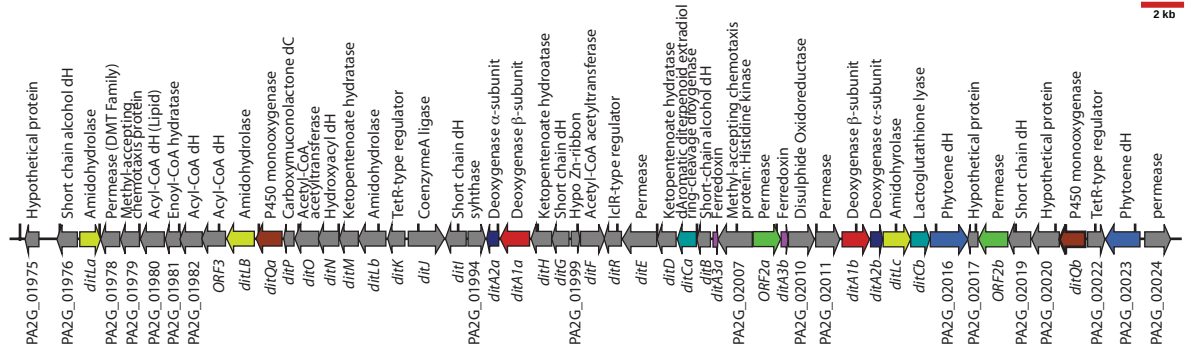
CCTGTTGAATCAAGAATCTATAGACTTTGGTGGACGTCTATAGATGAAGAA

TCATGCACTAGAGCGCTAAG<<PAGI_2>>TTCCGTGGCTTTGGCTGGAATGAAAAAGGGGCCGAAG

CCCCTAATTTCAATGACTTACAGACGTCAGTGGAAGTCTGTAGATCAAAAC

TTGGAGCGGGAAACGAG

RGP29-Dit Island

*B. xenovorans*

Bxe_C0621

Bxe_C0602

Bxe_C0603

Bxe_C0604

Bxe_C0605

Bxe_C0606

Bxe_C0599

Bxe_C0598

Bxe_C0597

Bxe_C0596

Bxe_C0595

Bxe_C0594

Bxe_C0593

Bxe_C0592

Bxe_C0591

Bxe_C0590

Bxe_C0587

Bxe_C0586

Bxe_C0585

Bxe_C0583

Bxe_C0582

Bxe_C0581

Bxe_C0608

Bxe_C0640

Bxe_C0639

Bxe_C0638

Bxe_C0580

Bxe_C0579

Bxe_C0606

Bxe_B2541

Bxe_C0630

Bxe_C0631

Bxe_C0636

Bxe_C0637

B. abietaniphila

ORF3

ORF4

ditQ

ditP

ditO

ditN

ditM

ditL

ditK

ditJ

ditI

ditA2

ditA1

ditH

ditG

ditF

ditE

ditD

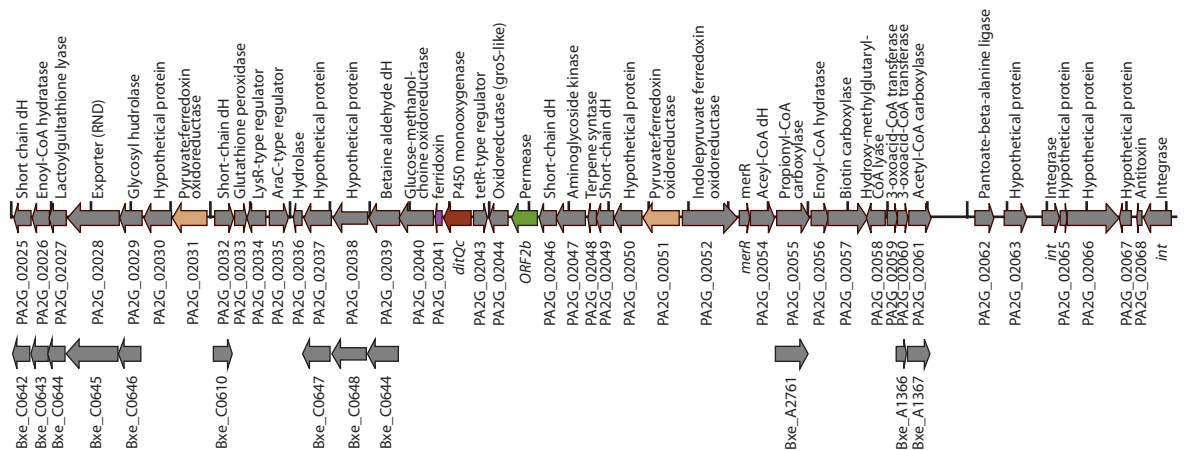
ditC

ditB

ditA3

ORF2

RGP29-Dit Island

*B. xenovorans*

Bxe_C0642

Bxe_C0643

Bxe_C0644

Bxe_C0645

Bxe_C0646

Bxe_C0610

Bxe_C0647

Bxe_C0648

Bxe_C0644

Bxe_A2761

Bxe_A1366

Bxe_A1367

Table 1. Genome properties of *P. aeruginosa* strains and their genomes

Properties	PA2192	C3719	PAO1	PA14	PACS2
Origin	Cystic Fibrosis	Cystic Fibrosis	Burn wound	Burn victim	Cystic Fibrosis
O-antigen biosynthetic genes *	O1 [†]	O3 [‡]	O5	10	O1
Genome size (bp)	6,905,121	6,222,097	6,264,404	6,537,648	6,492,423
ORFs	6,191	5,578	5,571	5,905	5,676
tRNAs	46 [§]	40 [§]	64	63	65 *****
rRNAs	4	4	4	4	4
GC content (%)	66.2	66.5	66.6	66.3	66
Leading strand - No. of Genes (%)	3386 (55)	3043 (55)	3127 (56)	3212 (54)	3088 (54)
Leading strand - % essential genes	65	66	67	66	58
Total RGPs	35	31	32	37	27
No of Unique RGPs (RGP No.)	6 (RGP12, 17, 19, 30, 35, 42)	3 (RGP8, 16, 46)	1 (RGP34)	7 (RGP10, 11, 26, 32, 33, 38, 50)	1(RGP62)
Reference	This Study	This Study	1, 2	3, 4	

* As shown by Raymond *et al.* (5)

[†] PA2G_02562 is a pseudogene (four nucleotide insertions creating a premature termination codon)

[‡] *PACG_02083* is a pseudogene (a single nucleotide insertion that results in a premature stop codon). The cluster also contains a 1137-bp insertion sequence.

[§] Based on the information from the draft genomes PA2192 and PAC3719

References

1. Holloway BW (1955) *J Gen Microbiol* 13:572–581.
2. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrenner P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, *et al.* (2000) *Nature* 406:959-964.
3. Schroth MN, Cho JJ, Green SK, Kominos SD (1977) in *Pseudomonas aeruginosa: Ecological Aspects and Patient Colonization*, ed. Young VM (Raven, New York).
4. Lee DG, Urbach JM, Wu G, Liberati NT, Feinbaum RL, Miyata S, Diggins LT, He J, Saucier M, Deziel E, *et al.* (2006) *Genome Biol* 7:R90.
5. Raymond CK, Sims EH, Kas A, Spencer DH, Kuttyavin TV, Ivey RG, Zhou Y, Kaul R, Clendenning JB, Olson MV (2002) *J Bacteriol* 184:3614-3622.

Table 2: List of phylogtically useful genes (PUG)

PA Number	Gene	Alt gene	Functional Annotation	Functional Category	Parsimony-Informative Sites
PA2002	<i>atoE</i>		conserved hypothetical protein	b	4
PA4372			hypothetical protein	b	4
PA5386			probable 3-hydroxyacyl-CoA dehydrogenase	b	4
PA1164			conserved hypothetical protein	c	4
PA1421	<i>gbuA</i>	<i>speB2</i>	guanidinobutyrase	e	4
PA3792	<i>leuA</i>		2-isopropylmalate synthase	e	4
PA1475	<i>ccmA</i>	<i>helA/cycV</i>	heme exporter protein CcmA	h	4
PA2408		<i>znuC</i>	probable ATP-binding component of ABC transporter	h	4
PA2711	<i>potF4</i>		probable periplasmic spermidine/putrescine-binding protein	h	4
PA0130		<i>mmsA</i>	probable aldehyde dehydrogenase	i	4
PA0773	<i>pdxJ</i>		pyridoxal phosphate biosynthetic protein PdxJ	j	4
PA3625	<i>surE</i>		survival protein SurE	k	4
PA3548	<i>algI</i>		alginate o-acetyltransferase AlgI	n	4
PA4956	<i>rhdA</i>		thiosulfate:cyanide sulfurtransferase	q	4
PA1524	<i>xdhA</i>		xanthine dehydrogenase	r	4
PA4720	<i>trmA</i>		tRNA (uracil-5-)-methyltransferase	zB	4
PA4693	<i>pssA</i>		phosphatidylserine synthase	b	5
PA3897			hypothetical protein	d	5
PA1162	<i>dapE</i>		succinyl-diaminopimelate desuccinylase	e	5
PA3666	<i>dapD</i>		tetrahydrodipicolinate succinylase	e	5
PA3814	<i>iscS</i>		L-cysteine desulfurase (pyridoxal phosphate-dependent)	e	5
PA1651			probable transporter	h	5
PA2341	<i>mtlK</i>		probable ATP-binding component of ABC maltose/mannitol transporter	h	5
PA5095		<i>proW</i>	probable permease of ABC transporter	h	5
PA4438		<i>yhcM</i>	conserved hypothetical protein	i	5
PA2826			probable glutathione peroxidase	k	5
PA2843			probable aldolase	o	5
PA1687	<i>speE</i>		spermidine synthase	q	5
PA5260	<i>hemC</i>	<i>popE</i>	porphobilinogen deaminase	q	5
PA0441	<i>dht</i>	<i>hyuA</i>	dihydropyrimidinase	r	5
PA0590	<i>apaH</i>		bis(5'-nucleosyl)-tetrphosphatase	r	5
PA3751	<i>purT</i>		phosphoribosylglycinamide formyltransferase 2	r	5
PA5387			conserved hypothetical protein	c	6
PA0229	<i>pcaT</i>	<i>kgtP</i>	dicarboxylic acid transporter PcaT	d	6
PA0392		<i>yggT</i>	conserved hypothetical protein	d	6
PA0440	<i>dypA</i>		probable oxidoreductase	e	6
PA3108	<i>purF</i>		amidophosphoribosyltransferase	e	6
PA5373	<i>betB</i>		betaine aldehyde dehydrogenase	e	6
PA5451	<i>wzm</i>		membrane subunit of A-band LPS efflux transporter	g	6
PA0301	<i>spuE</i>	<i>potF3</i>	polyamine transport protein	h	6

PA1207	<i>kefB</i>	<i>trkB/kefC</i>	glutathione-regulated potassium-efflux system protein KefB	h	6
PA3865		<i>hisJ</i>	probable amino acid binding protein	h	6
PA5479	<i>gltP</i>		proton-glutamate symporter	h	6
PA2776	<i>ordL</i>		conserved hypothetical protein	i	6
PA0182		<i>yjgI</i>	probable short-chain dehydrogenase	j	6
PA4606	<i>cstA</i>		conserved hypothetical protein	k	6
PA3970	<i>amn</i>		AMP nucleosidase	r	6
PA0211	<i>mdcD</i>		malonate decarboxylase beta subunit	s	6
PA5028			conserved hypothetical protein	y	6
PA2545	<i>xthA</i>		exodeoxyribonuclease III	a	7
PA1640			conserved hypothetical protein	b	7
PA2001	<i>atoB</i>		acetyl-CoA acetyltransferase	b	7
PA0240	<i>opdF</i>		probable porin	d	7
PA2751		<i>yfkH</i>	conserved hypothetical protein	d	7
PA5172	<i>arcB</i>		ornithine carbamoyltransferase, catabolic	e	7
PA2278	<i>arsB</i>		ArsB protein	h	7
PA3920		<i>yvgX</i>	probable metal transporting P-type ATPase	h	7
PA1318	<i>cyoB</i>		cytochrome o ubiquinol oxidase subunit I	i	7
PA5192	<i>pckA</i>		phosphoenolpyruvate carboxykinase	i	7
PA0500	<i>bioB</i>		biotin synthase	j	7
PA2611	<i>cysG</i>		siroheme synthase	j	7
PA0181		<i>yeaT</i>	probable transcriptional regulator	l	7
PA4367			conserved hypothetical protein	m	7
PA2823			conserved hypothetical protein	o	7
PA4974	<i>opmH</i>		probable outer membrane protein precursor	x	7
PA1528	<i>zipA</i>		cell division protein ZipA	y	7
PA3617	<i>recA</i>		RecA protein	a	8
PA3164			still frameshift 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE prephenate dehydrogenase	e	8
PA4696	<i>ilvI</i>		acetolactate synthase large subunit	e	8
PA3047	<i>dacB</i>	<i>pbp</i>	probable D-alanyl-D-alanine carboxypeptidase	g	8
PA3186	<i>oprB</i>		Glucose/carbohydrate outer membrane porin OprB precursor	h	8
PA2623	<i>icdA</i>		isocitrate dehydrogenase	i	8
PA5034	<i>hemE</i>		uroporphyrinogen decarboxylase	j	8
PA4184			probable transcriptional regulator	l	8
PA5550	<i>glmR</i>		GlmR transcriptional regulator	l	8
PA3828		<i>yjgP</i>	conserved hypothetical protein	d	9
PA4370	<i>icmP</i>		Insulin-cleaving metalloproteinase outer membrane protein precursor	d	9
PA2249	<i>bkdB</i>		branched-chain alpha-keto acid dehydrogenase (lipoamide component)	e	9
PA2309			hypothetical protein	h	9
PA2250	<i>lpdV</i>		lipoamide dehydrogenase-Val	i	9
PA4470	<i>fumC1</i>		fumarate hydratase	i	9
PA2903	<i>cobJ</i>	<i>cbiH</i>	precorrin-3 methylase CobJ	j	9

PA5237		<i>yigC</i>	conserved hypothetical protein	j	9
PA4398			probable two-component sensor	m	9
PA2494	<i>mexF</i>		Resistance-Nodulation-Cell Division (RND) multidrug efflux transporter MexF	zA	9
PA2840	<i>deaD</i>		probable ATP-dependent RNA helicase	zB	9
PA0507		<i>fad?</i>	probable acyl-CoA dehydrogenase	b	10
PA5546			conserved hypothetical protein	b	10
PA3424			hypothetical protein	e	10
PA3624	<i>pcm</i>		L-isoaspartate protein carboxymethyltransferase type II	f	10
PA3824	<i>queA</i>		S-adenosylmethionine:trna ribosyltransferase-isomerase	f	10
PA0185	<i>atsB</i>		probable permease of ABC transporter	h	10
PA0281	<i>cysW</i>		sulfate transport protein CysW	h	10
PA3677			probable Resistance-Nodulation-Cell Division (RND) efflux membrane fusion protein precursor	h	10
PA2867			probable chemotaxis transducer	u	10
PA1730			conserved hypothetical protein	c	11
PA1566			conserved hypothetical protein	e	11
PA4548	<i>dadA</i>	<i>yfiT</i>	probable D-amino acid oxidase	e	11
PA4344			probable hydrolase	f	11
PA0464	<i>creC</i>	<i>phoM</i>	two-component sensor CreC	m	11
PA2346			conserved hypothetical protein	o	11
PA3769	<i>guaA</i>		GMP synthase	r	11
PA4701			conserved hypothetical protein	c	12
PA1638		<i>yneH</i>	conserved hypothetical protein	e	12
PA5235	<i>glpT</i>		glycerol-3-phosphate transporter	h	12
PA4733	<i>acsB</i>		acetyl-coenzyme A synthetase	q	12
PA3950			probable ATP-dependent RNA helicase	zB	12
PA5443	<i>uvrD</i>	<i>recL/rad/pdeB/mutU</i>	DNA helicase II	a	13
PA0454		<i>yccS</i>	conserved hypothetical protein	d	13
PA3257	<i>prc</i>	<i>tsp</i>	periplasmic tail-specific protease	f	13
PA1946	<i>rbsB</i>		binding protein component precursor of ABC ribose transporter	h	13
PA3896			probable 2-hydroxyacid dehydrogenase	i	13
PA0757		<i>tctE</i>	probable two-component sensor	m	13
PA3710			probable GMC-type oxidoreductase	o	13
PA1101	<i>fliF</i>		Flagella M-ring outer membrane protein precursor	t	13
PA1886	<i>polB</i>	<i>dinA</i>	DNA polymerase II	a	14
PA1689			conserved hypothetical protein	d	14
PA1272	<i>cobO</i>	<i>cobA/btuR</i>	cob(I)alamin adenosyltransferase	j	14
PA3821	<i>secD</i>		secretion protein SecD	x	14
PA4483	<i>gatA</i>		Glu-tRNA(Gln) amidotransferase subunit A	f	15
PA3759			probable aminotransferase	g	15
PA3028	<i>moeA2</i>		molybdenum cofactor biosynthesis protein A2	j	15
PA4692			conserved hypothetical protein	o	15
PA3271		<i>flhS</i>	probable two-component sensor	m	16
PA0559		<i>yhiN</i>	conserved hypothetical protein	d	17

PA5529			probable sodium/proton antiporter	h	17
PA1585	<i>sucA</i>		2-oxoglutarate dehydrogenase (E1 subunit)	i	17
PA4843			probable two-component response regulator	m	17
PA0183	<i>atsA</i>		arylsulfatase	q	17
PA1532	<i>dnaX</i>		DNA polymerase subunits gamma and tau	a	18
PA4115		<i>ygdH</i>	conserved hypothetical protein	o	21
PA3893			conserved hypothetical protein	x	22
PA2018	<i>mexH</i>	<i>amrB/mexY</i>	Resistance-Nodulation-Cell Division (RND) multidrug efflux transporter	zA	23
PA5447	<i>wbpZ</i>		glycosyltransferase WbpZ	g	25
PA3297	<i>hrpA</i>		probable ATP-dependent helicase	a	26
PA4686			hypothetical protein	c	26
PA4772			probable ferredoxin	i	27
PA4513	<i>piuB</i>		probable oxidoreductase	o	29
PA3961	<i>hrpB</i>		probable ATP-dependent helicase	a	30
PA3258			hypothetical protein	o	31
PA5556	<i>atpA</i>	<i>papA/uncA</i>	ATP synthase alpha chain	i	55
PA2393			probable dipeptidase precursor	q	118
PA2403			hypothetical protein	d	138
PA0595	<i>ostA</i>	<i>imp</i>	organic solvent tolerance protein OstA precursor	k	376

Table 4. Summary of BLAST results for RGP sequences

	Core	Accessory
<i>Acidovorax</i>	4	1
<i>Comamonas</i>	0	0
<i>Delftia</i>	2	0
<i>Bordetella</i>	168	21
<i>Burkholderia</i>	870	142
<i>Cupriavidus</i>	403	245
<i>Chromobacterium</i>	124	42
<i>Nitrosomonas</i>	56	1
<i>Azotobacter</i>	2802	188
<i>P. abietaniphila</i>	7	17
<i>P. entomophila</i>	0	0
<i>P. fluorescens</i>	5994	343
<i>P. mendocina</i>	24	6
<i>P. putida</i>	2944	208
<i>P. syringae</i>	2208	310
Other Pseudomonads	378	49
<i>E. coli</i>	42	14
<i>Shewanella</i>	106	20
γ proteobacterium	0	0
<i>Xanthomonas</i>	136	35
<i>Sinorhizobium</i>	33	8
Bacteriophage	22	68
Eukaryotes	8	1
Other	2558	448
<i>Pseudomonas</i> Total	11555	933
Others Total1	4429	640
Others Total2	2905	594
Others Total	7334	1234
Remaining genes (No Hits or Paralogs)	1195	687
Number of Predicted Paralogs	388	90
Genome Size (ORFs)	23251	23251
<i>Pseudomonas</i> Percent of Total	0.50	0.04
Others Percent	0.32	0.05
CORE or ACCESSORY Size	20084	2854
<i>Pseudomonas</i> Percent of CORE or ACCESSORY	0.58	0.33
Others Percent of CORE or ACCESSORY	0.37	0.43
Corrected Percentages (Counting Remaining as Non- <i>Pseudomonas</i> Hits)		
<i>Pseudomonas</i> Percentage	57.53	32.69
Others (Non- <i>Pseudomonas</i>) Percentage	40.53	64.16

Table 5. Primers used to detect the excision and presence of a circular intermediates of the tRNA-linked RGPs

RGP	Primer sequence (5' to3')
RGP5	P1: CCGCTGAACTTCTGGGAATA P2: GCGGAATTCCACACCATC P3: GCCATCCGACTTGACGTAGT P4: ATGTACGGGATGATGGTGGT
RGP8	P1: CGTCTGTGGGACAAGGAAAC P2: ACTCGGCACAAACCGATAAC P3: ACGCCGAAGGAAAGGAATAC P4: CTACGAGCACGAAAACGTGA
RGP17	P1: GTAGGCCTGGGAGAGAAAGC P2: CCATCTGCTTGAGAGCCTTG P3: GATACGTCCCAAGCGACACT P4: GAATCAAGGCTCTGGAGTCG
RGP41	P1: GACAAGACCAGCCACAACCT P2: GTGGTGATGACCTCCAACCT P3: CGAGCACAGAAATGTCCTGA P4: AGCTACATCGAGGCCGACTA
RGP62	P1: CCAGCAGGAGATGGACAAGT P2: GATCGACACCTCGATCCTGT P3: CAGATCTGGACGGGTGATCT P4: CTGTCAGGTGCTTTGCTCAG

Table 6. Sequences of Direct Repeats (DR) identified in those RGPs that can excise from their chromosomal locations

RGP	Associated tRNA	Strains	Detection of excision	Presence of circular form	Sequence of DR (5'-3')
RGP5	Gly	PAO1	+	-	TGGAGCGGGCGAAGGGAATCGAACCT
		PA14	+	-	
		PA2192	+	+	
RGP8	Ser	C3719	+	+	AGGGTTCAAATCCCTATCTCTCCGCCATTACATATGAGCAAAGCCCCTG
RGP17	Leu	PA2192	+	+	TCTCTCCGTCCGCACCAT
RGP41 (PAPI-1)	Lys	PA14	+	+	CTTTTAACCAATTGGTCGTAGGTTCGAATCCTACACGACCCACCAT
		PA2192	+	-	
		C3719	+	+	
		PACS2	+	+	
RGP62	Phe	PACS2	+	+	GTTTCGATTCCGTCCCTGGCACCAT