

SUPPLEMENTARY MATERIAL

Table of contents

Table S1	pages S2-S8
List of phage clusters present in dot plot figures	page S9
Legends for figures S1-S7	pages S10-S14
Figure S1A-S1W	page S15-S37
Figure S2A-S2C	page S38-S40
Figure S3	page S41
Figure S4	page S42
Figure S5	page S43
Figure S6	page S44
Figure S7	page S45

Table S1

Tailed Phages Genomes Sequences Used in this Study

Footnotes:

- * Not proven to be temperate phages. Homologous DNA molecules are present as plasmids in bacterial genome sequences.
- # Sequence nearly complete, but clearly incomplete as reported.
- † Prophage sequence that has not been proven to be a fully functional phage.

Table S1

CLUSTER	Prototype	SubClusters	Member phages	Host	Accession number	Sequence publication	Genome size (kbp)
Lytic1	T1-like	A	T1	<i>Escherichia coli</i>	AY216660	Virology 318:245	48836
		A	Shf1	<i>Shigella flexneri</i>	NC_015456	–	50661
		A	ADB-2	<i>Escherichia coli</i>	JX912252	GenomeA 1:e00043-13	50552
		B	Rogue1	<i>Escherichia coli</i>	JQ182736	ViroJ 9:207	45805
		B	JK06	<i>Escherichia coli</i>	DQ121662	–	46072
		B	øJLA23	<i>Escherichia coli</i>	KC333879	GenAnn 1:e00219-12	43017
		B	øKP26	<i>Escherichia coli/S. enterica</i>	KC579452	ArchViro1 158:2395	47285
		B	øEB49	<i>Escherichia coli</i>	JF770475	AEM 77:6630	47180
		C	Rtp	<i>Escherichia coli</i>	AM156809	JBACT 188:1419	46219
		C	ACG-M12	<i>Escherichia coli</i>	NC_019404	Viruses 4:471	46054
		D	F20#	<i>Enterobacter aerogenes</i>	JN672684	JGenViro1 93:2310	51543
		D	KP36	<i>Klebsiella pneumoniae</i>	NC_019781	ViroJ 10:100	49818
		E	TLS	<i>Escherichia coli</i>	AY308796	JMolBiol 308:579	49902
		E	FSL_SP-126 #	<i>Salmonella enterica</i>	KC139521	BMCgenomics 14:481	51092
		F	ESP2949-1	<i>Cronobacter sakazakii</i>	JF912400	ArchViro1 157:199	49116
		G	pSf-1	<i>Shigella flexneri</i>	KC710998	ResMicro 164:979	51821
		Lytic2	T4-like	A	T4	<i>Escherichia coli</i>	AY318471
A	wV7			<i>Escherichia coli</i>	HM997020	JViro1 88:1026	166452
A	RB14			<i>Escherichia coli</i>	NC_012638	–	165429
A	RB32			<i>Escherichia coli</i>	NC_008515	–	165890
A	RB51			<i>Escherichia coli</i>	NC_012635	–	168394
A	ECML-134			<i>Escherichia coli</i>	JX128259	–	166783
A	IME09			<i>Escherichia coli</i>	JN202312	–	166499
A	ACG-C40			<i>Escherichia coli</i>	JN986846	Viruses 4:471	167396
A	AR1			<i>Escherichia coli</i>	AP011113	JViro1 85:6567	167435
A	Shf2			<i>Shigella flexneri</i>	NC_015457	–	165919
A	PST			<i>Yersinia pestis</i>	KF208315	–	167785
A	øD1			<i>Yersinia pestis</i>	HE956711	–	167063
A	e11/2			<i>Escherichia coli</i>	NC_024125	–	168470
B	RB69			<i>Escherichia coli</i>	NC_004928	ViroJ 3:30	167560
B	HX01			<i>Escherichia coli</i>	JX536493	JViro1 86:13871	161158
B	JS09			<i>Escherichia coli</i>	NC_024124	–	169148
C	JS10			<i>Escherichia coli</i>	EU863409	Virology 388:21	171451
C	JS98			<i>Escherichia coli</i>	EF469154	JBact 189:8206	170523
C	Bp7			<i>Escherichia coli</i>	HQ829472	JViro1 86:13832	168066
C	IME08			<i>Escherichia coli</i>	NC_014260	ArchViro1 156:1489	172253
D	VR7			<i>Escherichia coli</i>	HM563683	ArchViro1 155:871	169285
D	SP18			<i>Shigella sonnei</i>	GQ981382	JMicrobiol 48:213	170605
E	STML_198			<i>Salmonella enterica</i>	JX181825	–	158099
E	S16			<i>Salmonella enterica</i>	HQ331142	MolMicro 87:818	160221
F	øR1-RT			<i>Yersinia enterocolitica</i>	HE956709	–	168809
G	CC31			<i>Escherichia coli</i>	GU323318	–	166540
G	PG7			<i>Enterobacter cloacae</i>	KJ101592	–	173276
H	RB49			<i>Escherichia coli</i>	NC_005066	ViroJ 3:30	164018
H	ø1			<i>Escherichia coli</i>	NC_009821	–	164270
H	JSE			<i>Escherichia coli</i>	EU863408	Virology 388:21	166418
I	RB16			<i>Escherichia coli</i>	HM134276	–	176789
I	RB43			<i>Escherichia coli</i>	NC_007023	ViroJ 3:30	180500
I	Lw1			<i>Escherichia coli</i>	NC_021344	GenomeA 1:e00743-13	176227
I	KP15			<i>Klebsiella pneumoniae</i>	GU295964	ViroJ 10:100	174436
I	KP27			<i>Klebsiella pneumoniae</i>	HQ918180	ViroJ 10:100	174413
I	GAP161			<i>Cronobacter sakazakii</i>	JN882287	JViro1 86:13806	178193
J	PS2	<i>Serratia marcescens</i>	KJ025957	–	167266		
not Entero	44RR2.t8	<i>Aeromonas</i>	AY375531	–	173591		
not Entero	31	<i>Aeromonas</i>	AY962392	–	172963		
not Entero	25	<i>Aeromonas</i>	DQ529280	–	161475		
not Entero	IME13	<i>Stenotrophomonas</i>	JX306041	–	162327		
not Entero	Acj61	<i>Acinetobacter</i>	GU911519	ViroJ 7:292	164093		
not Entero	Ac42	<i>Acinetobacter</i>	HM032710	ViroJ 7:292	167716		
Lytic3	Vi01-like	A	CBA120	<i>Escherichia coli</i>	JN593240	ViroJ 8:430	157304
		A	Phax1	<i>Escherichia coli</i>	JN673056	Microbiology 159:1629	156628
		A	ECML-4	<i>Escherichia coli</i>	JX128257	–	157308
		A	SKML-39	<i>Salmonella enterica</i>	JX181829	–	159624
		A	øSH19	<i>Salmonella enterica</i>	JN126049	ViroJ 8:498	157785
		A	SFP10	<i>Salmonella enterica</i>	HQ259103	ApplEnvMicro 78:58	157950
		A	Vi01 (VII)	<i>Salmonella enterica</i>	FQ312032	JBact 192:5746	157061
		A	Det7	<i>Salmonella enterica</i>	–	Roger Hendrix, pers. com.	157498
		A	Marshall	<i>Salmonella enterica</i>	KF669653	GenomeA 1:e00867	156338
		A	Maynard	<i>Salmonella enterica</i>	KF669654	GenomeA 1:e00867	154701
		A	SJ2	<i>Salmonella enterica</i>	KJ174317	–	152460
		A	SJ3	<i>Salmonella enterica</i>	KJ174318	–	162910
		B	LIMEstone1	<i>Dickeya solani</i>	HE600015	PLoSOne 7:e33227	152472
		B	øSboM-AG3	<i>Shigella boydii</i>	FJ373894	ViroJ 8:242	158006
		C	0507-KN2-1	<i>Klebsiella pneumoniae</i>	AB797215	–	159991
		D	øMAM1	<i>Serratia plymuthica</i>	JX878496	JViro1 86:13872	157834
		Lytic4	T5-like	A	T5	<i>Escherichia coli</i>	AY543070
A	EPS7			<i>Salmonella enterica</i>	CP000917	FemsMicroLett 289:202	111382
A	SPC35			<i>Salmonella/Escherichia coli</i>	HQ406778	ApplEnvMicro 77:2042	118351
A	AKFV33			<i>Escherichia coli</i>	NC_017969	PLoSOne e34585	108853
A	H8#			<i>Salmonella enterica</i>	AC171169	JBact 189:5658	104373
A	øR201			<i>Yersinia enterocolitica</i>	HE956708	–	112795
B	My1			<i>Pectobacterium carotovorum</i>	JX195166	JViro1 86:11410	122024
not entero	Vpv-1	<i>Vibrio parahaemolyticus</i>	JQ340389	JViro1 86:7103	111506		

Table S1

CLUSTER	Prototype	SubClusters	Member phages	Host	Accession number	Sequence publication	Genome size (kbp)
Lytic5	T7-like	A	T7	<i>Escherichia coli</i>	V01146	JMolBiol 166: 477	39936
	T7 supercluster	A	øA1122	<i>Yersinia pestis</i>	AY247822	JBact 185:5248	37555
		A	YpsP-G	<i>Yersinia pestis</i>	JQ965703	Bacteriophage 2:168	38288
		A	YpP-R	<i>Yersinia pestis</i>	JQ965701	Bacteriophage 2:168	38284
		A	YpP-Y	<i>Yersinia pestis</i>	JQ965700	Bacteriophage 2:168	37714
		A	R	<i>Yersinia pestis</i>	JX000007	–	37260
		A	Y	<i>Yersinia pestis</i>	JQ957925	–	37432
		A	13a	<i>Escherichia coli</i>	EU734174	–	38841
		A	Vi VI (VI06)	<i>Salmonella enterica</i>	FR667955	JBact 192:5746	38368
		A	Vi VII (VI07)	<i>Salmonella enterica</i>	Derek Pickard, pers. comm.	JBact 192:5746	39248
		A	Vi V (Vi05)	<i>Salmonella enterica</i>	Derek Pickard, pers. comm.	JBact 192:5746	38582
		A (not entero)	IME15	<i>Stenotrophomonas maltophilia</i>	JX872508	JVirol 86:13839	38,513
		B	T3	<i>Escherichia coli</i>	AJ318471	JMolBiol 319:1115	38208
		B	øYeO3-12	<i>Yersinia enterocolitica</i>	AJ251805	JBact 183:1928	39600
		B	øSG-JL2	<i>Salmonella enterica</i>	EU547803	ApplEnvMicro 74:6970	38815
						JGenVirol 62(Pt 2):331; Gene98:89	
		C	BA14	<i>Escherichia coli</i>	NC_011040	–	39816
		C	YpP-G	<i>Yersinia pestis</i>	JQ965702	–	39415
		C	Berlin	<i>Yersinia pestis</i>	NC_008694	–	37555
		C	285p	<i>Escherichia coli</i>	GQ468526	VirusGenes 48:528	39270
		C	Yepe2	<i>Yersinia pestis</i>	NC_011038	JGenVirol 92(Pt 1):216	38677
		C	Yep-ø	<i>Yersinia pestis</i>	HQ333270	JGenVirol 92(Pt 1):216	38616
		C	Kvp1	<i>Kluyvera cryocrescens</i>	FJ194439	VirolJ 5:122	39472
		C	FE44	<i>Erwinia</i>	NC_022744	–	39860
		D	K30	<i>Escherichia coli</i>	HM480846	–	40940
		D	KP32	<i>Klebsiella pneumoniae</i>	GQ413937	VirolJ 10:100	41119
		D	K11	<i>Klebsiella sp. 390</i>	EU734173	G3 2:825	41181
		E	L1	<i>Erwinia amylovora</i>	HQ728265	ApplEnvMicro 77:5945	39282
		F	Mmp1	<i>Morganella morganii</i>	EU652770	Genomics 96:167	38233
		G	CR8	<i>Citrobacter rodentium</i>	HG818824	–	39651
		H	CR44b	<i>Citrobacter rodentium</i>	HG818822	–	39207
		H	EcoDS1	<i>Escherichia coli</i>	NC_011042	–	39252
		H	Dev2	<i>Cronobacter turicensis</i>	HG813241	–	38966
		H	K1F	<i>Escherichia coli</i>	AM084414	JBact 187:8499	39699
		H	Vi III (VI03)	<i>Salmonella enterica</i>	Derek Pickard, pers. comm.	JBact 192:5746	38969
		not entero	ICP3_2009_B	<i>Vibrio cholerae</i>	HQ641341	–	39,042
		not entero	VP3	<i>Vibrio cholerae</i>	JQ780163	–	38,927
		not entero	N4	<i>Vibrio cholerae</i>	JQ780163	–	38,983
		not entero	ICP3_2009_A	<i>Vibrio cholerae</i>	HQ641342	–	39,503
		not entero	ICP3	<i>Vibrio cholerae</i>	NC_015159	–	40973
		not entero	VP4	<i>Vibriophage</i>	NC_007149	–	41386
		not entero	øBB-PF7A	<i>Pseudomonas fluorescens</i>	NC_015264	BMCBiotechnol 8:80	39562
		not entero	PPpW-4	<i>Pseudomonas plecoglossicida</i>	NC_023005	–	40192
		not entero	ø15	<i>Pseudomonas putida</i>	NC_015208	PLoS One 6:e18597	37359
		not entero	ø-S1	<i>Pseudomonas fluorescens</i>	NC_021062	JVirol 86:10239	40441
		not entero	gh-1	<i>Pseudomonas putida</i>	NC_004665	Virology311:305	43444
		not entero	RSB2	<i>Ralstonia solanacearum</i>	NC_023736	–	43769
		not entero	RHEph01	<i>Rhizobium etli</i>	JX483873	ApplEnvMicro 80:446	44371
Lytic6	SP6-like	A	SP6	<i>Salmonella enterica</i>	AY288927	JBact 186:1933	43769
	T7 supercluster	A	ACG-C91	<i>Escherichia coli</i>	NC_019403	Viruses 4:471	44385
		A	K1-5 (K1-dep(3))	<i>Escherichia coli</i>	AY370674	JMolBiol 335:1151	45251
		A	K1E (K1-dep(2))	<i>Escherichia coli</i>	AM084415	MolMicro 60:1123	43931
		A	UAB-ø78	<i>Salmonella enterica</i>	GU595417	IntJFoodMicrobiol 165:169	44618
						Derek Pickard, pers. comm.	
		B	Vi IV (Vi04)	<i>Salmonella enterica</i>		JBact 192:5746	44400
		C	PP1	<i>Pectobacterium carotovorum</i>	JQ837901	JVirol 86:8899	45445
		D	ERA103	<i>Erwinia amylovora</i>	EF160123	ApplEnvMicro 51:862	45554
		D	øEa100	<i>Erwinia amylovora</i>	FQ482086	JBact 193:795	45522
		D	øEa1h	<i>Erwinia amylovora</i>	FQ482084	JBact 193:795	44541
		not entero	Vc1	<i>Vibrio sp.</i>	KJ502657	–	43809
Lytic7	KP34-like	A	KP34	<i>Klebsiella pneumoniae</i>	NC_013649	AppMicroBiotech 90:1333	43766
	T7 supercluster	A	F19	<i>Klebsiella pneumoniae</i>	KF765493	–	44546
		B	LIMElight	<i>Pantoea agglomerans</i>	FR687252	ApplEnvMicro 77:3443	43931
		not entero	VP93	<i>Vibrio parahApp/EnvMicroolytic</i>	NC_012662	–	
Lytic8	LIMEzero	A	LIMEzero	<i>Pantoea agglomerans</i>	FR751545	ApplEnvMicro 77:3443	43032
Lytic9	T7 supercluster	A	øKT-like	<i>Escherichia coli</i>	JN882298	–	42608
Lytic10	GAP227-like	A	GAP227	<i>Coronobacter sakazakii</i>	NC_020078	GenomeA 1:e00122-12	41796
	T7 supercluster	B	øR8-01	<i>Yersinia enterocolitica</i>	HE956707	–	42081
		C	ø80-18	<i>Yersinia enterocolitica</i>	HE956710	–	41572
		not entero	øAS7	<i>Aeromonas salmonicida</i>	NC_019528	–	

Table S1

CLUSTER	Prototype	SubClusters	Member phages	Host	Accession number	Sequence publication	Genome size (kbp)
not Entero	T7 Supercluster	not entero	Cd1	<i>Caulobacter crescentus</i>	GU393987	–	41581
		not entero	JG068	<i>Burkholderia cenocepacia</i> K56	NC_022916	–	41604
		not entero	Paz	<i>Xyella fastidiosa</i>	NC_022982	–	43869
		not entero	Prado	<i>Xyella fastidiosa</i>	NC_022987	–	43940
		not entero	Bf7	<i>Pseudomonas tolaasii</i> LMG 23	NC_016764	–	40058
		not entero	RSB1	<i>Ralstonia</i>	NC_011201	–	43079
		not entero	RHEph02	<i>Rhizobium etli</i>	JX483874	ApplEnvMicro 80:446	46486
		not entero	RHEph08	<i>Rhizobium etli</i>	JX483879	ApplEnvMicro 80:446	43619
		not entero	øAB1	<i>Acinetobacter baumannii</i>	HQ186308	Genomics91:249	41526
		not entero	Abp1	<i>Acinetobacter baumannii</i>	NC_021316	CurrMicrobiol 66:535	42185
		not entero	AB3	<i>Acinetobacter baumannii</i>	NC_021337	–	31185
		not entero	phi-2	<i>Pseudomonas fluorescens</i>	NC_013638	–	43114
		not entero	RSB3	<i>Ralstonia solanacearum</i>	NC_022917	–	44578
		not entero	øKMV	<i>Pseudomonas aeruginosa</i>	EU236438.1	Virology 312:49	42519
		not entero	LKA1	<i>Pseudomonas aeruginosa</i>	AM265639.1	JBact 188:6924	41593
		Lytic11	N4-like	A	N4	<i>Escherichia coli</i>	EF056009
A	EC1-UPM			<i>Escherichia coli</i>	KC206278	ViroJ 10:308	70831
A	IME11			<i>Escherichia coli</i>	NC_019423	JViro 86:13861	77570
A	G7C			<i>Escherichia coli</i>	NC_015933	Virology 426:93	72917
A	KNBP21 (ECBP1?)			<i>Escherichia coli</i>	JX415535	JViro 86:12439	69855
A	pSb-1			<i>Shigella boydii</i>	KF620435	–	71629
B	FSL_SP-058			<i>Salmonella enterica</i>	KC139517	BMCgenomics 14:481	72369
B	FSL_SP-076			<i>Salmonella enterica</i>	KC139520	BMCgenomics 14:481	72098
C	EcP1			<i>Enterobacter cloacae</i>	NC_019485	–	59080
D	Ea9-2			<i>Erwinia amylovora</i>	KF806588	–	75568
E	S6			<i>Erwinia amylovora</i>	HQ728266	ApplEnvMicro 77:5945	74669
not entero	JWAalpha			<i>Achromobacter xylosoxidans</i>	KF787094	Viro J 11:14	72329
not entero	JWDelta			<i>Achromobacter xylosoxidans</i>	KF787095	Viro J 11:14	73659
Lytic12	9NA-like	A	9NA	<i>Salmonella enterica</i>	KJ802832	–	52869
		A	FSL_SP-069 #	<i>Salmonella enterica</i>	KC139632-37	BMCgenomics 14:481	56593
		A	FSL_SP-062 #	<i>Salmonella enterica</i>	KC139649-51	BMCgenomics 14:481	56593
Lytic13	Chi-like	A	Chi (X)	<i>Salmonella enterica</i>	JX094499	ArchViro 158:2179; Roger Hendrix, unpub	59578
		A	FSL_SP-030	<i>Salmonella enterica</i>	KC139519	BMCgenomic 14:481	59746
		A	FSL_SP-039	<i>Salmonella enterica</i>	KC139514	BMCgenomic 14:481	59815
		A	FSL_SP-088	<i>Salmonella enterica</i>	KC139512	BMCgenomic 14:481	59454
		A	FSL_SP-124	<i>Salmonella enterica</i>	KC139515	BMCgenomic 14:481	59245
		A	iEPS5	<i>Salmonella enterica</i>	KC677662	ApplEnvMicro 79:4829	59254
		A	SPN19	<i>Salmonella enterica</i>	JN871591	–	59203
		B	Enc34	<i>Enterobacter cancerogenus</i>	JQ340774	JViro 86:11403	60384
		C	RedJac	<i>Providencia stewartii</i>	NC_018832	PLoS One 8:e61762	58104
Lytic14	øEco32-like	A	Eco32	<i>Escherichia coli</i>	EU330206	JMolBiol 377:774	77554
		A	KBNP135 (ECBP2)	<i>Escherichia coli</i>	JX415536	JViro 86:12439	77315
		A	NJ01	<i>Escherichia coli</i>	JX867715	JViro 86:13874	77448
		A	KBNP1711	<i>Escherichia coli</i>	KF981730	–	76184
		B	7-11	<i>Salmonella enterica</i>	HM997019	ArchViro156:149	89916
C	GAP52	<i>Cronobacter sakazakii</i>	JN882286	–	76631		
Lytic15	Felix-O1-like	A	Felix-O1	<i>Salmonella enterica</i>	AF320576	Viruses 2:710	86155
		A	UAB-ø87	<i>Salmonella enterica</i>	JN225449	–	87603
		A	FO1a	<i>Salmonella enterica</i>	JF461087	–	88331
		A	EC6	<i>Escherichia coli</i>	JX560968	GenomeA 1: e00085-12	86231
		A	wV8	<i>Escherichia coli</i>	EU877232	ViroJ 6:41	88487
		A	JH2	<i>Escherichia coli</i>	KF055347	–	87721
		B	øEa21-4	<i>Erwinia amylovora</i> +others	EU710883	ApplEnvMicro 75:2139	84567
		B	øEa104	<i>Erwinia amylovora</i>	FQ482083	JBact 193:795	84565
		B	M7	<i>Erwinia amylovora</i>	HQ728263	ApplEnvMicro 77:5945	84694
		Lytic16	SETP3-like SETP3 Supercluster	A	SETP3	<i>Salmonella enterica</i>	EF177456
A	AG11			<i>Salmonella enterica</i>	JX297445	–	41546
A	SETP7			<i>Salmonella enterica</i>	KF562865	–	42789
A	SETP13			<i>Salmonella enterica</i>	KF562864	–	42665
A	Ent1			<i>Salmonella enterica</i>	HE775250	JGenViro 93:2046	42391
A	SE2			<i>Salmonella enterica</i>	JQ007353	JViro 86:7712	43221
A	SSe3 (KST)			<i>Salmonella enterica</i>	AY730274	–	40794
A	wsk13			<i>Salmonella enterica</i>	JX202565	ApplEnvMicro 79:1958	42633
A	FSL_SP-101			<i>Salmonella enterica</i>	KC139511	BMCgenomics 14:481	41873
A	Jersey			<i>Salmonella enterica</i>	KF148055	–	43447
B	K1-dep(4) / (K1G)			<i>Escherichia coli</i>	GU196277	Virology 398:79	43587
B	K1-dep(1) / (K1H)			<i>Escherichia coli</i>	GU196278	Virology 398:79	41632
B	K1-ind(1)			<i>Escherichia coli</i>	GU196279	Virology 398:79	42292
B	K1-ind(2)			<i>Escherichia coli</i>	GU196280	Virology 398:79	42765
B	K1-ind(3)			<i>Escherichia coli</i>	GU196281f	Virology 398:79	43461
C	FSL_SP-031			<i>Salmonella enterica</i>	KC139518	BMCgenomics 14:481	42215
D	Eta (η)	<i>Serratia marcescens</i>	KC460990	ViroJ 11:6	42724		

Table S1

CLUSTER	Prototype	SubClusters	Member phages	Host	Accession number	Sequence publication	Genome size (kbp)
Lytic17	SO-1-like	A	SO-1	<i>Sodalis glossinidius</i>	GQ502199	–	45169
	SETP3 Supercluster	A	EP23	<i>Shigella sonnei</i>	JN984867	JMicro 49:927	44077
		A	EEP (SSL-2009a)	<i>Escherichia coli</i>	FJ750948	Interviro 53:211	39792
		A	HK578	<i>Escherichia coli</i>	JQ086375	–	43741
		A	JL1	<i>Escherichia coli</i>	JX865427	ArchViro 158:2429	43457
		B	eiAU	<i>Edwardsiella ictaluri</i>	Mark Liles	Viro 8:6	42807
		B	eiAu-183	<i>Edwardsiella ictaluri</i>	KF772234	–	43017
		B	eiDWF	<i>Edwardsiella ictaluri</i>	Mark Liles	Viro 8:6	41999
		B	eiMSLS	<i>Edwardsiella ictaluri</i>	Mark Liles	Viro 8:6	42460
Lytic18	ECO1230-10-like	A	ECO1230-10	<i>Escherichia coli</i>	GU903191	Ve Microb 148:267	41666
Lytic19	Gj1-like	A	Gj1 (øEcoM-Gj1)	<i>Escherichia coli</i>	EF460875	AEM 74:516	52975
		B	PM1	<i>Pectobacterium carotovorum</i>	KF534715	ArchViro 2014 in press	55098
		C	Y2	<i>Erwinia amylovora</i>	HQ728264	AEV 77:5945	56621
Lytic20	PY100-like	A	PY100	<i>Yersinia enterocolitica</i>	AM076770	JBact 190:332	50291
Lytic21	ø92-like	A	ø92	<i>Escherichia coli</i>	FR775895	J Virol 86:10384	148612
	rV5 Supercluster	A	phAPEC8	<i>Escherichia coli</i>	JX561091	J Virol 86:13117	147737
Lytic22	rV5-like	A	rV5	<i>Escherichia coli</i>	NC_011041	Viro 10:76	137947
	rV5 Supercluster	A	FV3	<i>Escherichia coli</i>	JQ031132	Arch Viro 157:2431	136947
		A	2_JES-2013	<i>Escherichia coli</i>	NC_022323	–	136910
		B	PVP-SE1	<i>Salmonella enterica</i>	GU070616	J Virol 85:11265	145964
		B	SSE-121	<i>Salmonella enterica</i>	JX181824	–	147745
		B	GAP31	<i>Cronobacter sakazakii</i>	JN882284	J Virol 86:13830	147940
		B	4MG	<i>Escherichia coli</i>	KF550303	ArchViro 2014 in press	148567
		C	øTE	<i>Pectobacterium atrosepticum</i>	NC_020201	PLoS Genet 8:e1003023	142349
		C	CR3	<i>Cronobacter sakazakii</i>	JQ691612	JViro 86:6367	149273
		C	CR9	<i>Cronobacter sakazakii</i>	JQ691611	–	151924
Lytic23	SPN3US-like	A	SPN3US	<i>Salmonella enterica</i>	JN641803	JViro 85:13470	240413
		B	øEaH2	<i>Erwinia amylovora</i>	JX316028	JViro 86:10899	243050
		C	CR5	<i>Cronobacter sakazakii</i>	NC_021531	–	233989
Lytic24	Rak2-like	A	RaK2	<i>Klebsiella sp. KV-3</i>	JQ513383	PLoS One 8:e60717	345809
		B	GAP32	<i>Cronobacter sakazakii</i>	JN882285	–	358663
		C	PBECO 4	<i>Escherichia coli</i>	KC295538	ArchViro 158:2399	348133
Lytic25	øR1-37-like	A	øR1-37	<i>Yersinia enterocolitica</i>	AJ972879	JViro 86:12625	262391
Lytic26	E1-like	A	E1	<i>Salmonella enterica</i>	AM491472	JBact 190:2580	45362
Lytic27	EMCL-117-like	A	ECML-117	<i>Escherichia coli</i>	JX128258	–	66854
Lytic28	KF-1-like	A	K1-F	<i>Edwardsiella tarda</i>	AB757800	GenomeA 1:e00089-12	41549
		A	IW-1	<i>Edwardsiella tarda</i>	AB757801	GenomeA 1:e00089-12	41684
Lytic29	MSW-3-like	A	MSW-3	<i>Edwardsiella tarda</i>	AB767244	GenomeA 1:e00248-12	42746
		A	PEi2	<i>Edwardsiella ictaluri</i>	NC_021342	GenomeA 2:e00228-14	43418
		B	JD001	<i>Klebsiella pneumoniae</i>	JX866719	JViro 86:13843	48814
Lytic30	Ea35-70-like	A	Ea35-70	<i>Erwinia amylovora</i>	KF806589	–	271084
Lytic31	øEaH1-like	A	øEaH1	<i>Erwinia amylovora</i>	KF623294	FEMSMicroLett 350:25	218339
Lytic32	9g-like	A	9g	<i>Escherichia coli</i>	KJ419179	–	56702
Temperate							
Temp1	λ-like	A	lambda (λ)	<i>Escherichia coli</i>	J02459	JMolBiol 162: 729	48503
	λ Supercluster	A	HK630	<i>Escherichia coli</i>	JQ086376	–	47090
		A	HK629	<i>Escherichia coli</i>	JQ182735	–	47288
Temp2	ø80-like	A	ø80	<i>Escherichia coli</i>	JX871397	JBact 194:6802	46150
	λ Supercluster	B	HK225	<i>Escherichia coli</i>	JQ086371	–	45366
		B	mEp237	<i>Escherichia coli</i>	JQ182730	–	44375
Temp3	N15-like	A	N15	<i>Escherichia coli</i>	AF064539	JMolBiol 299:53	46363
	λ Supercluster	B	PY54	<i>Yersinia enterocolitica</i>	AJ564013	JMolBiol 331:605	46339
		C	øKO2*	<i>Klebsiella oxytoca</i>	AY374448	JBact 186:1818	51601
Temp4	HK97-like	A	HK97	<i>Escherichia coli</i>	AF069529	JMolBiol 299:27	39732
	λ Supercluster	A	HK022	<i>Escherichia coli</i>	AF069308	JMolBiol 299:27	40751
		A	HK75	<i>Escherichia coli</i>	HM173637	–	36661
		A	HK106	<i>Escherichia coli</i>	JQ086369	–	41468
		A	HK140	<i>Escherichia coli</i>	JQ086370	–	40710
		A	HK446	<i>Escherichia coli</i>	JQ086372	–	39026
		A	HK542	<i>Escherichia coli</i>	JQ086373	–	38964
		A	HK544	<i>Escherichia coli</i>	JQ086374	–	40155
		A	HK633	<i>Escherichia coli</i>	JQ086377	–	41528
		A	mEpX1	<i>Escherichia coli</i>	JQ182727	–	41567
		A	mEpX2	<i>Escherichia coli</i>	JQ182726	–	38759
		A	mEp234	<i>Escherichia coli</i>	JQ182732	–	39578
		A	mEp235	<i>Escherichia coli</i>	JQ182731	–	37595
		B	mEp390	<i>Escherichia coli</i>	JQ182729	–	40029
		C	ENT39118	<i>Cronobacter sakazakii</i>	HQ201307	JViro 86:5400	39012
Temp5	ES18-like	A	ES18	<i>Salmonella enterica</i>	AY736146	JBact 187:1091	46900
	λ Supercluster	A	Oslo	<i>Salmonella enterica</i>	NC_018279	–	49116
		B	SPN3UB	<i>Salmonella enterica</i>	JQ288021	JViro 86:3404	47355

Table S1

CLUSTER	Prototype	SubClusters	Member phages	Host	Accession number	Sequence publication	Genome size (kbp)
Temp6	gifsy-2-like λ Supercluster	A	gifsy-2	<i>Salmonella enterica</i>	NC_010393	Nature 413:852	45841
		A	gifsy-1	<i>Salmonella enterica</i>	NC_010392	Nature 413:852	51021
		B	Fels-1	<i>Salmonella enterica</i>	NC_010391	Nature 413:852	43667
		C	mEp043	<i>Escherichia coli</i>	JQ182734	–	42780
		C	mEp213	<i>Escherichia coli</i>	JQ182733	–	44120
		D	CP-1639†	<i>Escherichia coli</i>	AJ304858	Microbiology 151:941	39445
		E	CTD-lø	<i>Escherichia coli</i>	AB285204	PNAS 104:14483	47021
		E	mEp460	<i>Escherichia coli</i>	JQ182728	–	44510
		F	FSL_SP-016 #	<i>Salmonella enterica</i>	KC139516	BMCgenomics 14:481	46698
Temp7	BP-4795-like λ Supercluster	A	BP-4795	<i>Escherichia coli</i>	AJ556162	JBact 187:8498	57930
		A	2851	<i>Escherichia coli</i>	FM180578	JBact 76:5466	57248
		A	Stx2-1717†	<i>Escherichia coli</i>	NC_011357	–	62147
		A	YYZ-2008	<i>Escherichia coli</i>	FJ184280	–	54896
Temp8	SFV-like λ Supercluster	A	SFV	<i>Shigella flexneri</i>	AF339141	JBact 184:1974	37074
		A	SfII	<i>Shigella flexneri</i>	NC_021857	GenomeA 1: e00626-13	41459
		A	SfIV	<i>Shigella flexneri</i>	KC814930	BMCgenomics 14:677	39758
		B	SfI	<i>Shigella flexneri</i>	JX509734	BMCMicro 13:39	38389
		C	øP27	<i>Escherichia coli</i>	AJ298298	Inf Immun 70:1896	45575
		D	ST64B†	<i>Salmonella enterica</i>	AY055382	JBact 185:6481	40149
Temp9	P22-like λ Supercluster	A	P22	<i>Salmonella enterica</i>	BK000583	JBact 185:1475	41724
		A	L	<i>Salmonella enterica</i>	AY795968	–	40679
		A	SPN9CC	<i>Salmonella enterica</i>	JF900176	ApplEnvMicro 80:374	40128
		A	ST64T	<i>Salmonella enterica</i>	AY052766	JBact 185:3475	40679
		A	ST104	<i>Salmonella enterica</i>	AB102868	JBact 42:1807	41391
		A	ST160	<i>Salmonella enterica</i>	NC_014900	–	40986
		A	epsilon34 (ε34)	<i>Salmonella enterica</i>	EU570103	BMCMicro 8:227	43016
		A	g341	<i>Salmonella enterica</i>	NC_013059	–	40975
		A	SE1	<i>Salmonella enterica</i>	NC_011802	Virology 351:360	41941
		A	Emek	<i>Salmonella enterica</i>	JQ806763	–	39783
		A	ø20 #	<i>Salmonella enterica</i>	GQ422450	–	41809
		B	IME10	<i>Escherichia coli</i>	JN600960	–	39464
		B	Sf6	<i>Shigella flexneri</i>	AF547987	JMolBiol 339:379	39044
		B	HK620	<i>Escherichia coli</i>	AF335538	JMolBiol 311:657	38297
		B	CUS-3	<i>Escherichia coli</i>	CP000711	JBact 189:6447	40207
		B	SPC-P1	<i>Salmonella enterica</i>	CP000857*	BMCgenomics 11:729	39659
Temp10	APSE-1-like λ Supercluster	A	APSE-1†	<i>Hamiltonella defensa</i>	AF157835	Virology 262:104	36524
		A	APSE-2†	<i>Hamiltonella defensa</i>	NC_011551	PNAS 102:16919	39867
Temp11	933W-like λ Supercluster	A	933W	<i>Escherichia coli</i>	AF125520	JBact 181:1767	61670
		A	stx1ø	<i>Escherichia coli</i>	AP005153	JBact 185:3966	59886
		A	stx2ø-l	<i>Escherichia coli</i>	AP004402	Gene 309:35	61765
		A	stx2ø-ll	<i>Escherichia coli</i>	AP005154	JBact 185:3966	62706
		A	stx2-86†	<i>Escherichia coli</i>	NC_008464	–	60238
		A	min27	<i>Escherichia coli</i>	EU311208	–	63395
		A	ø24B	<i>Escherichia coli</i>	HM208303	BMCgenomics 13:311	57677
		A	P13374	<i>Escherichia coli</i>	HE664024	JVirol 86:10444	60894
		A	TL-2011c	<i>Escherichia coli</i>	JQ011318	PLoSOne 7:e31413	60523
		A	VT2-sakai†	<i>Escherichia coli</i>	AP000363	DNARes 6:235	60942
		A	VT2ø_272	<i>Escherichia coli</i>	HQ424691	–	65955
Temp12	HK639-like λ Supercluster	A	HK639	<i>Escherichia coli</i>	NC_016158	–	49576
Temp13	øES15-like λ Supercluster	A	øES15	<i>Cronobacter sakazakii</i>	JQ780327	JVirol 86:7713	39974
Temp14	HS2-like λ Supercluster	A	HS2†	<i>Sodalis glossinidius</i>	CP006569	GenBioEvol 6:76	58621
Temp15	ENT47970-like λ Supercluster	A	ENT47670	<i>Cronobacter sakazakii</i>	HQ201308	–	47611
Temp16	ZF40-like λ Supercluster	A	ZF40	<i>Pectobacterium carotovorum</i>	NC_019522	PLoS One 7:e40102	48454
Temp17	øEt88-like λ Supercluster	A	øEt88	<i>Erwinia tasmaniensis</i>	FQ482085	JBact 193:795	47279
Temp18	epsilon 15 (ε15)	A	epsilon 15 (ε15)	<i>Salmonella enterica</i>	AY150271	Virology 369:234	39671
		B	SPN1S	<i>Salmonella enterica</i>	JN391180	JVirol 86:1284	38684
		B	SPN9TCW	<i>Salmonella enterica</i>	JQ691610	–	38689
		B	SPC32H	<i>Salmonella enterica</i>	KC911856	JVirol 86:11775	38689
		B	SPC32N	<i>Salmonella enterica</i>	KC911857	JVirol 86:11775	38689
		C	TL-2011b	<i>Escherichia coli</i>	JQ011317	PLoS One 7:e31431	44784
		D	øV10	<i>Escherichia coli</i>	DQ126339	FEMSLett 292:182	39104
Temp19	P1-like	A	P1	<i>Escherichia coli</i>	AF234172	JBact 186:7032	93601
		A	P7	<i>Escherichia coli</i>	AF503408	–	91660
		A	p0111_2*†	<i>Escherichia coli</i>	AP010962	PNAS 106:17939	97897

Table S1

CLUSTER	Prototype	SubClusters	Member phages	Host	Accession number	Sequence publication	Genome size (kbp)		
Temp20	P2-like	A	P2	<i>Escherichia coli</i>	AF063097	numerous publications	33593		
		A	Wø	<i>Escherichia coli</i>	AY135739	–	32684		
	P2 Supercluster	A	L-413C	<i>Yersinia pestis</i>	AY251033	Virology 372:85	30728		
		A	øAA91-ss	<i>Escherichia coli</i>	KF322032	JVirol 87:13665	33628		
		B	186	<i>Escherichia coli</i>	U32222	numerous publications	30624		
		B	PsP3	<i>Salmonella enterica</i>	NC_005340	Virology 188:414	30636		
		B	FSL-SP-004	<i>Salmonella enterica</i>	KC139521	BMCgenomics 14:481	29742		
		C	SopEø	<i>Salmonella enterica</i>	AY319521	JBact 185:5182	35155		
		C	Fels-2	<i>Salmonella enterica</i>	NC_010463	Nature 413:852	33638		
		C	RE-2010 (ELøS)	<i>Salmonella enterica</i>	HM770079	ApplEnvMicro 78:1785	34117		
		D	ENT90	<i>Erwinia amylovora</i>	HQ110084	–	29564		
		Temperate21	ESSI-2-like	A	ESSI-2	<i>Cronobacter sakazakii</i>	HQ110083	ArchViro 156:721	28765
				P2 supercluster	not entero	øCTX	<i>Pseudomonas aeruginosa</i>	AB008550	MolMicro 31:399
	not entero		RSA1	<i>Ralstonia solanacearum</i>	AB276040	JBact 190:143	38760		
	not entero		KL3	<i>Burkholderia cepacia</i>	GU911304	BMCGenomics 11:599	40555		
	not entero		øMHaA1	<i>Mannheimia haemolytica</i>	NC_008201	Virology 350:79	34525		
	not entero		HP1	<i>Haemophilus influenzae</i>	NC_001697	NucleicAcidRes 24:685	32355		
	not entero		K139	<i>Vibrio cholerae</i>	NC_003313	JBact 184:6592	33106		
Temperate22	Mu-like	A	Mu	<i>Escherichia coli</i>	AF083977	JMolBiol 317: 337	36717		
		A	D108	<i>Escherichia coli</i>	GQ357916	–	37235		
		not entero	DMS3	<i>Pseudomonas aeruginosa</i>	DQ631426	JBact 191:210	36415		
		not entero	MP22	<i>Pseudomonas aeruginosa</i>	DQ873690	Microbiology 153:2885	34409		
		not entero	SuMu	<i>Haemophilus</i>	JF832915	BMCGenomics 13:331	37151		
Temperate23*	SSU5-like	A	SSU5	<i>Salmonella enterica</i>	JQ965645	JVirol 86:10894	103229		
		A	pECHO89	<i>Escherichia coli</i>	HG530657	Gen Ann 2: e00356-14	111741		
		A	pLF82*†	<i>Escherichia coli</i>	CU638872	PLoSOne 5:e12714	108379		
		A	pHCM2*†	<i>Salmonella enterica</i>	NC_003385	Plasmid 47:159	106516		
		A	pKPHS1*†	<i>Klebsiella pneumoniae</i>	CP003223	JBact 194:1841	122799		
		B	pMT1*†	<i>Yersinia pestis</i>	AF074611	Inf Immun 66:5731	100900		
Temperate24	PEP14-like	A	pEP-14	<i>Erwinia pyrifoliae</i>	JN585957	–	60714		

SUPPLEMENTARY FIGURES

Phage genome dot plot locations

Cluster	Prototype phage	Dot plot figures
Lytic 1	T1	2*, S1A, S2A*, S2B*
Lytic 2	T4	1*, 5, S2A*, S1B, S1M*
Lytic 3	Vi01	1*, S1C, S1M*, S2A*
Lytic 4	T5	1*, S1D, S1M*, S2A*
Lytic 5	T7	2*, 6, S1E, S1F, S2B*
Lytic 6	SP6	2*, 6, S1E, S1F, S2B*
Lytic 7	KP34	2*, 6, S1E, S1F, S2B*
Lytic 8	LIMEzero	2, 6, S1E, S1F, S2B*
Lytic 9	øKT	2, 6, S1E, S1F, S2B*
Lytic 10	GAP227	2*, 6, S1E, S1F, S2B*
Lytic 11	N4	2*, S1H, S2B*
Lytic 12	9NA	2*, S1G, S2B*
Lytic 13	Chi	2*, S1I, S2B*
Lytic 14	øEco32	2*, S1J, S2B*
Lytic 15	Felix-O1	2*, 4, S1K, S2B*
Lytic 16	SETP3	2*, S1L, S2B*
Lytic 17	SO-1	2*, S1L, S2B*
Lytic 18	ECO1230-10	2, S1G, S2B*
Lytic 19	Gj1	2*, S1G, S2B*
Lytic 20	PY100	2, S1G, S2B*
Lytic 21	Ø92	1*, S1M, S2A*
Lytic 22	rv5	1*, S1M, S2A*
Lytic 23	SPN3US	1*, S1M, S2A*
Lytic 24	Rak2	1, S1M, S2A*
Lytic 25	øR1-37	1, S1M, S2A*
Lytic 26	E1	2, S1G, S2B*
Lytic 27	EMCL-117	2, S1G, S2B*
Lytic 28	KF-1	2*, S1G, S2B*
Lytic 29	MSW-3	2*, S1G, S2B*
Lytic 30	Ea35-70	1, S1M, S2A*
Lytic 31	øEaH1	1, S1M, S2A*
Lytic 32	9g	2, S1G, S2B*
Temperate 1	Lambda	3*, 7, S1N, S1S, S2C*
Temperate 2	ø80	3*, 7, S1N, S1S, S2C*
Temperate 3	N15	3*, 7, S1N, S1S, S2C*
Temperate 4	HK97	3*, 7, S1O, S1S, S2C*
Temperate 5	ES18	3*, 7, S1O, S1S, S2C*
Temperate 6	Gifsy-2	3*, 7, S1P, S1S, S2C*
Temperate 7	BP-4795	3*, 7, S1P, S1S, S2C*
Temperate 8	SfV	3*, 7, S1R, S1S, S2C*
Temperate 9	P22	3*, 7, S1Q, S1S, S2C*
Temperate 10	APSE-1	3, 7, S1Q, S1S, S2C*
Temperate 11	933W	3*, 7, S1R, S1S, S2C*
Temperate 12	HK639	3, 7, S1S, S2C*
Temperate 13	øES15	3, 7, S1S, S2C*
Temperate 14	HS2	3, 7, S1S, S2C*
Temperate 15	ENT47670	3, 7, S1S, S2C*
Temperate 16	ZF40	3, 7, S1S, S2C*
Temperate 17	øEt88	3, 7, S1S, S2C*
Temperate 18	ε15	3*, S1T, S2C*
Temperate 19	P1	1*, S1U
Temperate 20	P2	3*, S1V, S2C*
Temperate 21	ESSI-2	3, S1V, S2C*
Temperate 22	Mu	3, S1W, S2C*
Temperate 23 ⁵	SSU5	1*, S1U
Temperate 24	pEP-14	3, S2C*

* Representative members of cluster shown.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1A. Genome and MCP dot plots of all the phages in the T1-like cluster.

Gepard dot plots of genome nucleotide sequence (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1B. MCP dot plot of all the phages in the T4-like cluster.

A Gepard dot plot of MCP sequences is shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1C. Genome and MCP dot plots of the phages in the Vi01-like cluster.

Gepard dot plots of genome nucleotide sequence (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1D. Genome and MCP dot plots of the phages in the T5-like cluster.

Gepard dot plots of genome nucleotide sequence (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated on the left.

Figure S1E. MCP dot plot of all the phages in the T7 supercluster.

A Gepard dot plot of MCP sequences is shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. An apparent frameshift sequencing error in the MCP of phage YpsP-G was “fixed” in this plot by comparison with close relatives. For ease of comparison to figure S1F, only those phages whose genomes have been suitably annotated for the whole proteome plots are included.

Figure S1F. Genome product dot plot of all the *Enterobacteriaceae* phages in the T7 supercluster.

A Gepard dot plot of complete genome product sequences is shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. Annotation for phages Vi03, Vi04, Vi05 and Vi06 has not been released so these phages are not shown in this plot.

Figure S1G. Genome and MCP dot plots of all the phages in 9NA-, ECO-1230-, Gj1-, PY100-, EMCL-117-, KF-1-, MSW-3- and 9g-like small lytic phage clusters.

Gepard dot plots of genome nucleotide sequence (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences,

and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1H. Genome and MCP dot plots of all the phages in the N4-like supercluster. Gepard dot plots of genome nucleotide sequence (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1I. Genome and MCP dot plots of all the phages in the Chi-like cluster. Gepard dot plots of genome nucleotide sequence (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1J. Genome and MCP dot plots of all of the phages in the øEco32-like cluster. Gepard dot plots of genome nucleotide sequences (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. We note that genomes in this cluster that have been carefully analyzed have direct terminal repeats. These have been variably reported as a single copy on either end of the sequence or with the full repeats on both ends. The sequences used were those reported in GenBank, so the apparent 'movement' of terminal sequence to different ends of the genomes in the genome plot at least in part reflect this reporting inconsistency.

Figure S1K. MCP dot plot of all the phages in the Felix-O1-like cluster. A gepard dot plot of MCP amino acid sequences (below) is shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1L. Genome and MCP dot plots of all the phages in the SETP3- and SO-1-like clusters. Gepard dot plots of genome nucleotide sequences (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1M. Genome and MCP dot plots of all the phages in ø92-, rV5-, SPN3US-, Rak2-, øR1-37-, Ea35-70- and øEaH1-like large lytic clusters. Gepard dot plots of genome nucleotide sequences (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1N. Genome dot plots of all the phages in the lambda-, ø80 and N15-like clusters. Gepard dot plots of genome nucleotide sequences are shown with Gepard word lengths. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. See figure S1S for MCP dot plots and figures 7 and S1S for phage hosts.

Figure S1O. Genome dot plots of all the phages in the HK97- and ES18-like clusters. A Gepard dot plots of genome nucleotide sequences are shown with Gepard word lengths. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. See figure S1S for MCP dot plots and figures 7 and S1S for phage hosts.

Figure S1P. Genome dot plots of all the phages in the gfsy-2- and BP4795-like clusters. Gepard dot plots of genome nucleotide sequences are shown with Gepard word lengths. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. See figure S1S for MCP dot plots and figures 7 and S1S for phage hosts.

Figure S1Q. Genome dot plots of all the phages in the P22- and APSE-1-like clusters. Gepard dot plots of genome nucleotide sequences are shown with Gepard word lengths. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. See figure S1S for MCP dot plots and figures 7 and S1S for phage hosts.

Figure S1R. Genome dot plot of all the phages in the SfV- and 933W-like cluster. A Gepard dot plot of genome nucleotide sequences are shown with Gepard word lengths. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left. See figure S1S for an MCP dot plot.

Figure S1S. MCP dot plot of all the phages in the lambda supercluster. A Gepard dot plot of MCP sequences is shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and blue lines separate clusters; cluster names are indicated on the left. An apparent frameshift error in the MCP of phage Stx2ø-I was “fixed” in this plot by comparison with close relatives.

Figure S1T. Genome and MCP dot plots of all the phages in the epsilon 15-like cluster. Gepard dot plots of genome nucleotide sequences (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, and thick red lines separate subclusters; subcluster names are indicated by red letters on the left.

Figure S1U. Genome and MCP dot plots of all the phages in the P1- and SSU5-like clusters. Gepard dot plots of genome nucleotide sequences (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, thick red lines separate subclusters, and blue lines separate clusters; subcluster names are indicated by red letters on the left.

Figure S1V. Genome and MCP dot plots of all the phages in the P2- and ESS1-2-like clusters. Gepard dot plots of genome nucleotide sequences (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, thick red lines separate subclusters, and blue lines separate clusters; subcluster names are indicated by red letters on the left.

Figure S1W. Genome and MCP dot plots of all the phages in the Mu-like cluster. Gepard dot plots of genome nucleotide sequences (above) and MCP amino acid sequence (below) are shown with Gepard word lengths and hosts indicated. Thin red lines separate the different phage sequences, thick red lines separate subclusters, and blue lines separate clusters; subcluster names are indicated by red letters on the left.

Figure S2A. MCP dot plot analysis of representative *Enterobacteriaceae* tailed phages with genomes larger than 90 kbp. See figure 1 in text of report for host genera.

Figure S2B. MCP dot plot analysis of representative lytic *Enterobacteriaceae* tailed phages with genomes smaller than 90 kbp. See figure 2 in text of report for host genera.

Figure S2C. MCP dot plot analysis of representative temperate *Enterobacteriaceae* tailed phages with genomes smaller than 90 kbp. See figure 3 in text of report for host genera.

Figure S3. Genomic maps of a representative phage from each of the six T7 supercluster clusters. The maps indicate conservation of T7 RNA polymerase as well as DNA metabolism, structural, and cell lysis proteins. Phage maps were created by Phamerator, and boxes for gene products are labeled with predicted function, occasionally numbered, and colored manually to indicate similar function. Abbreviations are: host restriction inhibitor (Ocr), deoxyguanosine triphosphate (dGTP), deoxynucleotide monophosphate (dNMP), deoxyribonucleotide monophosphatase (dRMPase), single strand binding protein (SSB), endonuclease (Endo.), exonuclease (Exo.), polymerase (pol.), DNA ligase (ligase), major capsid protein (MCP), small terminase (TerS), large terminase (TerL).

Figure S4. Genes of phages ZF40 and øEt88 that are homologous and syntenic with phage lambda cluster genomes. ORF maps were created by Phamerator and colored manually. Predicted ORFs that are transcribed rightward and leftward are shown as boxes above and below the central

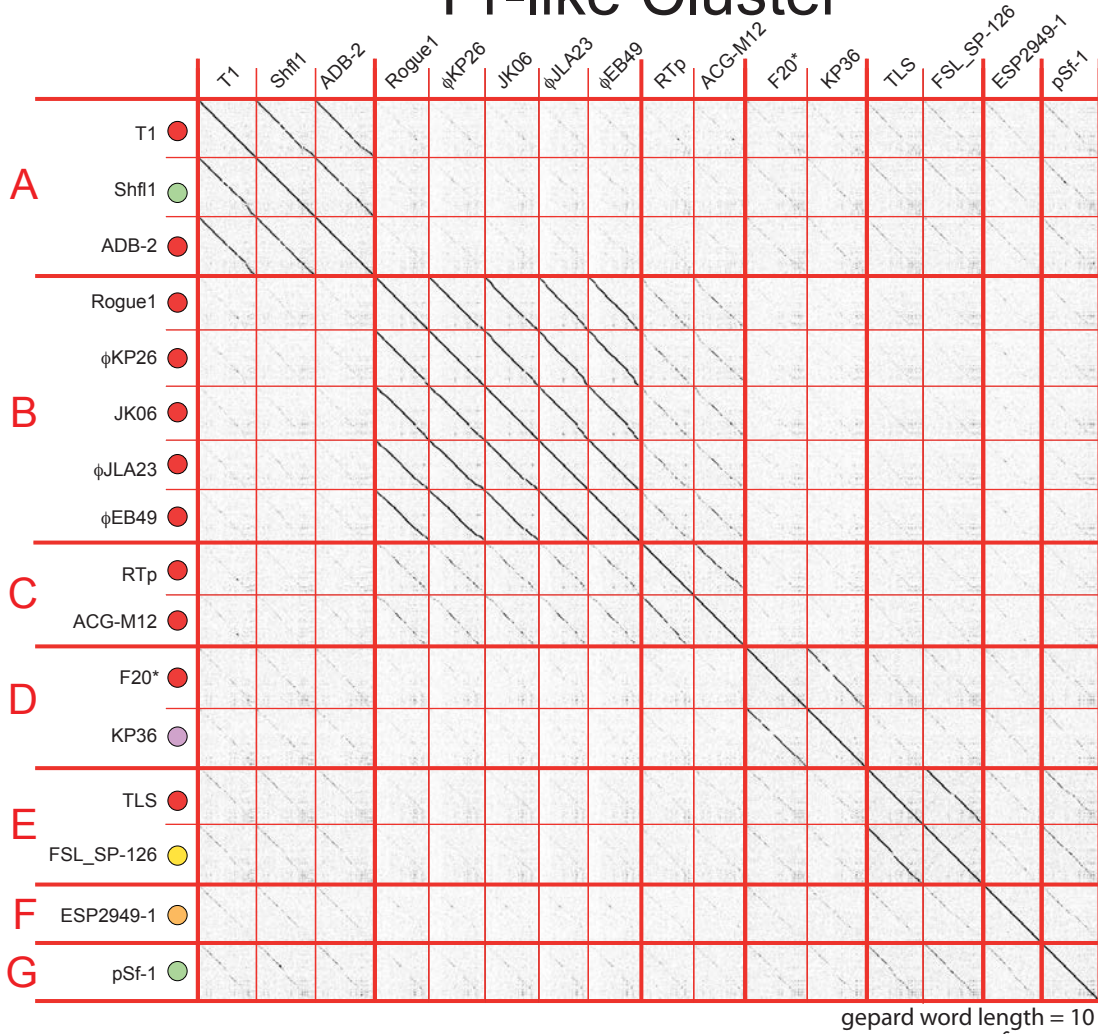
scale, respectively. The predicted encoded proteins were compared to the current protein database by blastp searches. Matches to other members of the lambda supercluster are indicated above each map with the percent AA identity and encoding phage of the best hits. These ORFs are colored yellow. Those few matches that are not syntenic (homologous genes lie in substantially different locations in the genome in these two phages than in other lambda supercluster phages) are marked by asterisks. Convincing matches to other phages outside this supercluster are indicated below the maps. These early region proteins have substantial syntenic homology with lambda cluster phages, and most of the lambda supercluster “non-matches” lie in the head and tail genes. We conclude that both ZF40 and øET88 are lambda supercluster members with unique sets of head and tail genes.

Figure S5. Neighbor-joining tree of the MCPs of the currently known members of the lambda supercluster. Clustal X was used to create the tree, and the nineteen major branches (MCP major sequence types) are indicated by colored boxes. The cluster to which the encoding phages belong is shown on the right. Note that lambda, ø80, gfsy-1 and N15 MCPs form four more closely related subtypes within the same branch. In addition, the MCPs of phage mEp235, FSL_SP-016, CUS-3/SPC-P1, Sf6/HK620 and SfV are very different from those of other members of their cluster. Finally, we note that the Sf6/HK620 MCP type, while being encoded by phages in the P22-like cluster, are quite similar to APSE-1-like phage MCPs.

Figure S6. Examples of horizontal exchange within the lambda supercluster. Nucleotide sequence dot plots created by DNA Strider with a scanning window stringency of 15/23 are shown. The genomes are all oriented so that *terS* is the leftmost gene and the rest of the late operon virion assembly genes at its immediate right. The late operon is indicated by black arrows and the early left and right operons are indicated by gray arrows. Phages SPN3UB and FSL_SP-016 have large genomic sections that are similar to different clusters within the lambda supercluster. The SPN3UB plots show that its late region is largely similar to ES18 and its early region is largely similar to FSL_SP-016.

Figure S7. Examples of mosaicism among phages in the HK97-like cluster. Nucleotide sequence dot plots created by DNA Strider with a scanning window stringency of 15/23 are shown. The genome orientation is the same as in figure S5, and gene functions and the location of the MCP gene are indicated above and below. Phage mEp253 has head genes that are unique and very different from HK97 and the rest of the cluster (see text), while HK542 and mEpX1 have virion assembly genes very similar to HK97 and early regions that show different mosaic relationships with HK97.

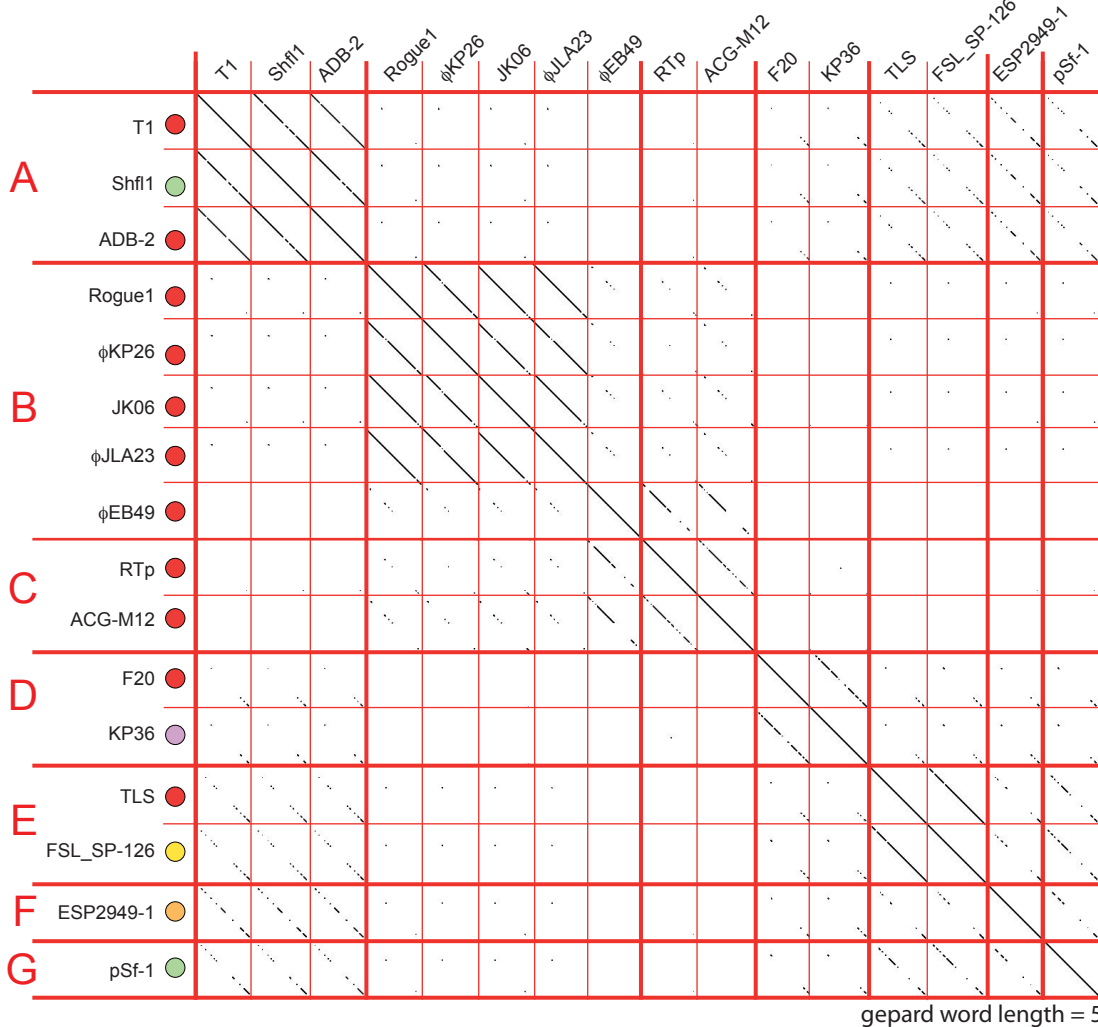
T1-like Cluster



Genomes

- *Coronobacter*
- *Escherichia*
- *Klebsiella*
- *Salmonella*
- *Shigella*

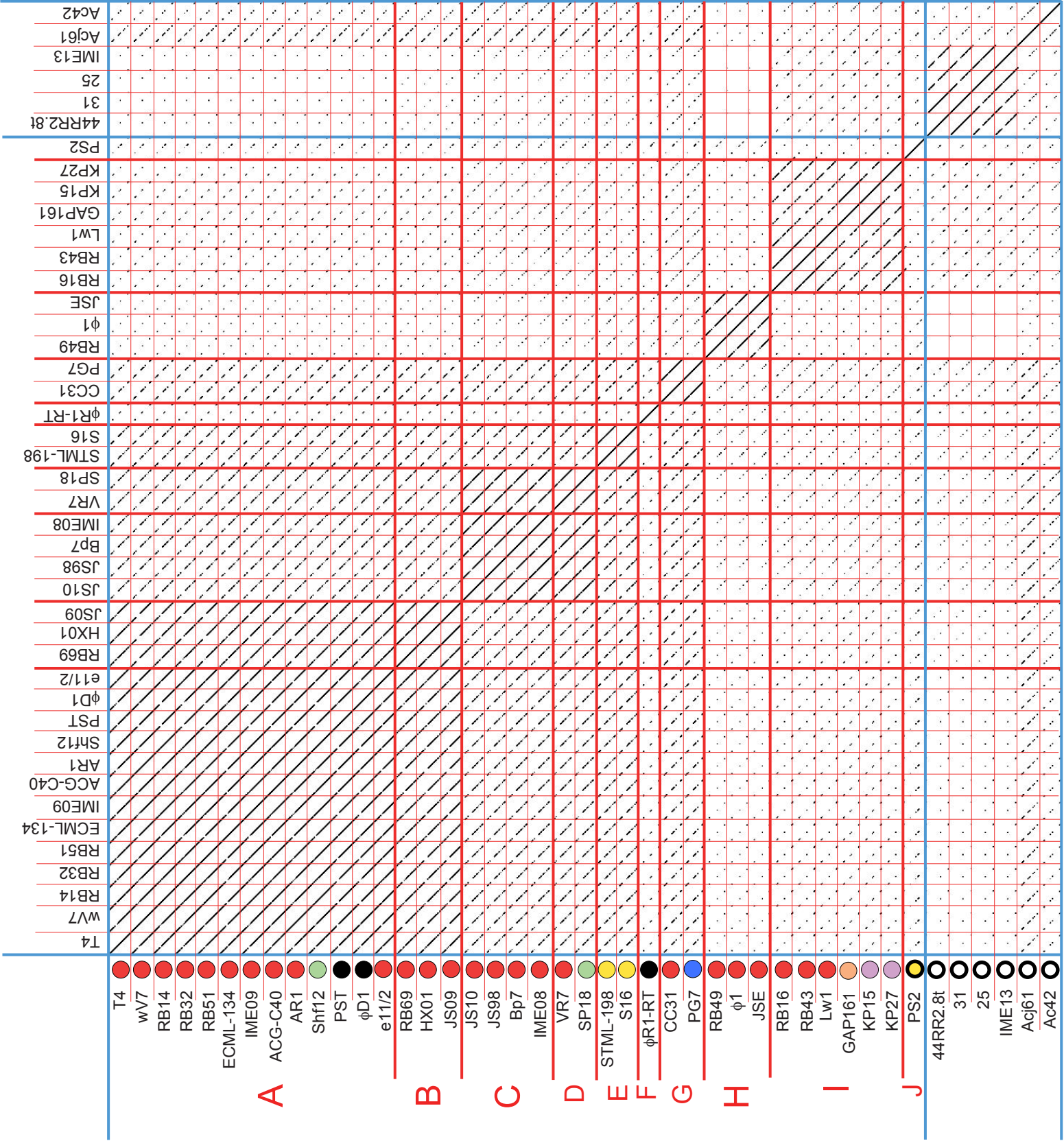
* Sequence not complete



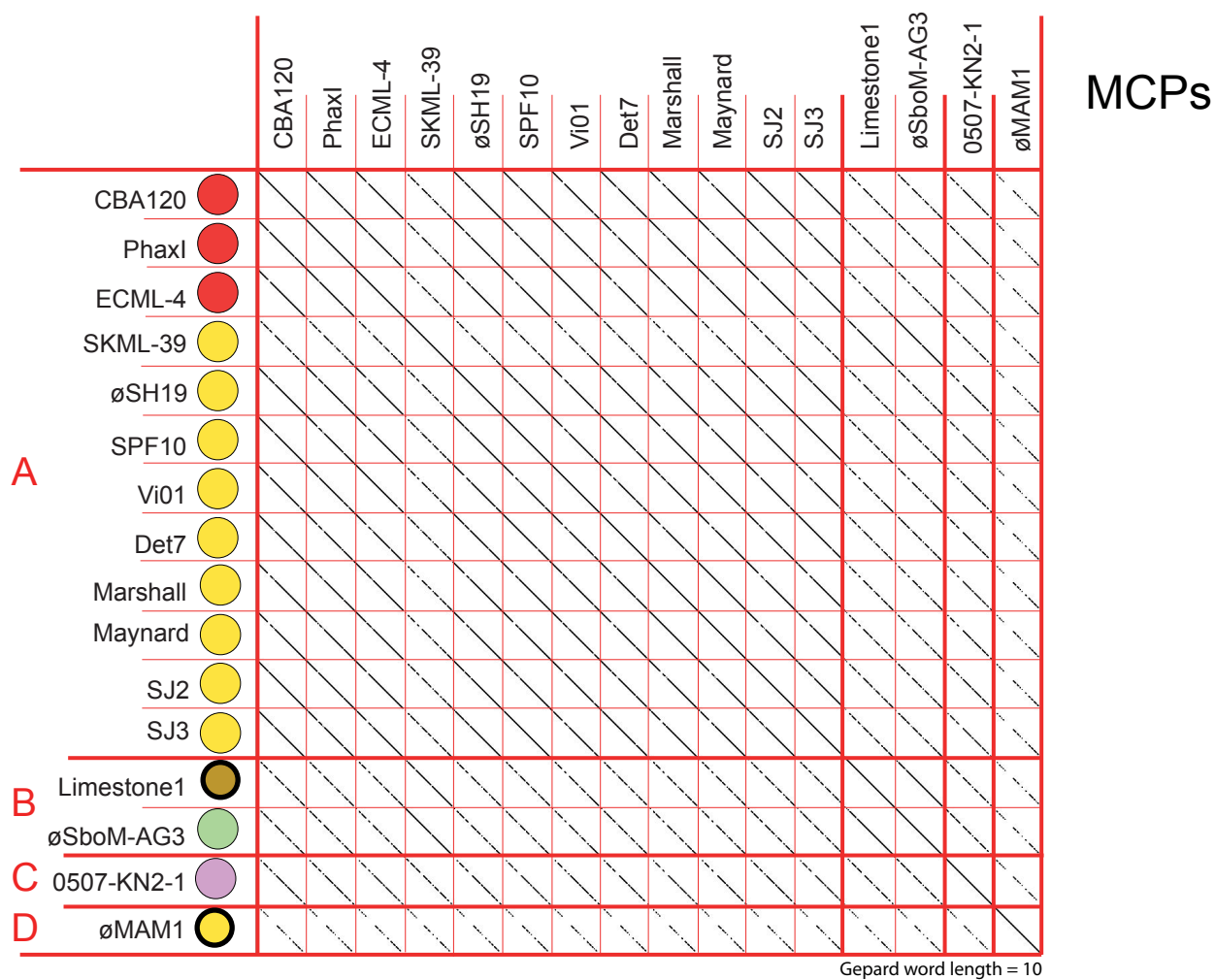
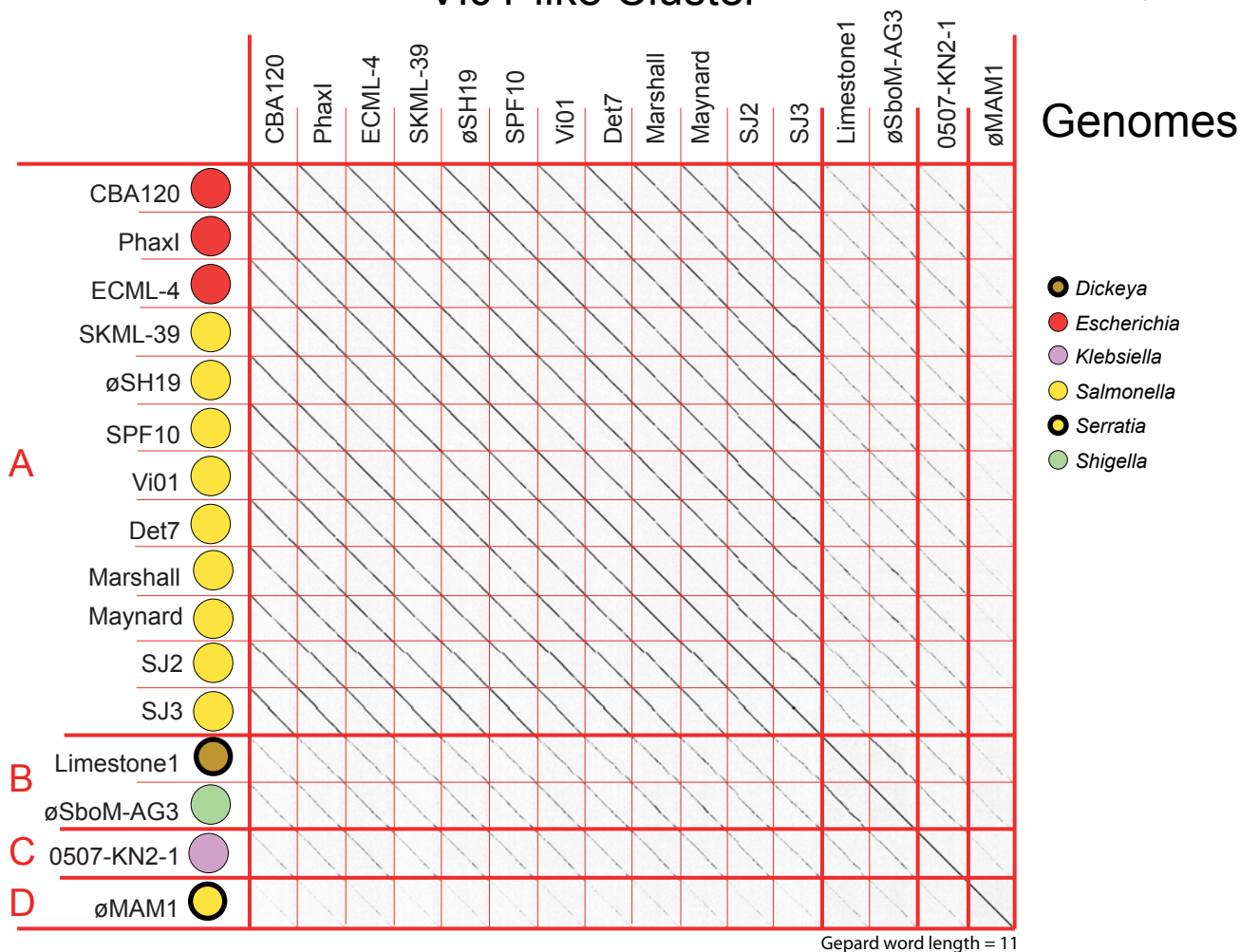
MCPs

T4-like Cluster MCPs

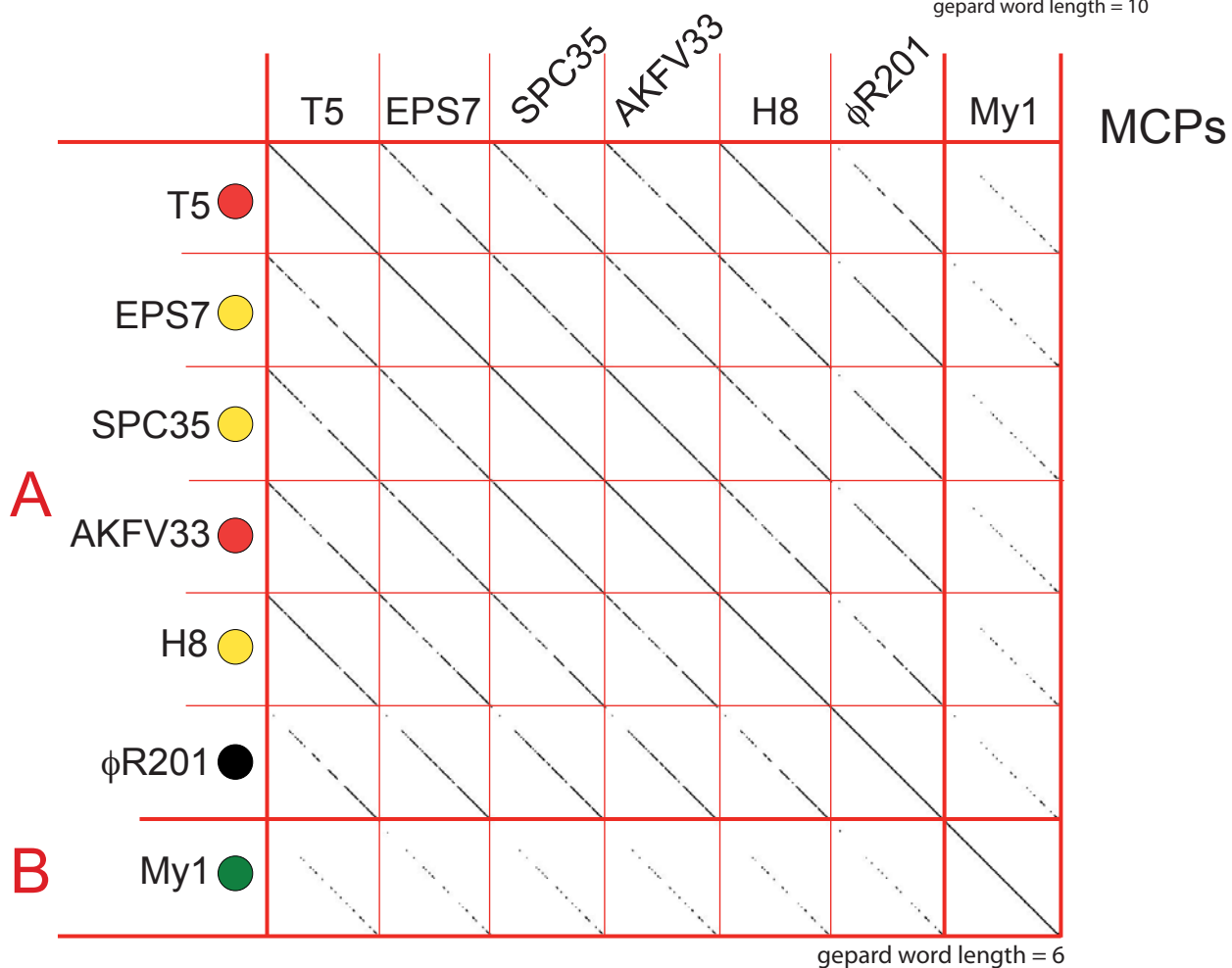
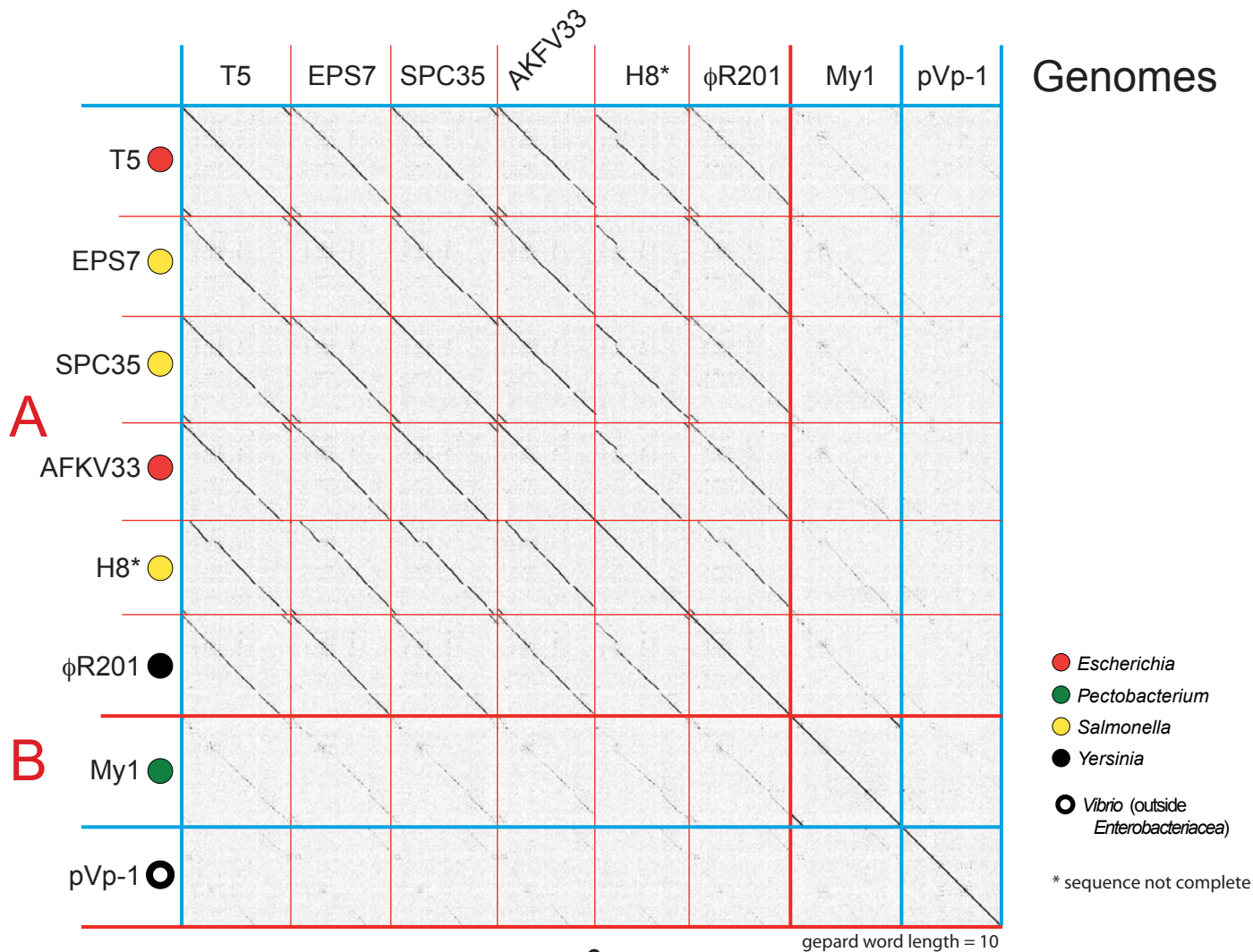
- *Coronobacter*
- *Enterobacter*
- *Escherichia*
- *Klebsiella*
- *Salmonella*
- *Serratia*
- *Shigella*
- *Yersinia*
- outside the
- *Enterobacteriaceae*
- *Acinetobacter*
- *Sterotrophomonas*
- *Aeromonas*



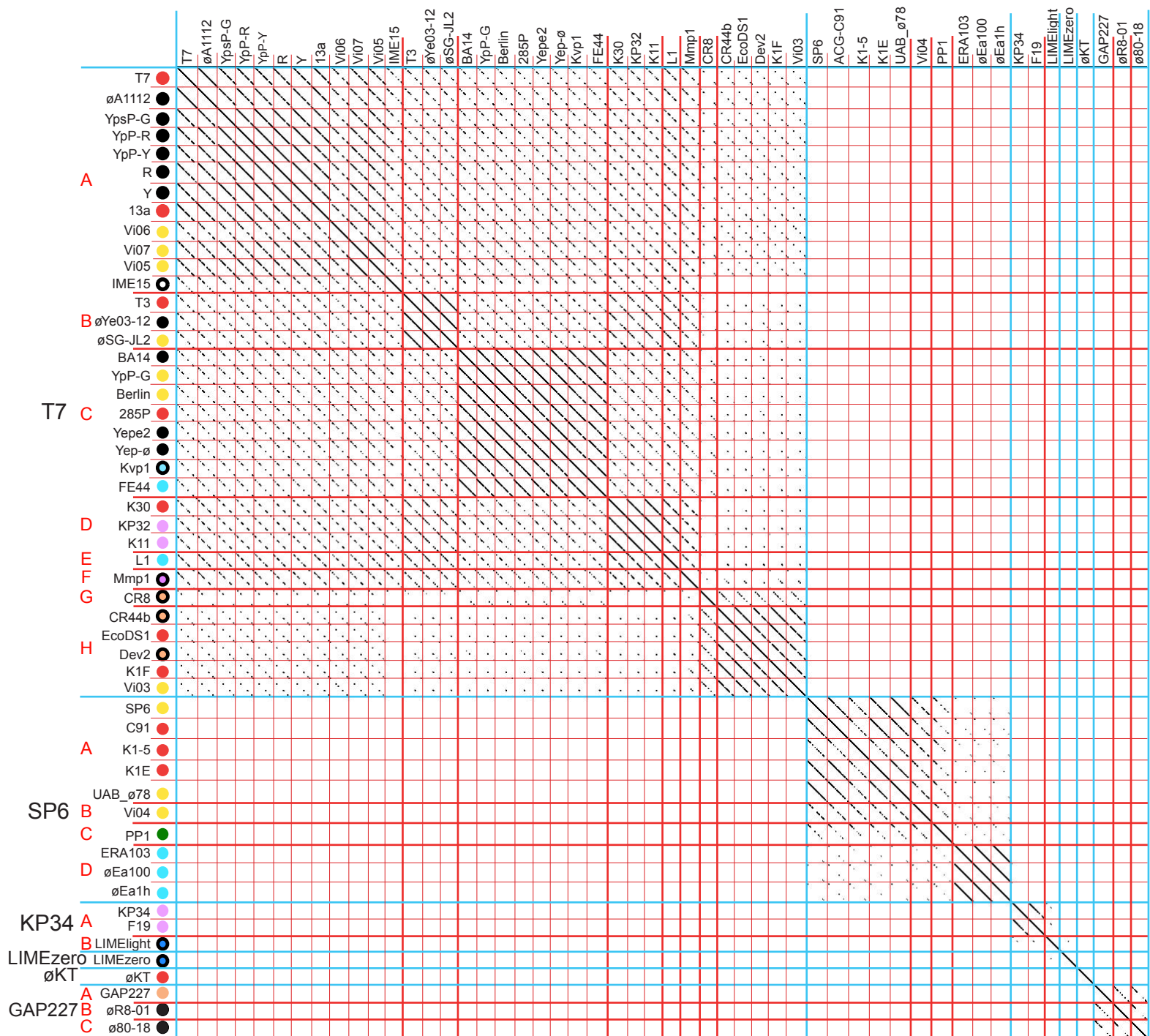
Gepard word length = 10



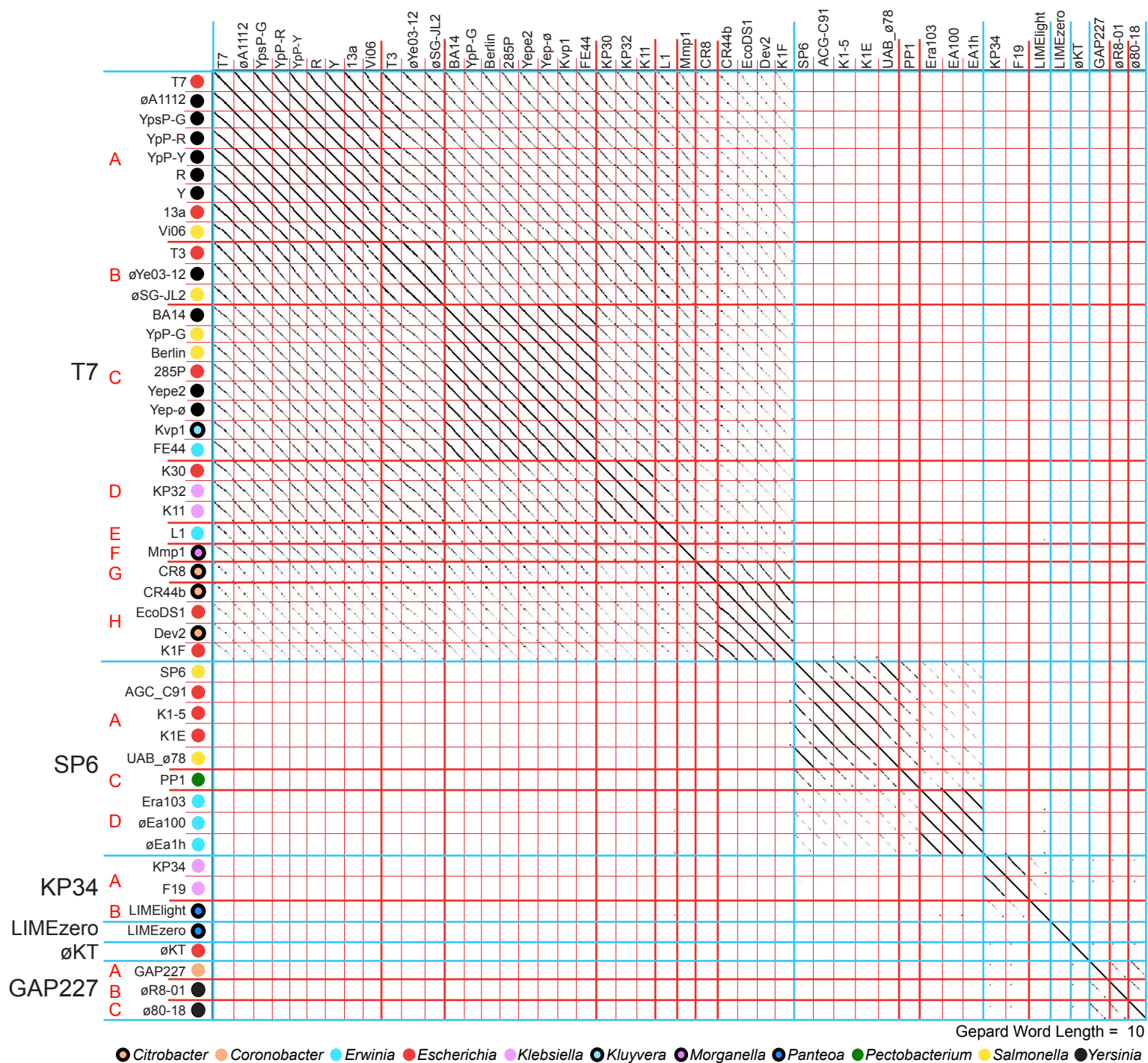
T5-like Cluster



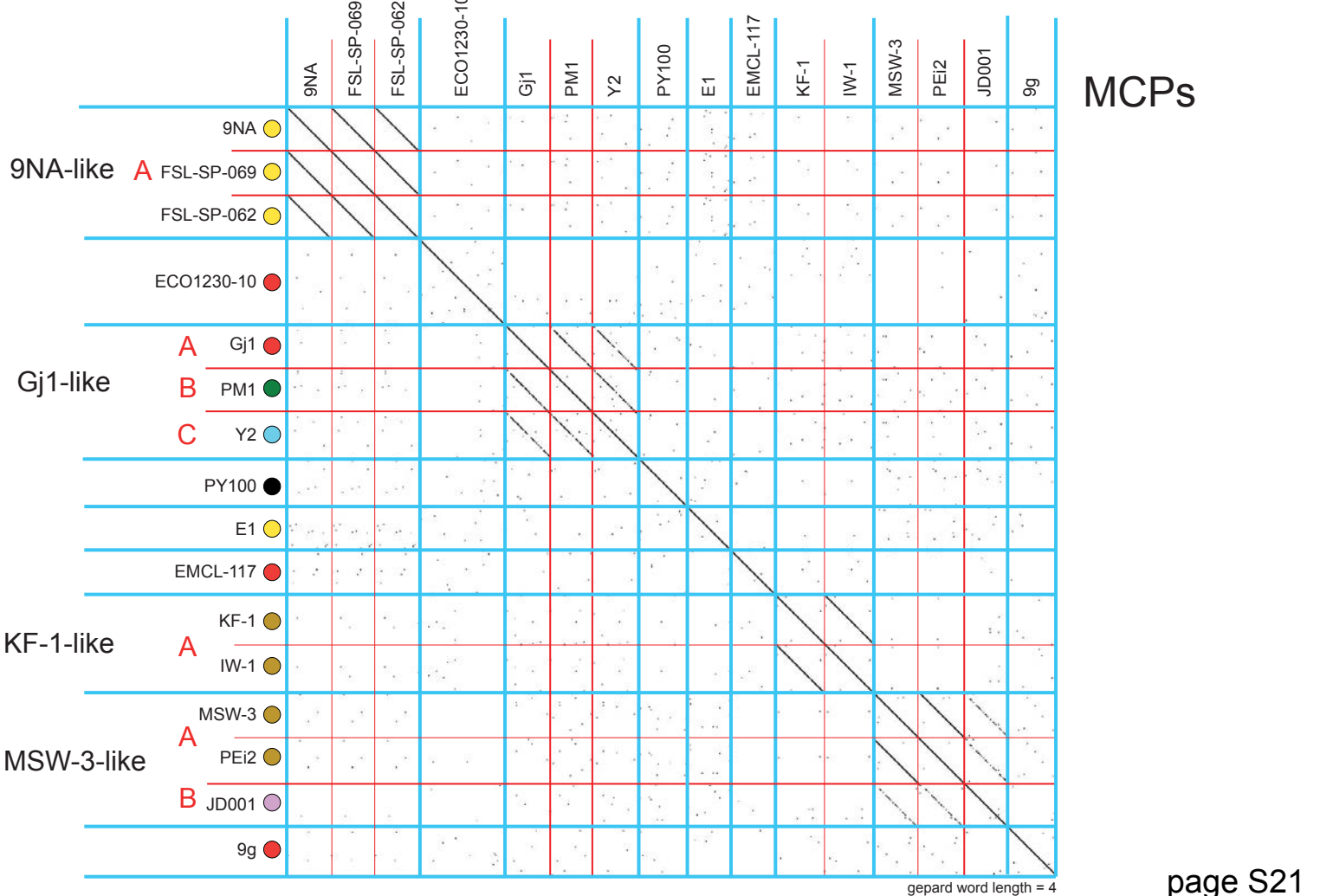
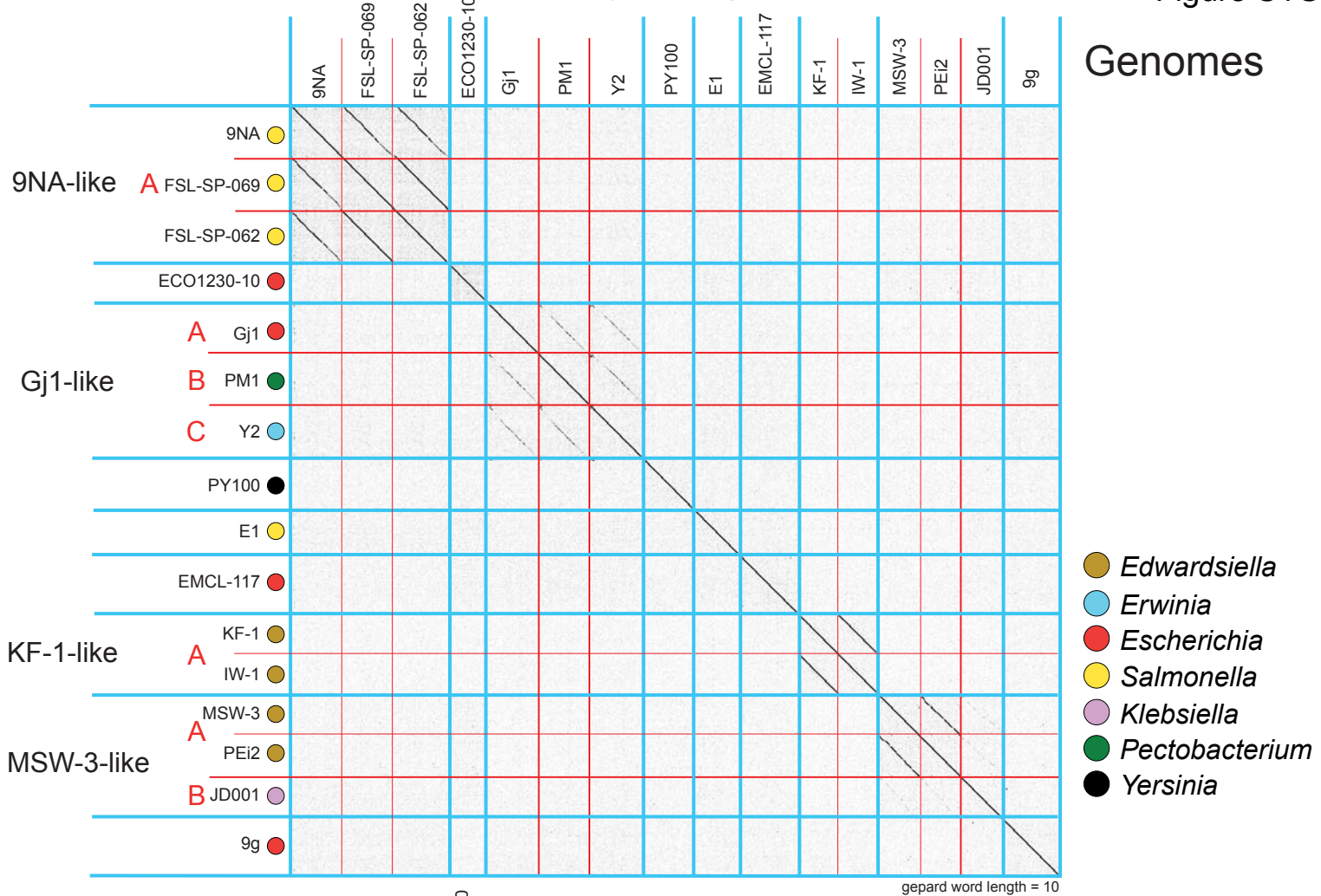
T7 MCP dot plot



T7 whole gene product dot plot

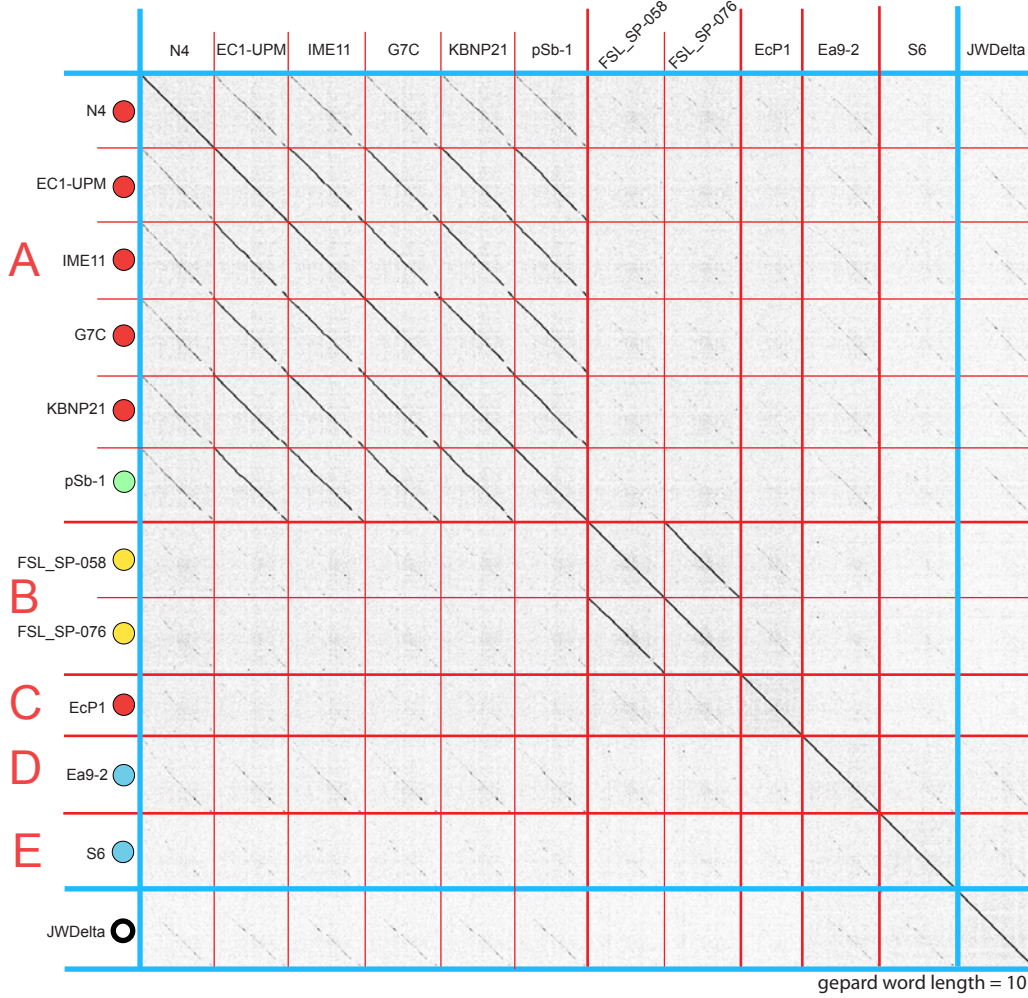


Nine Small Lytic Phage Clusters



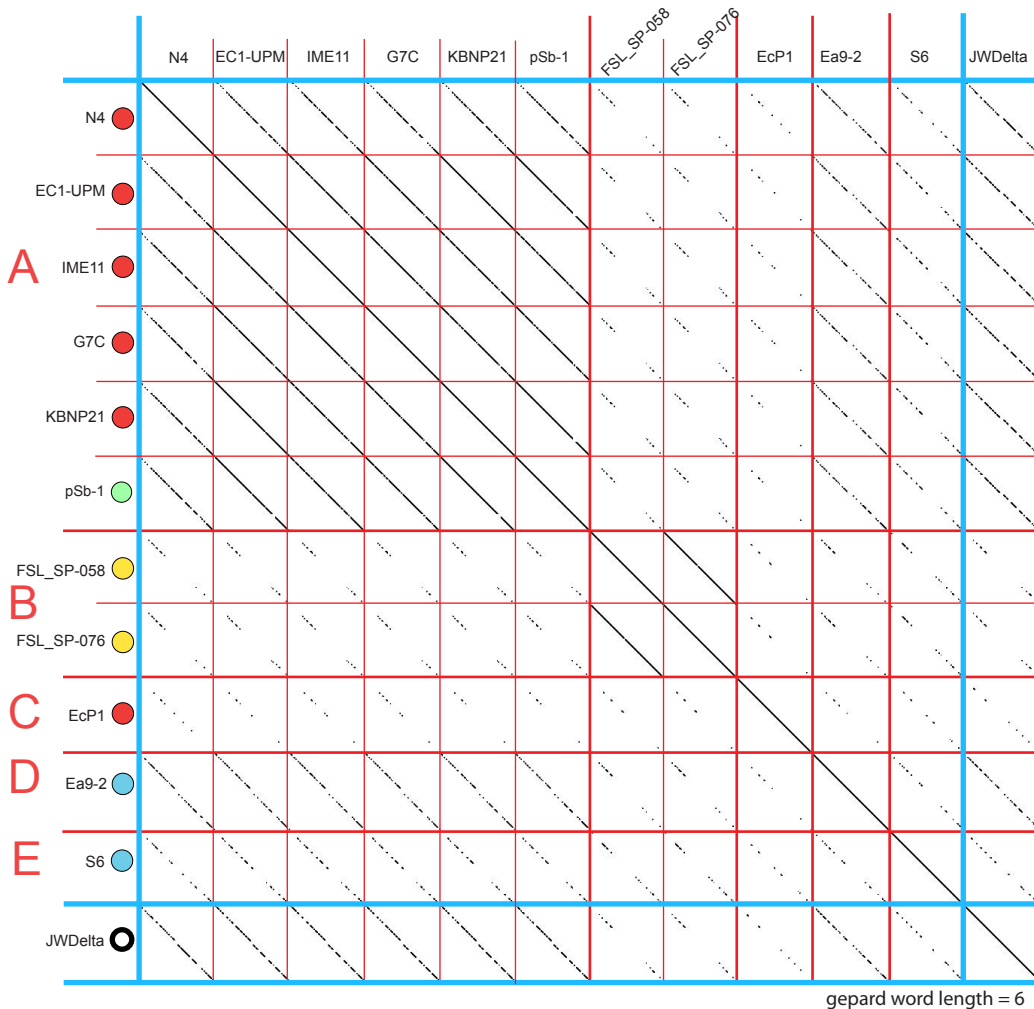
N4-like Cluster

Genomes

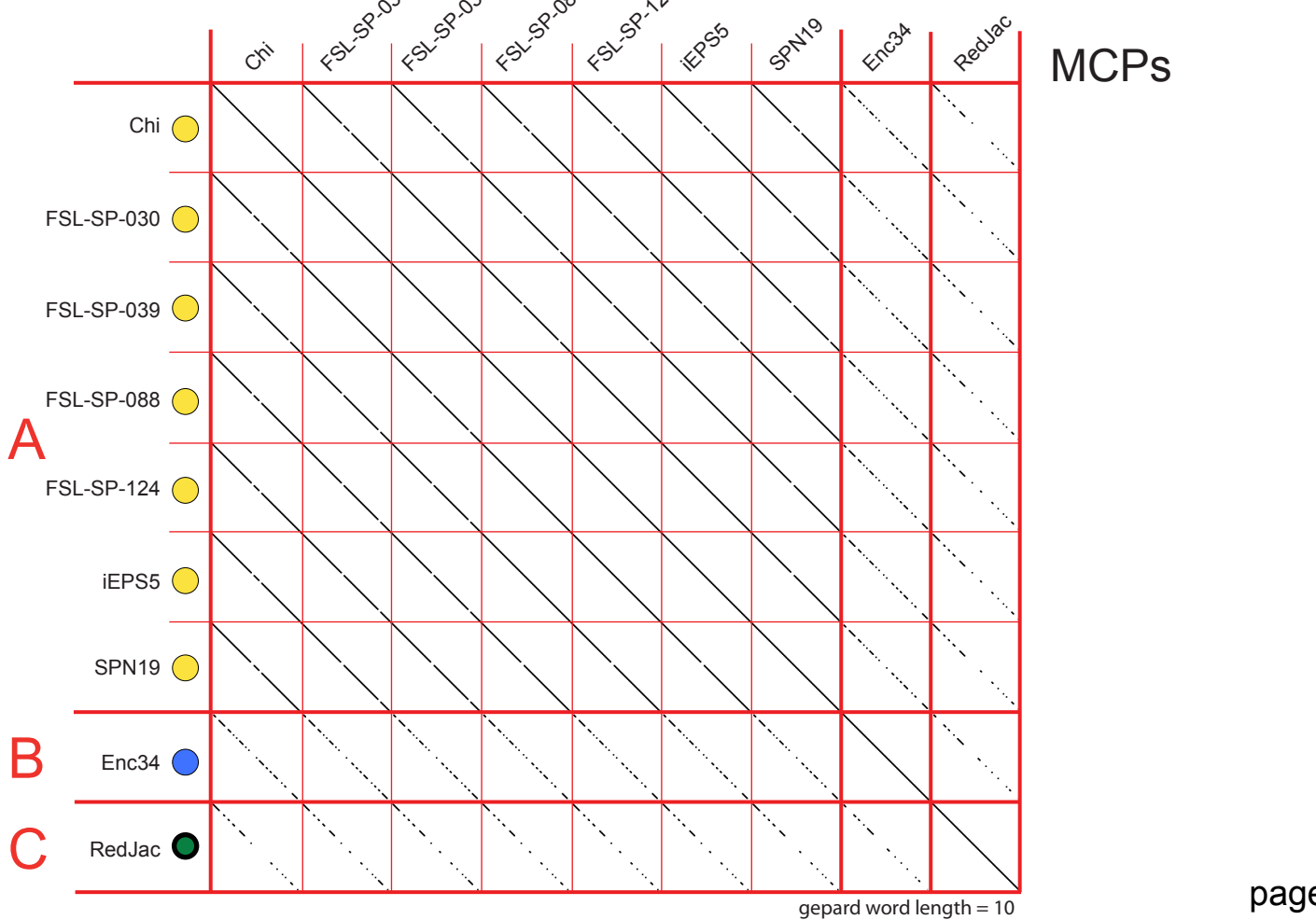
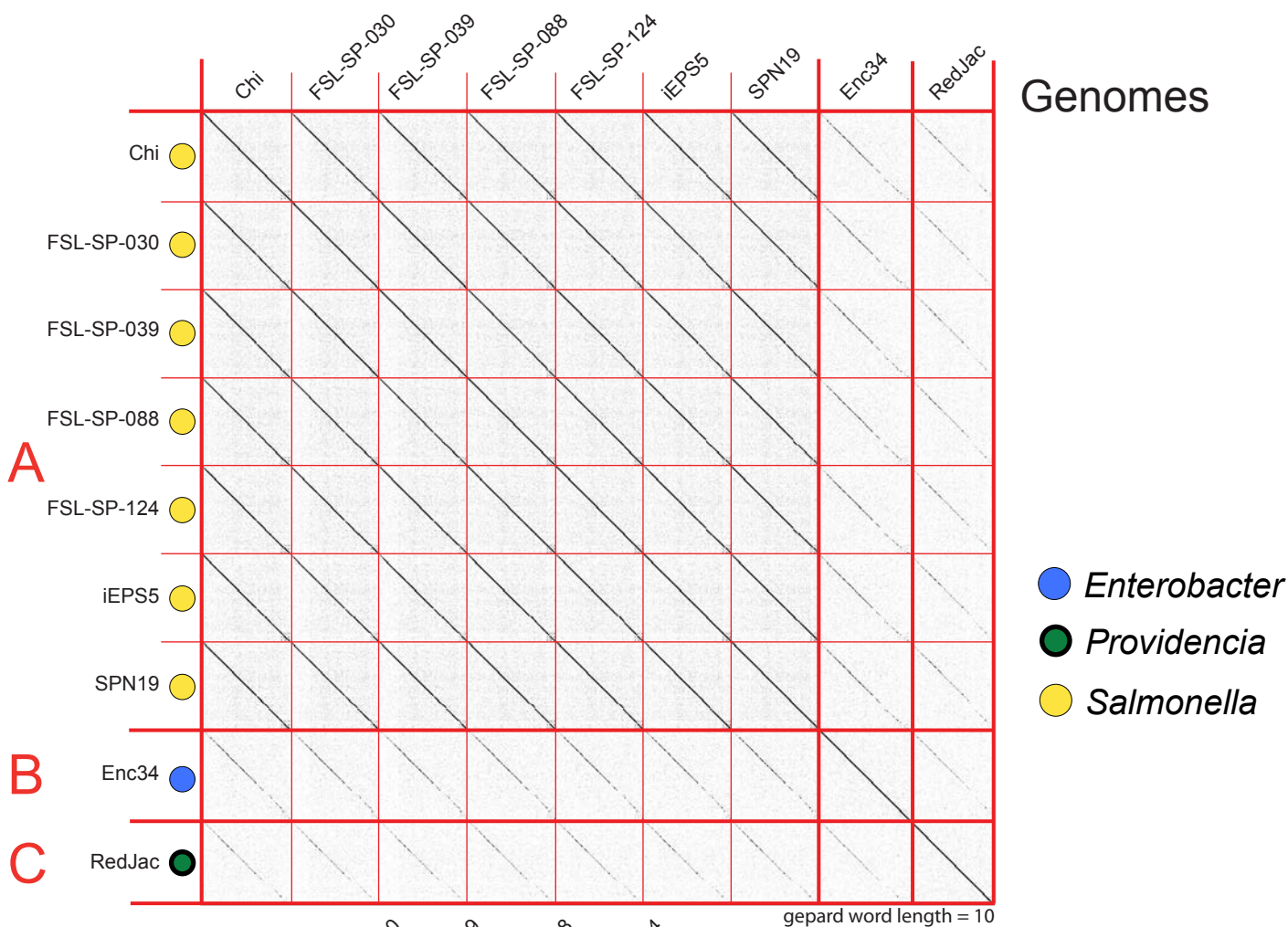


- *Erwinia*
- *Escherichia*
- *Salmonella*
- *Shigella*
- Outside the *Enterobacteriaceae* *Achromobacter*

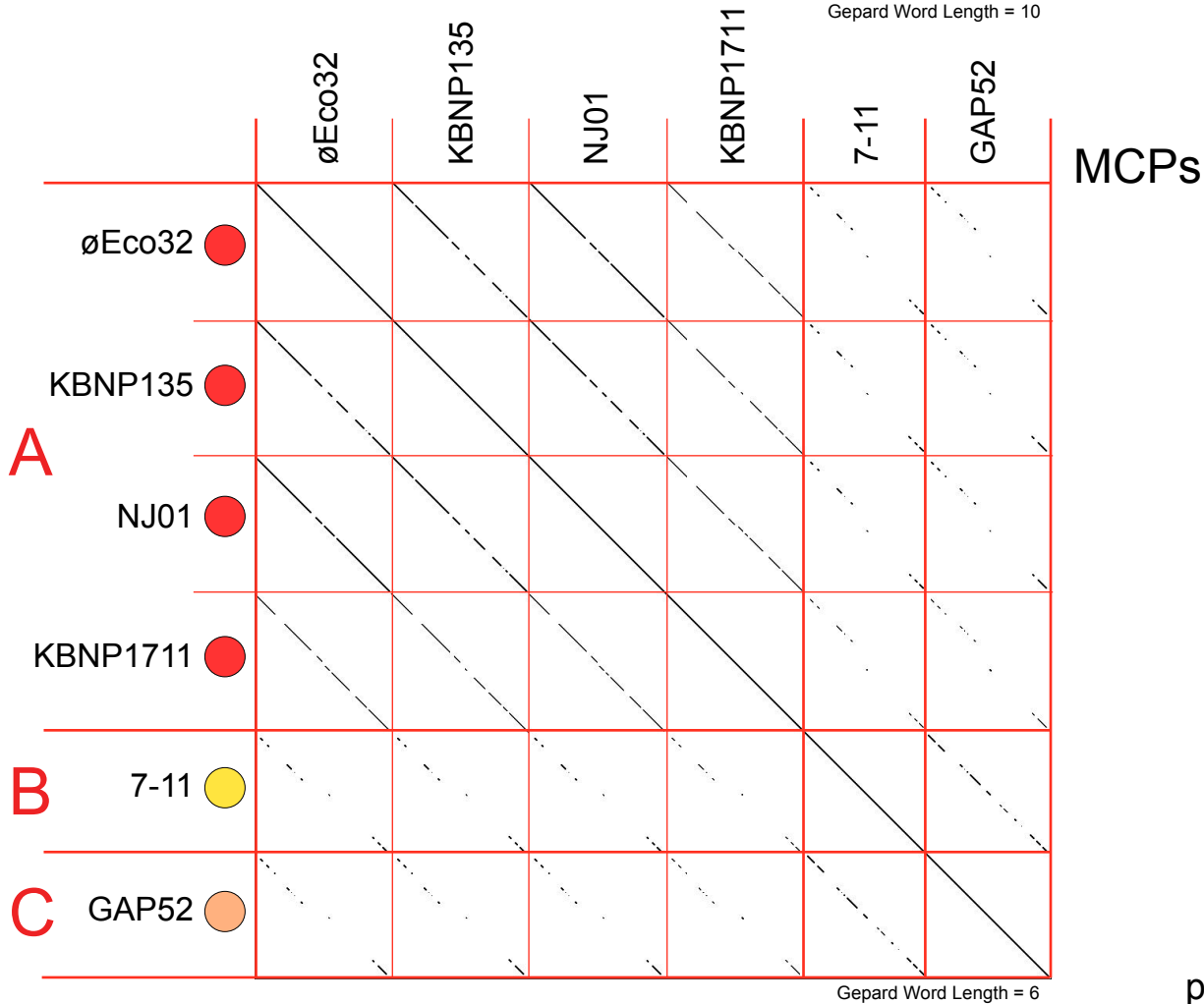
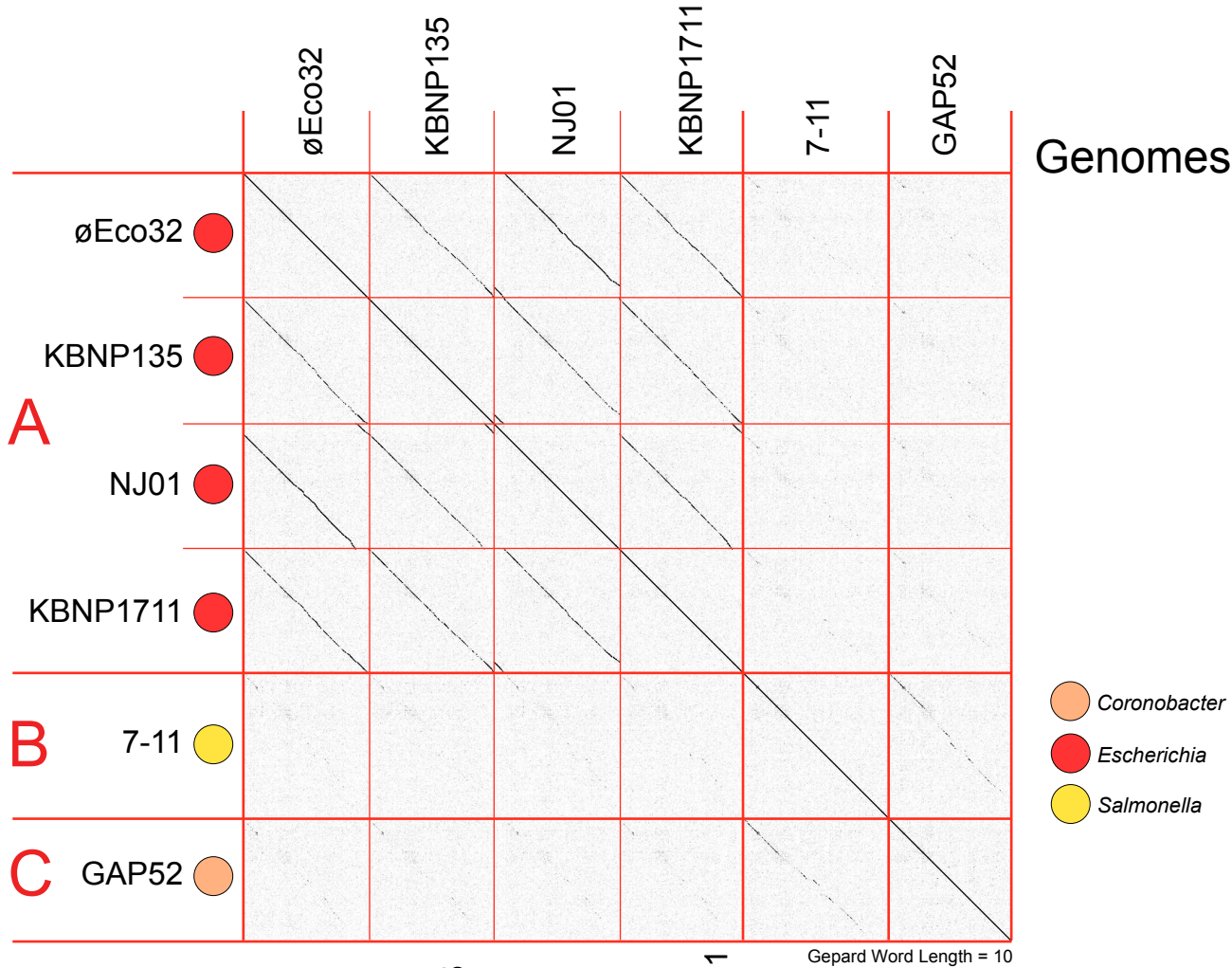
MCPs



Chi-like Cluster



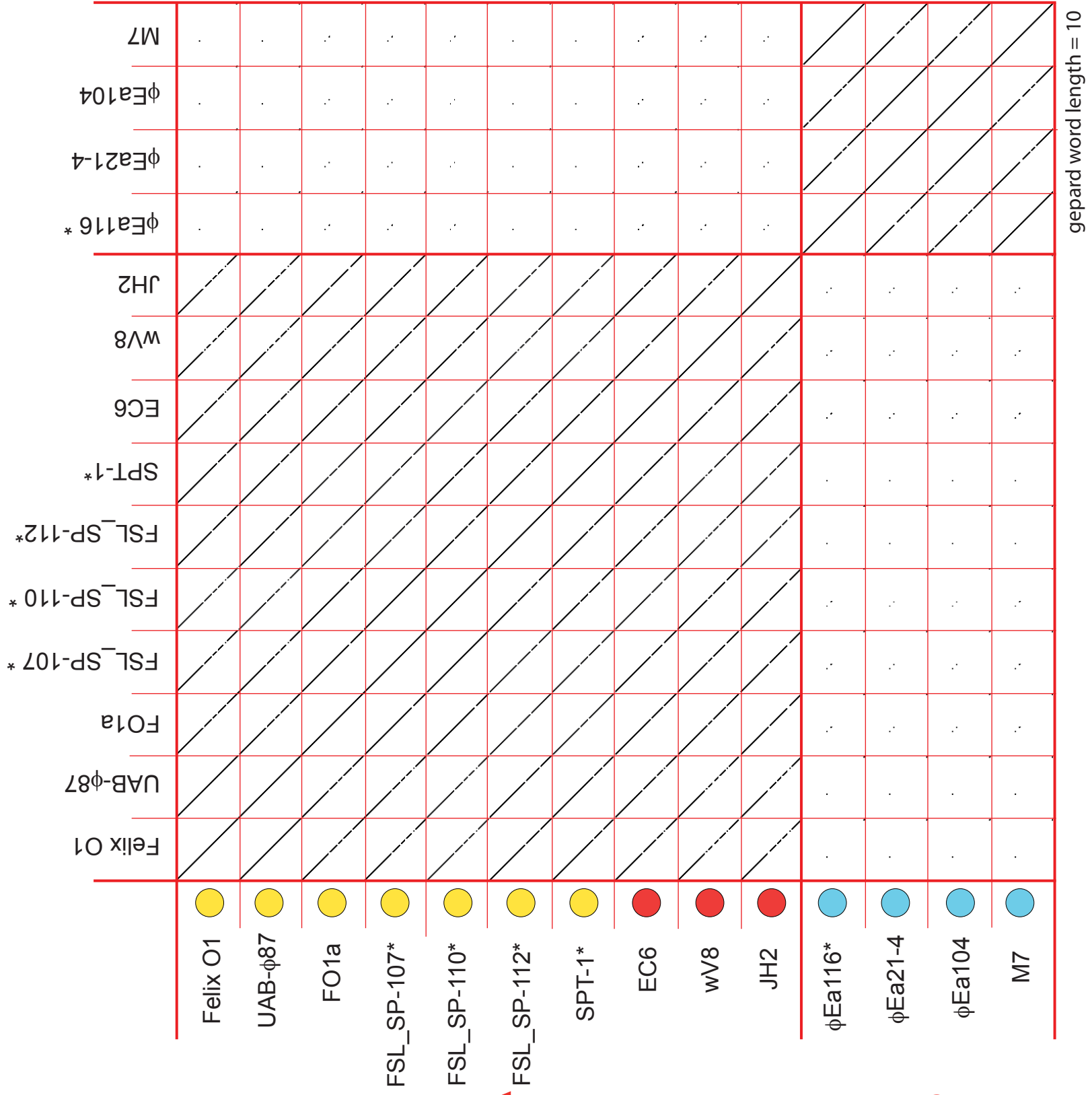
øEco32-like Cluster



Felix O1-like
Cluster
MCPs

- *Erwinia*
- *Escherichia*
- *Salmonella*

* genome not completely sequenced

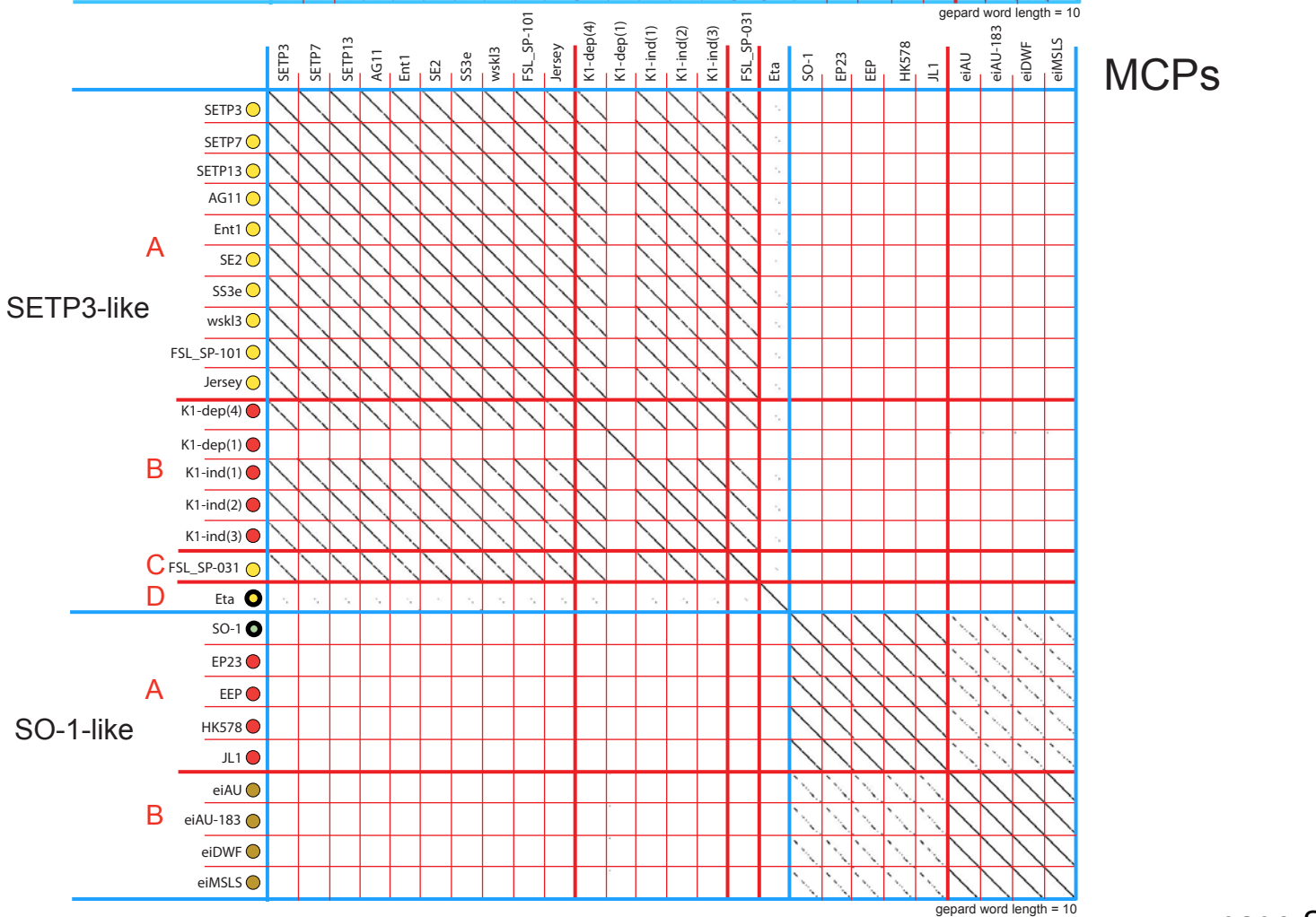


gepard word length = 10

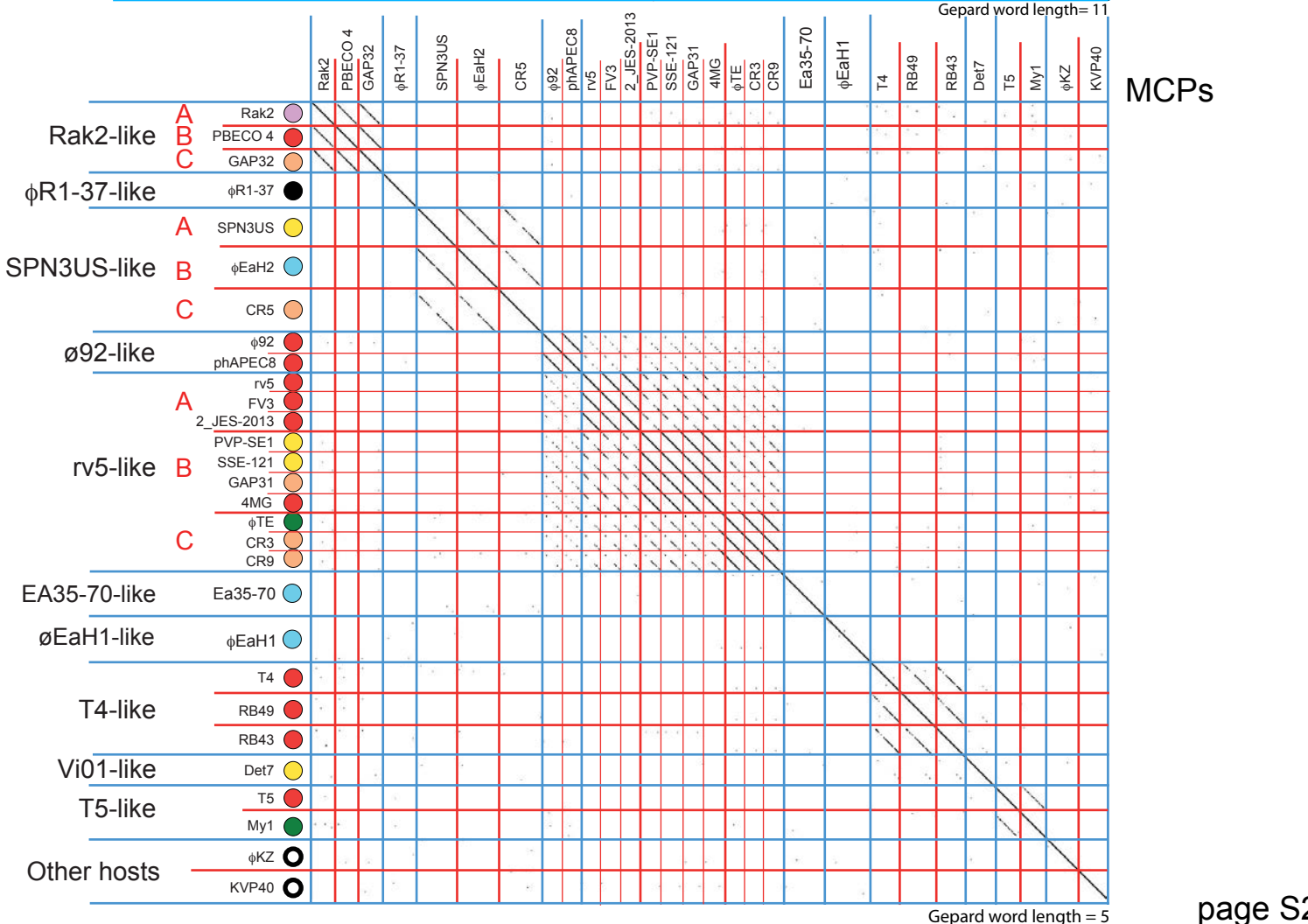
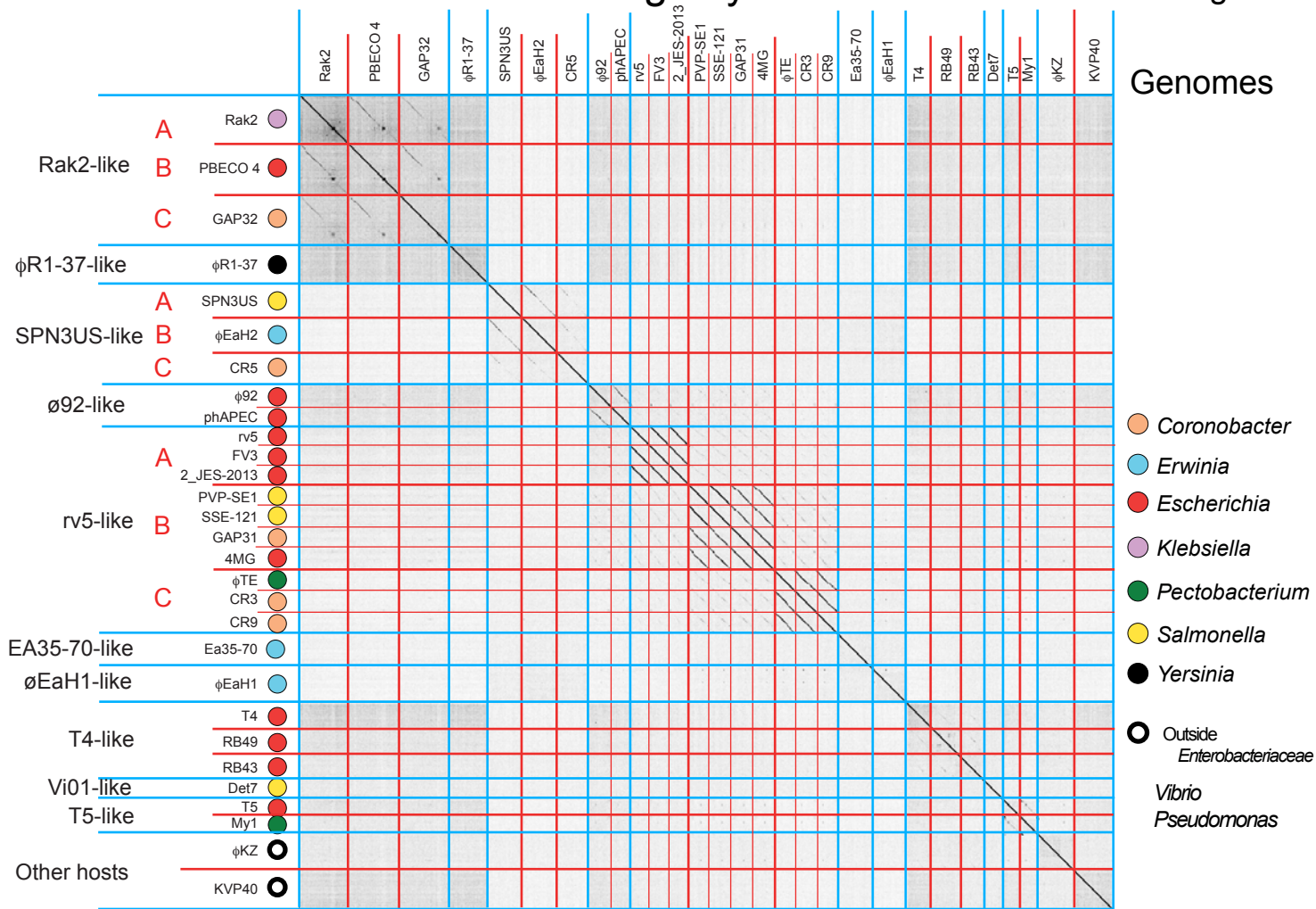
A

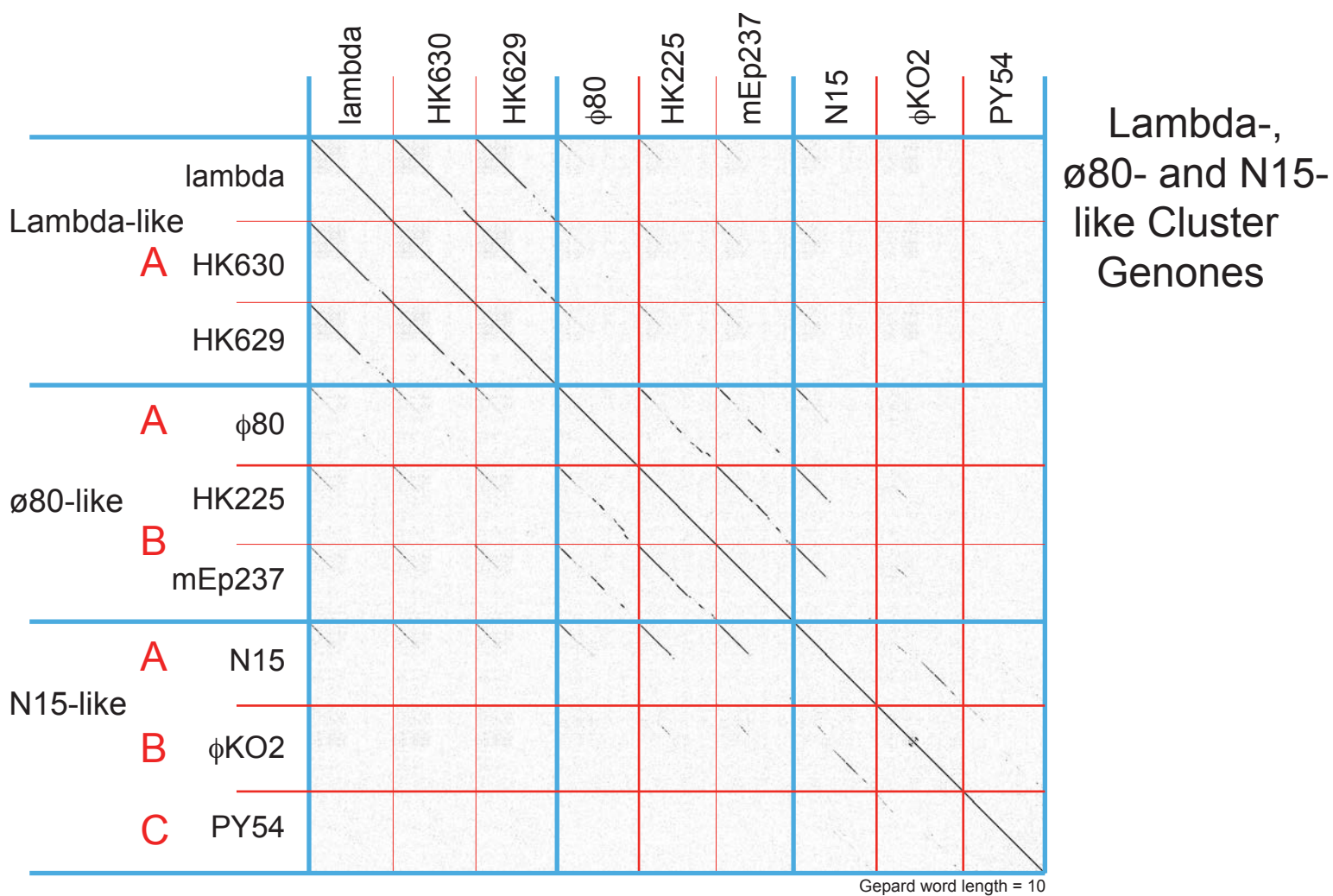
B

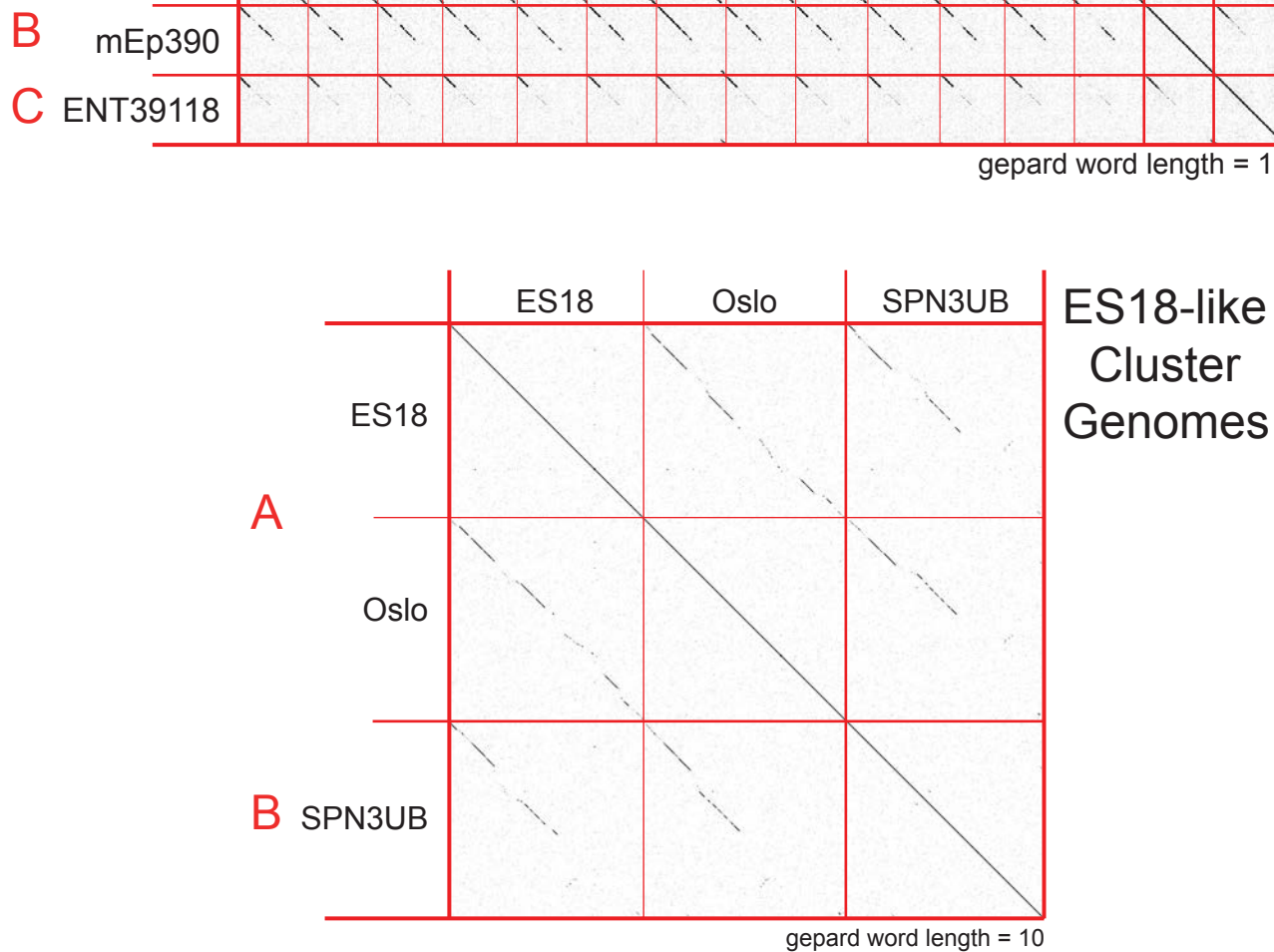
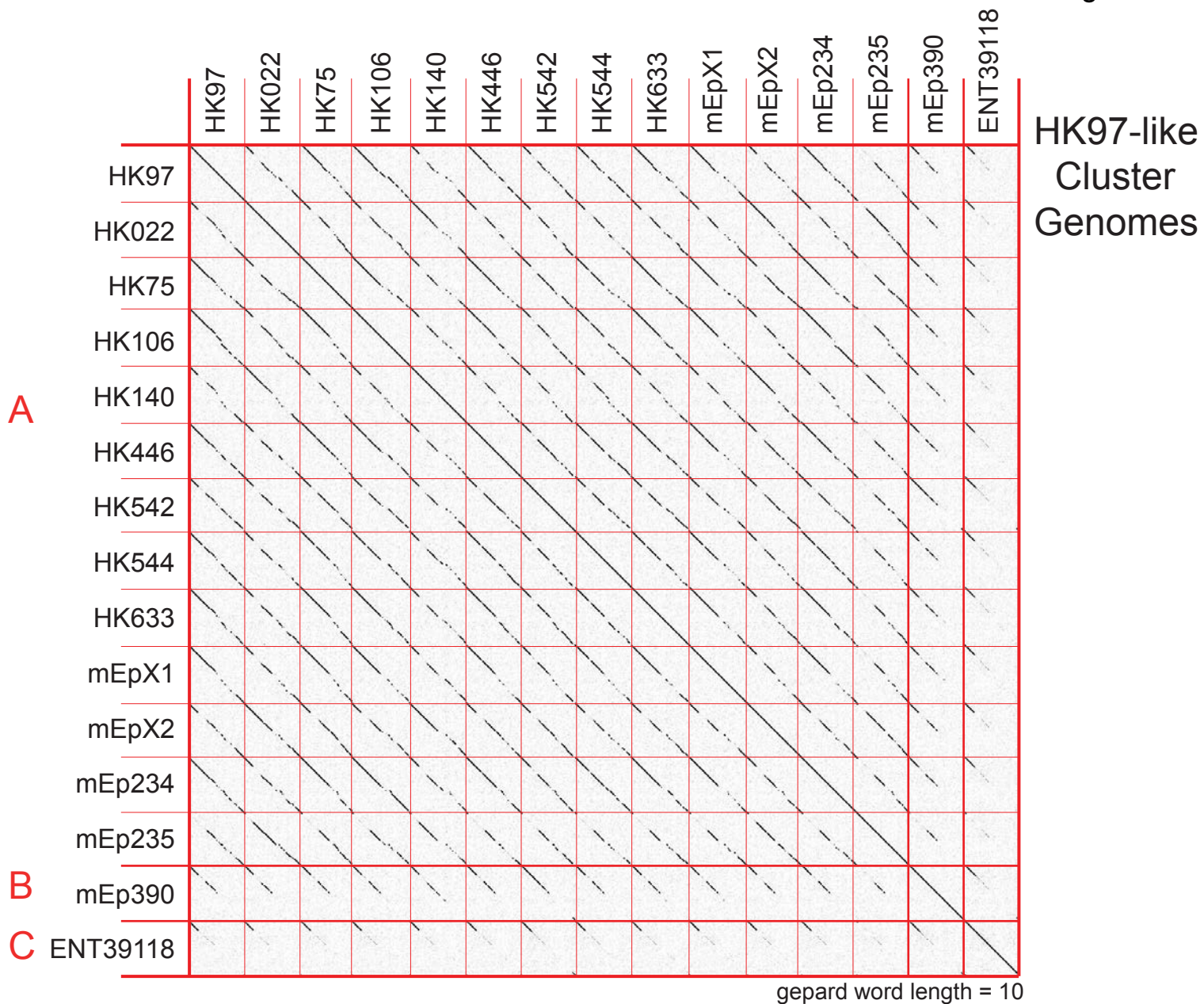
SETP3-like and SO-1-like Clusters

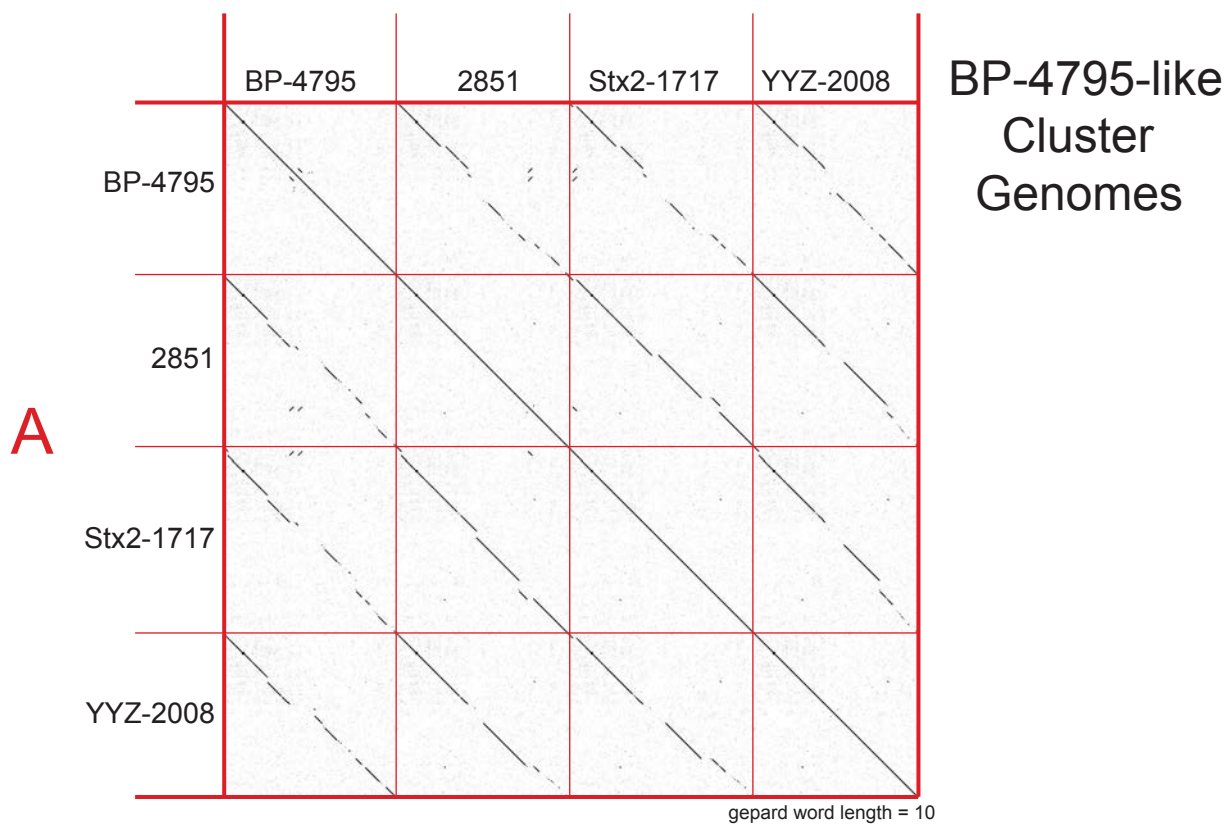
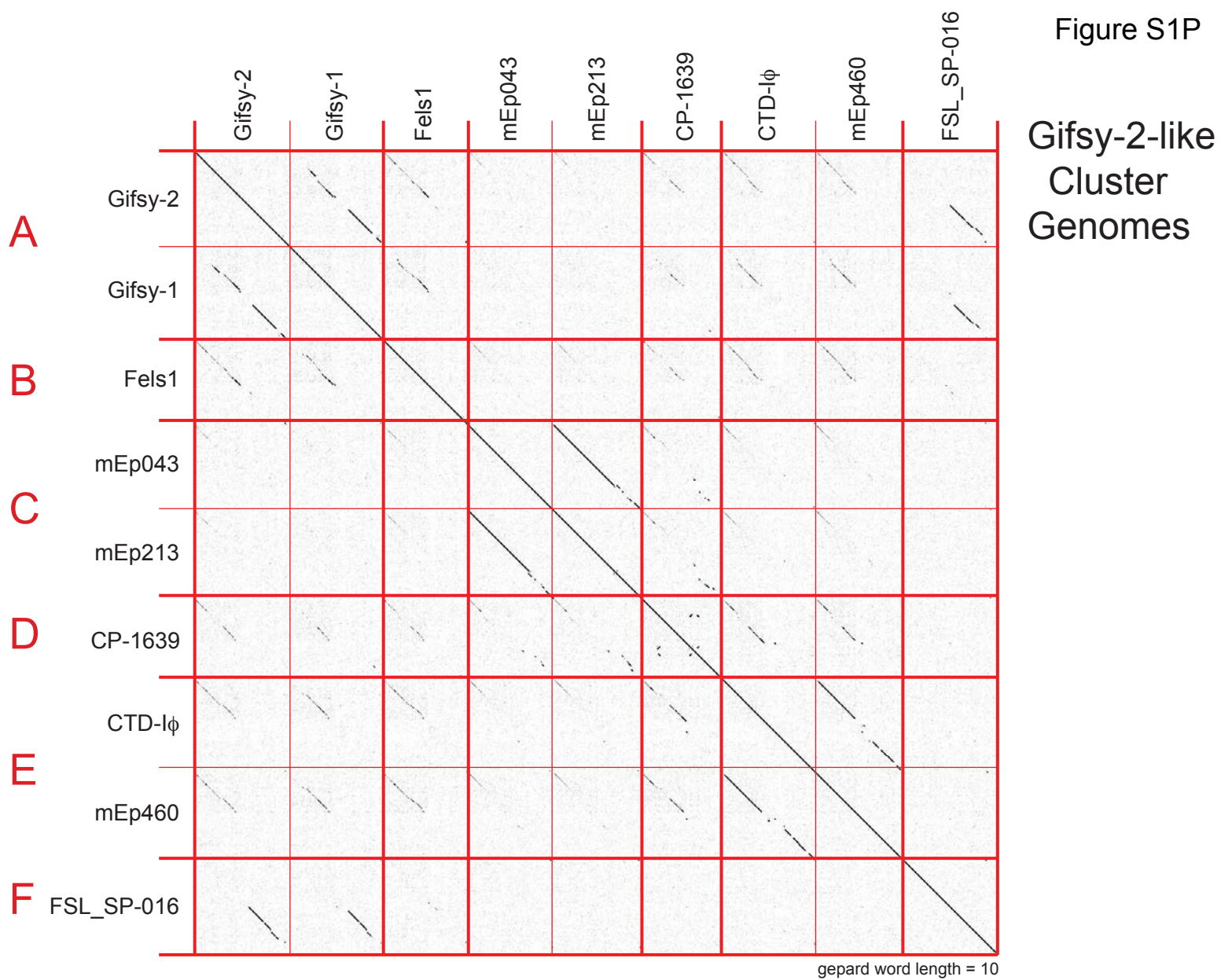


Seven Large Lytic Clusters

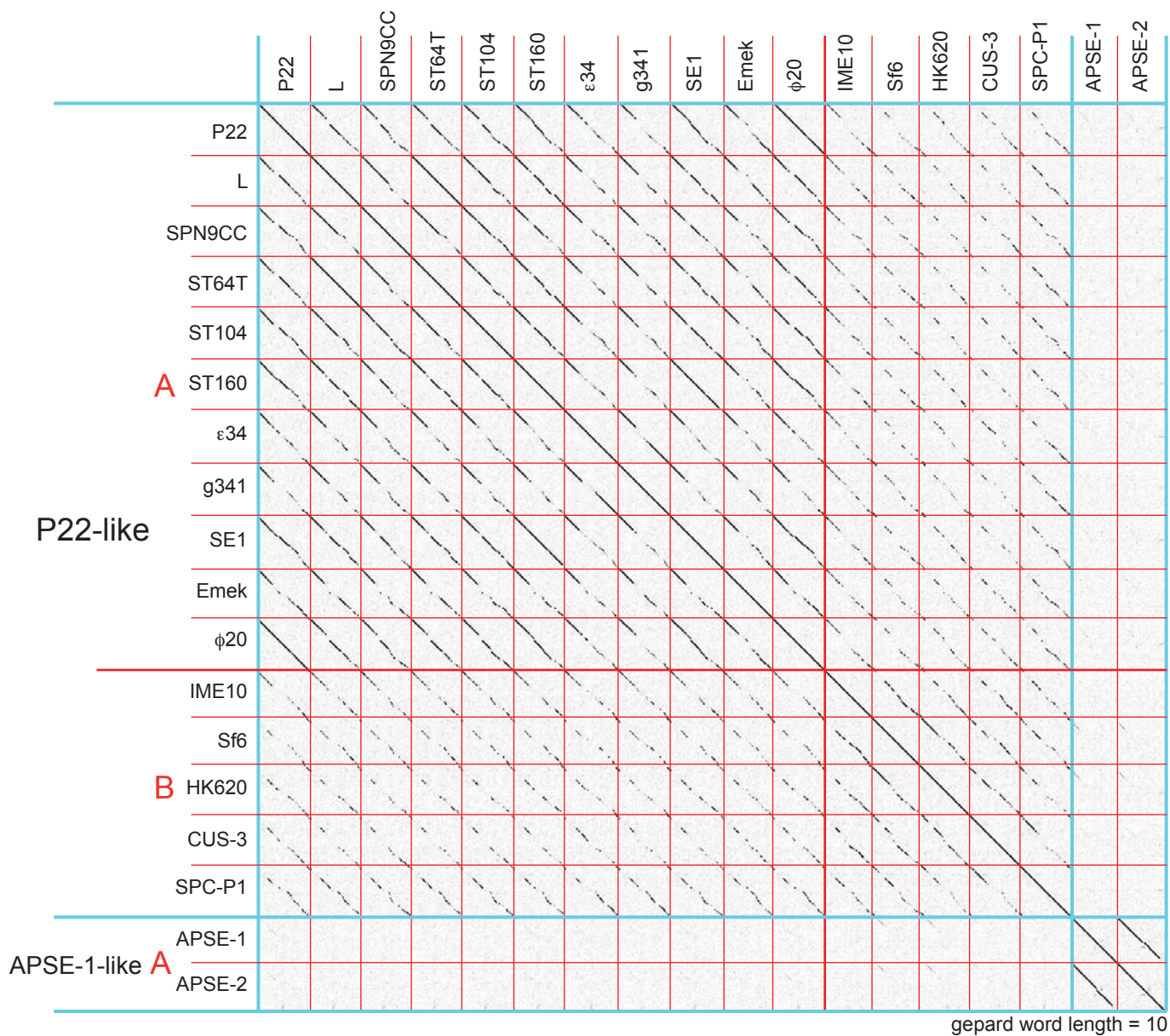


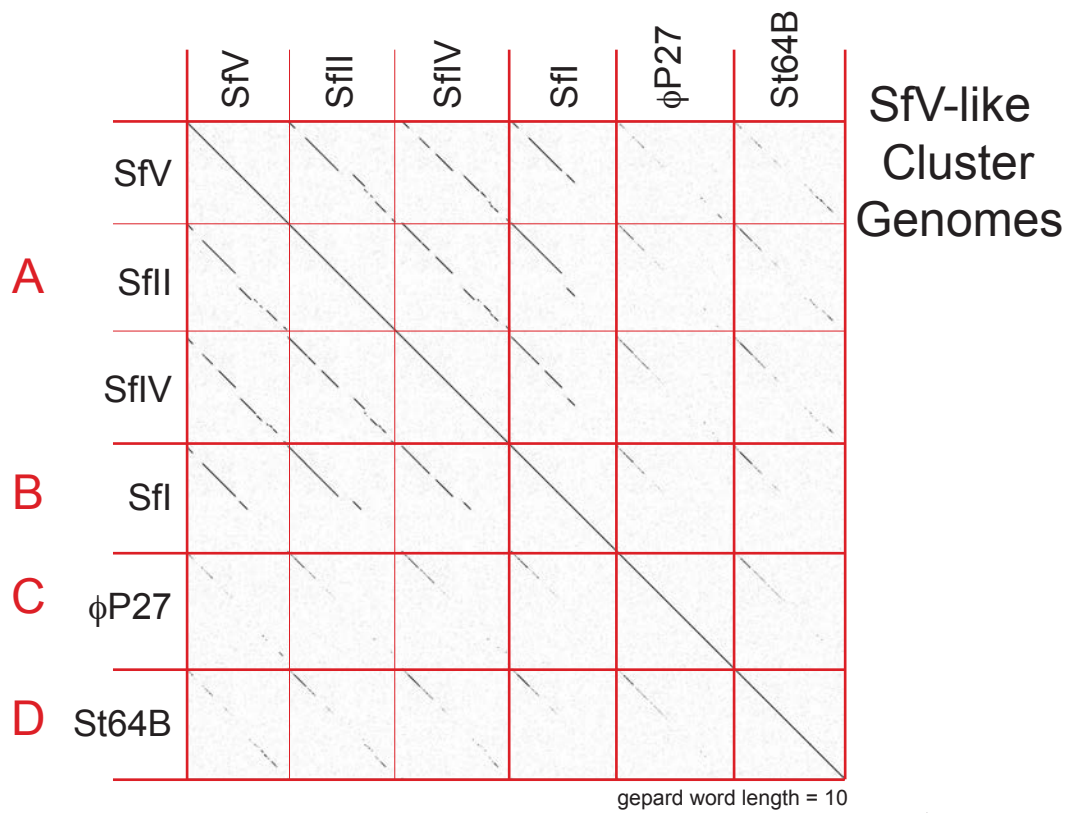




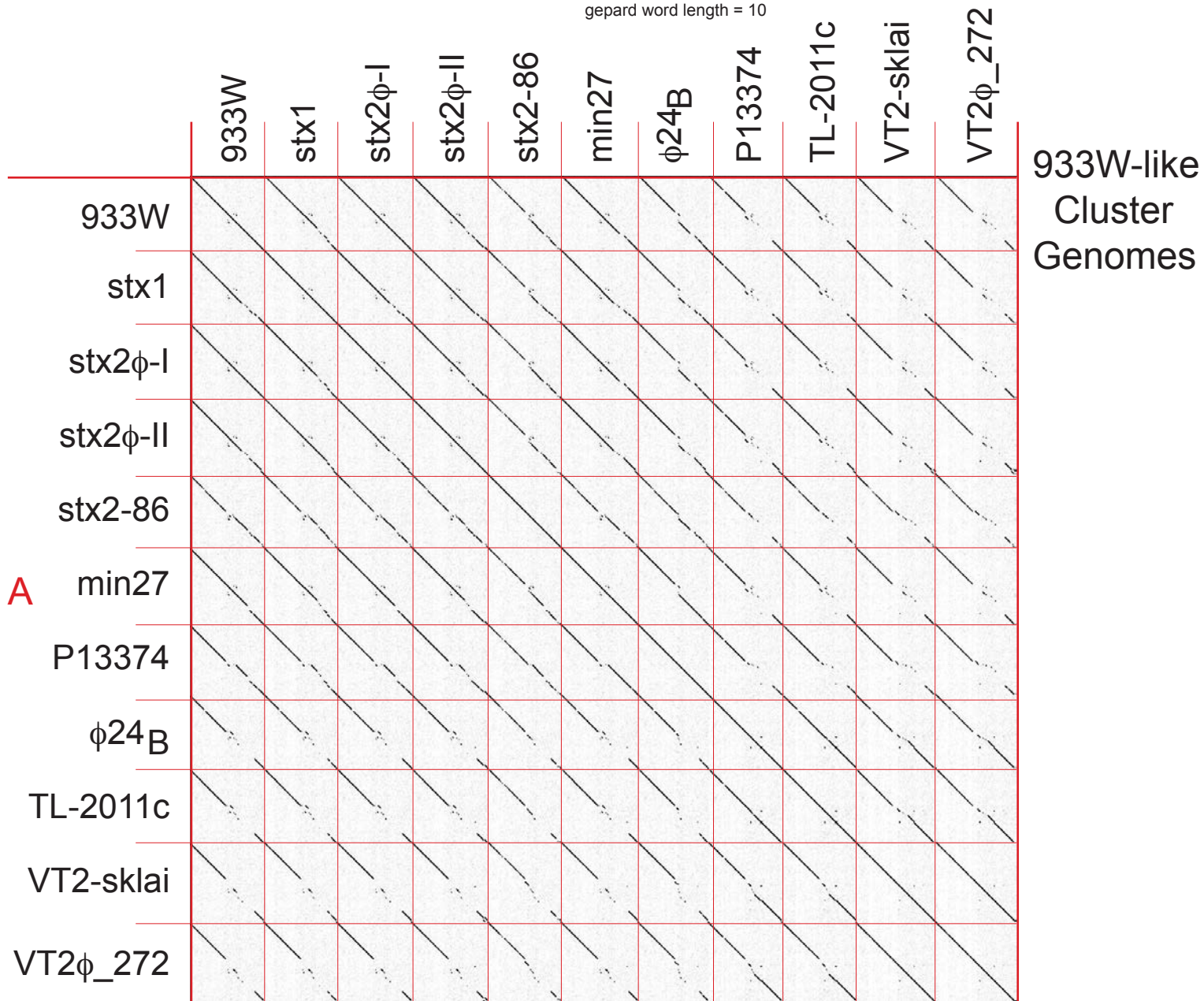


P22- and APSE-1-like Cluster Genomes

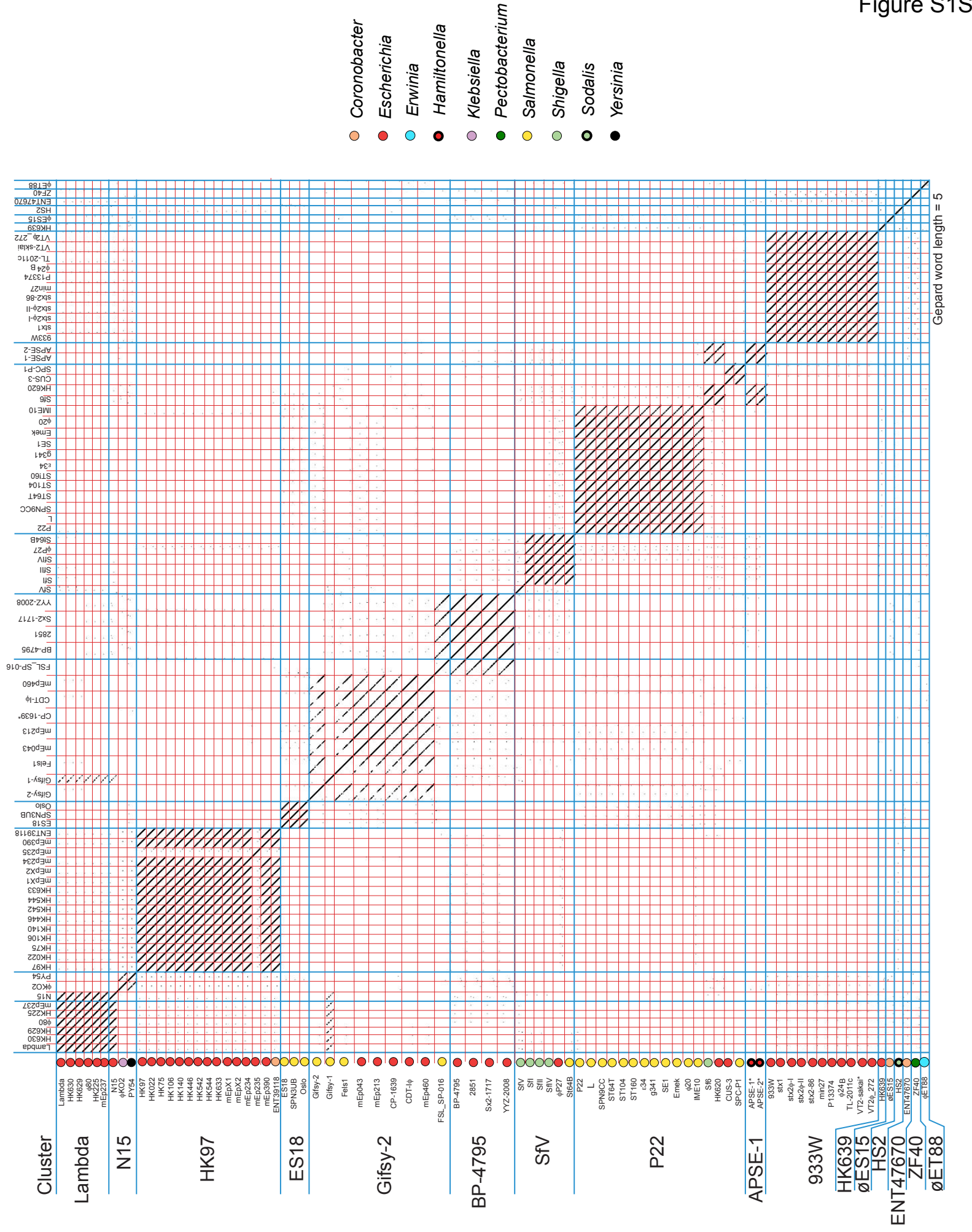




gepard word length = 10

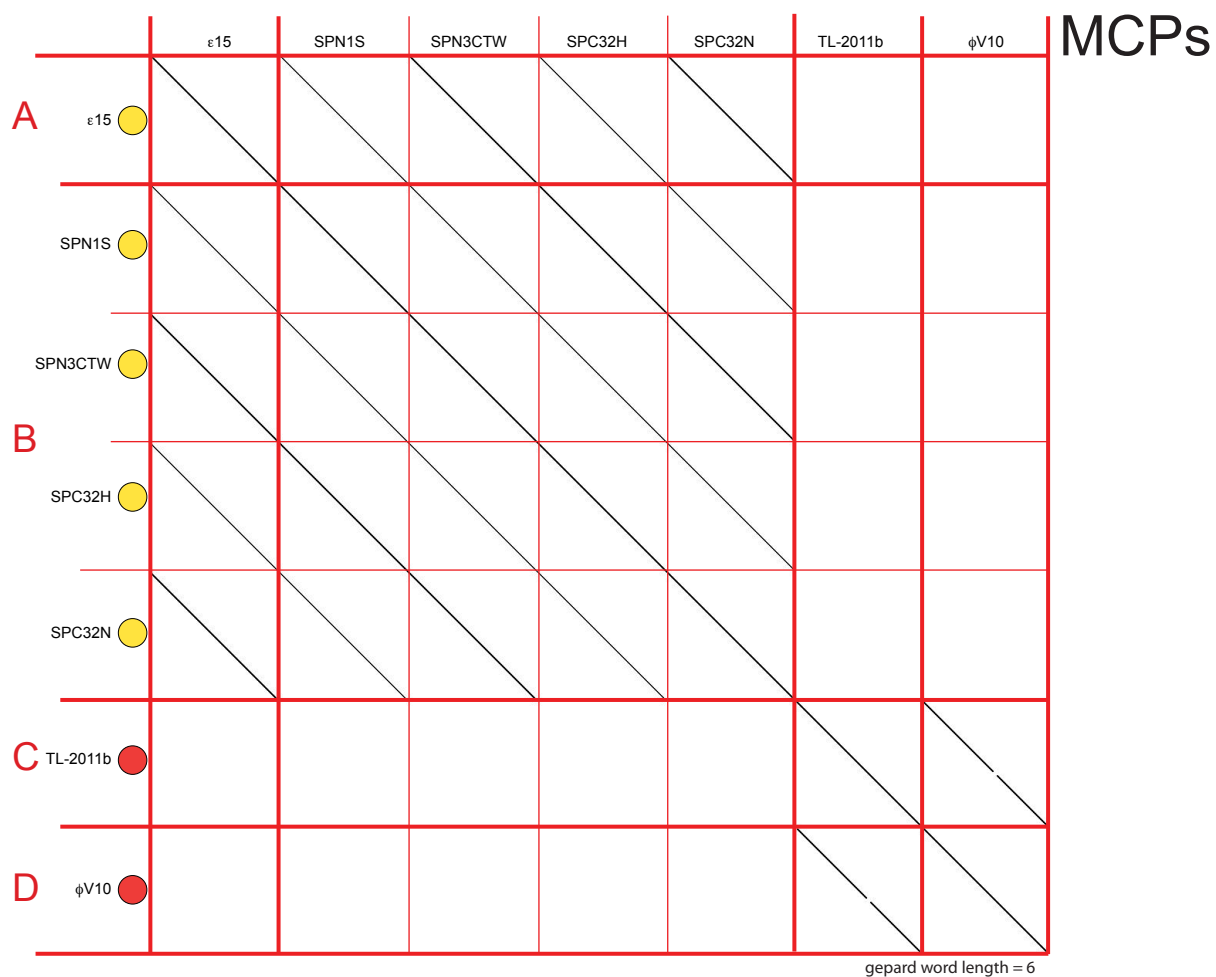
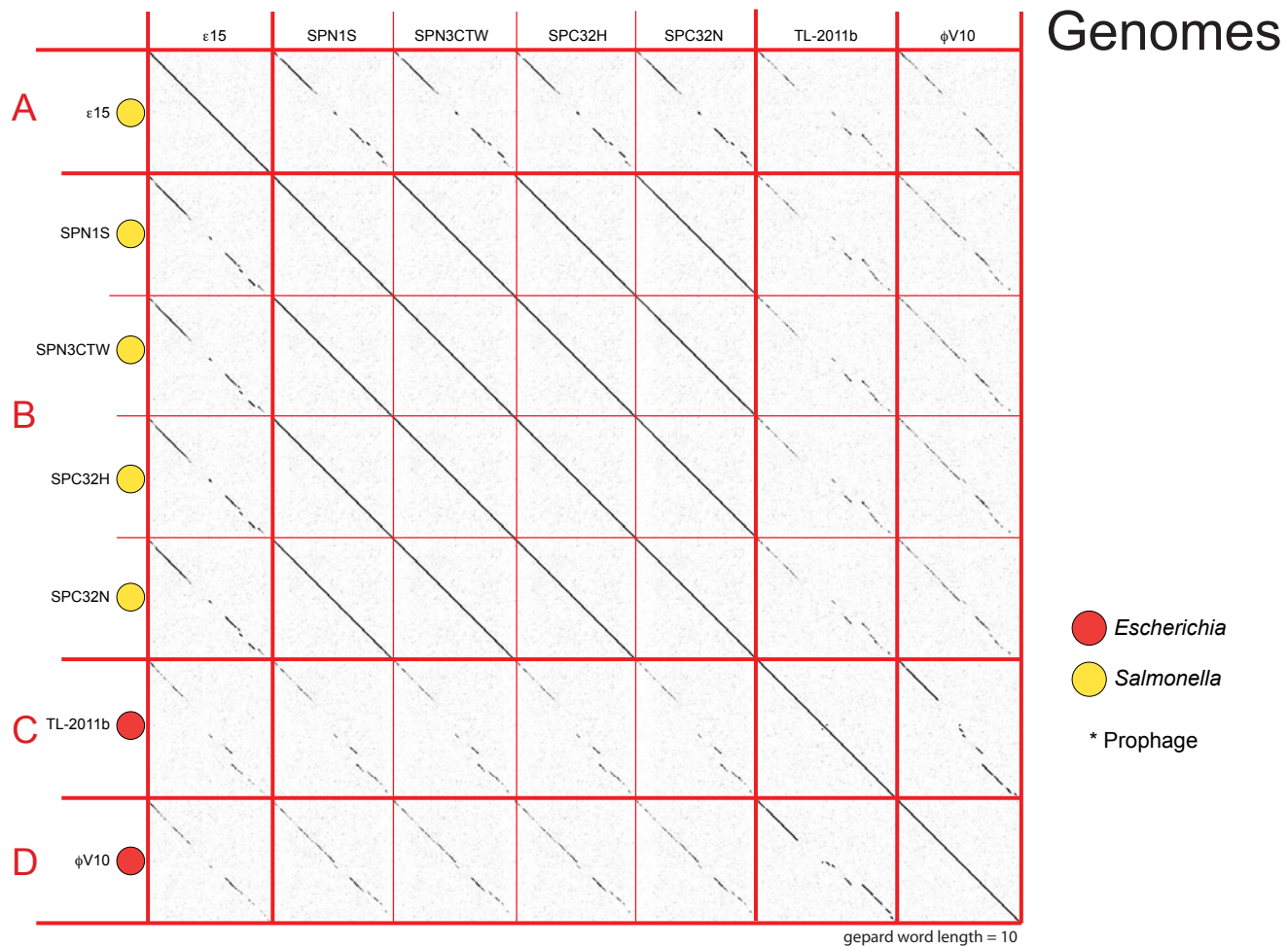


gepard word length = 10

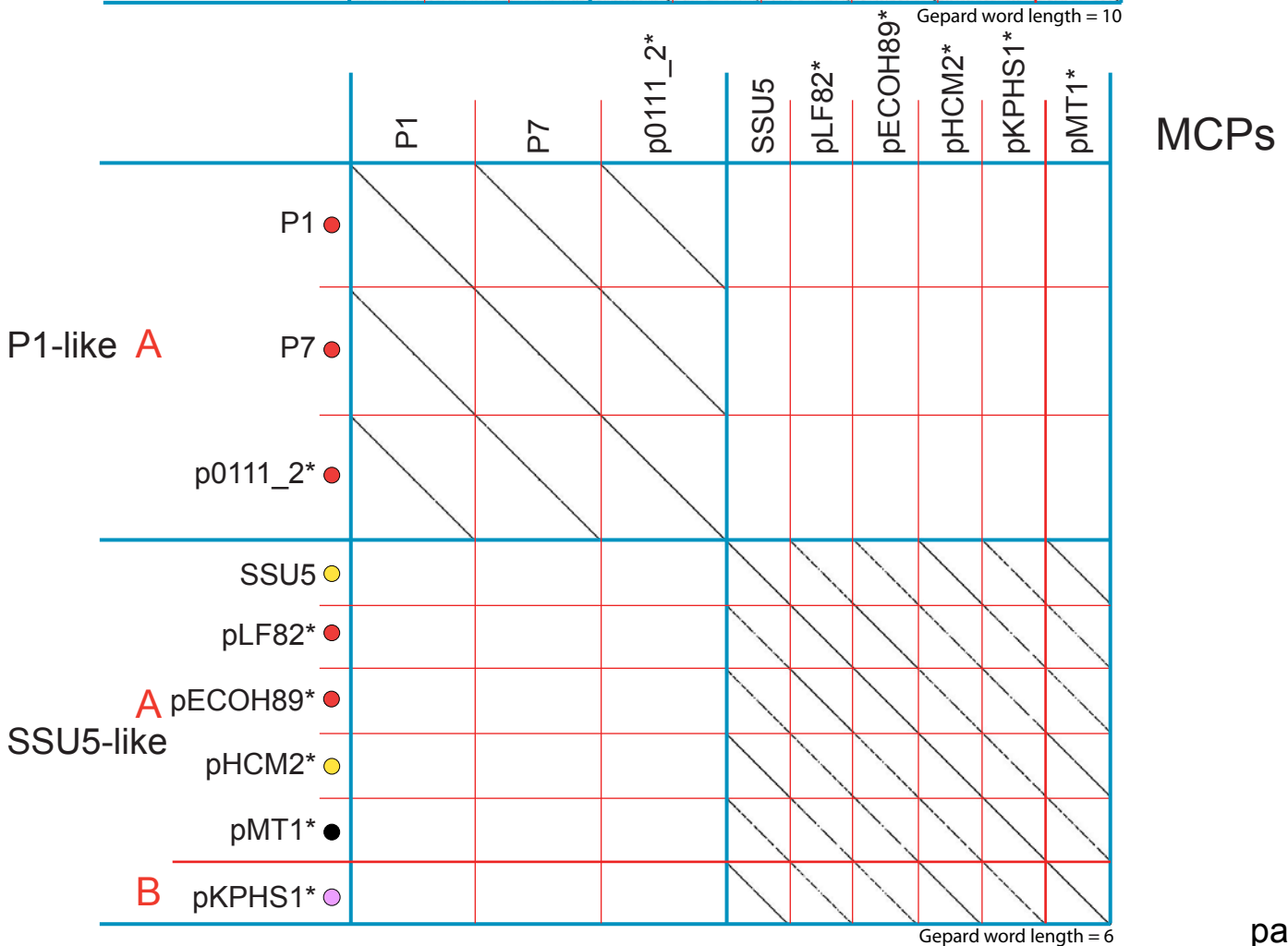
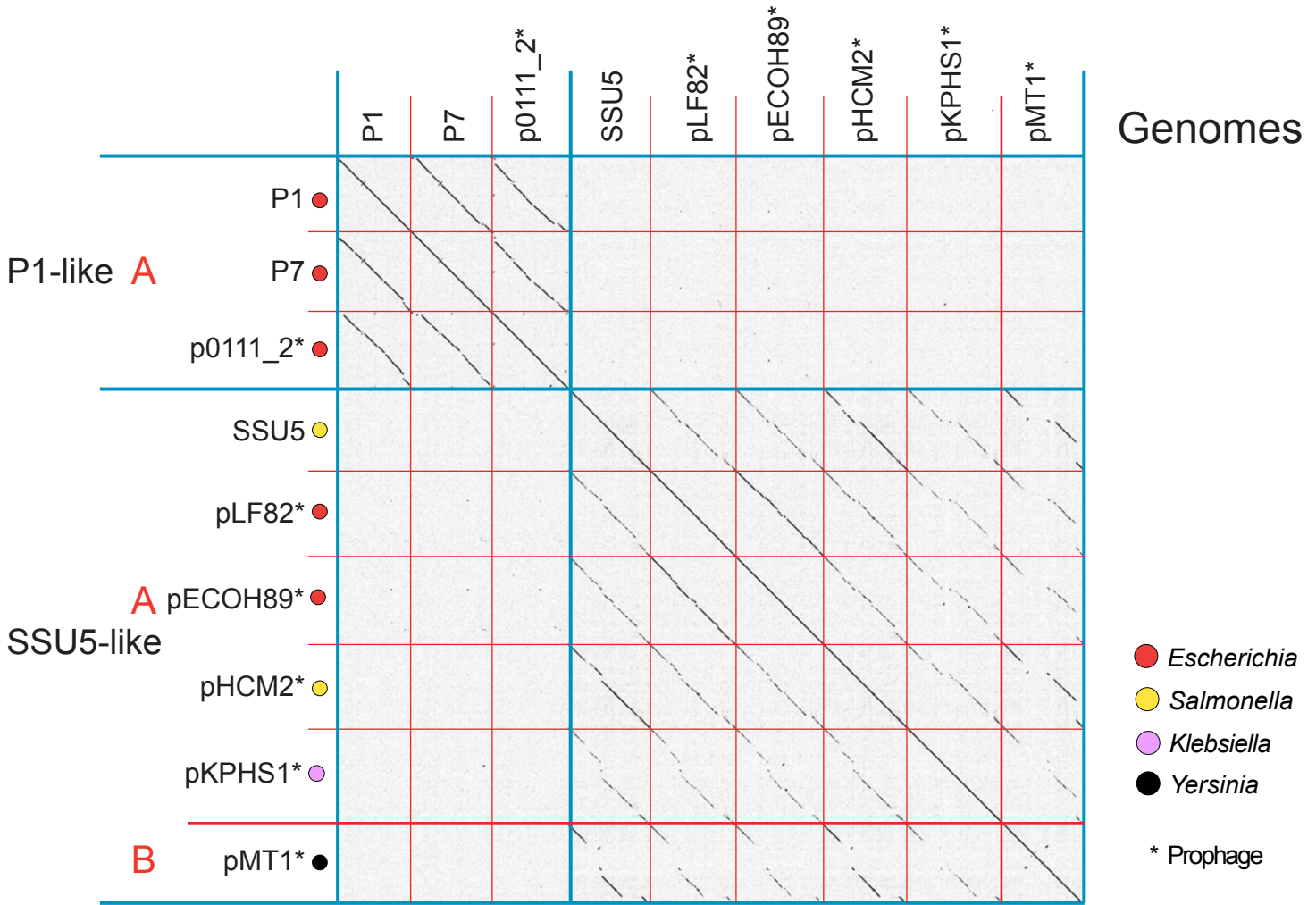


- *Coronobacter*
- *Escherichia*
- *Erwinia*
- *Hamiltonella*
- *Klebsiella*
- *Pectobacterium*
- *Salmonella*
- *Shigella*
- *Sodalis*
- *Yersinia*

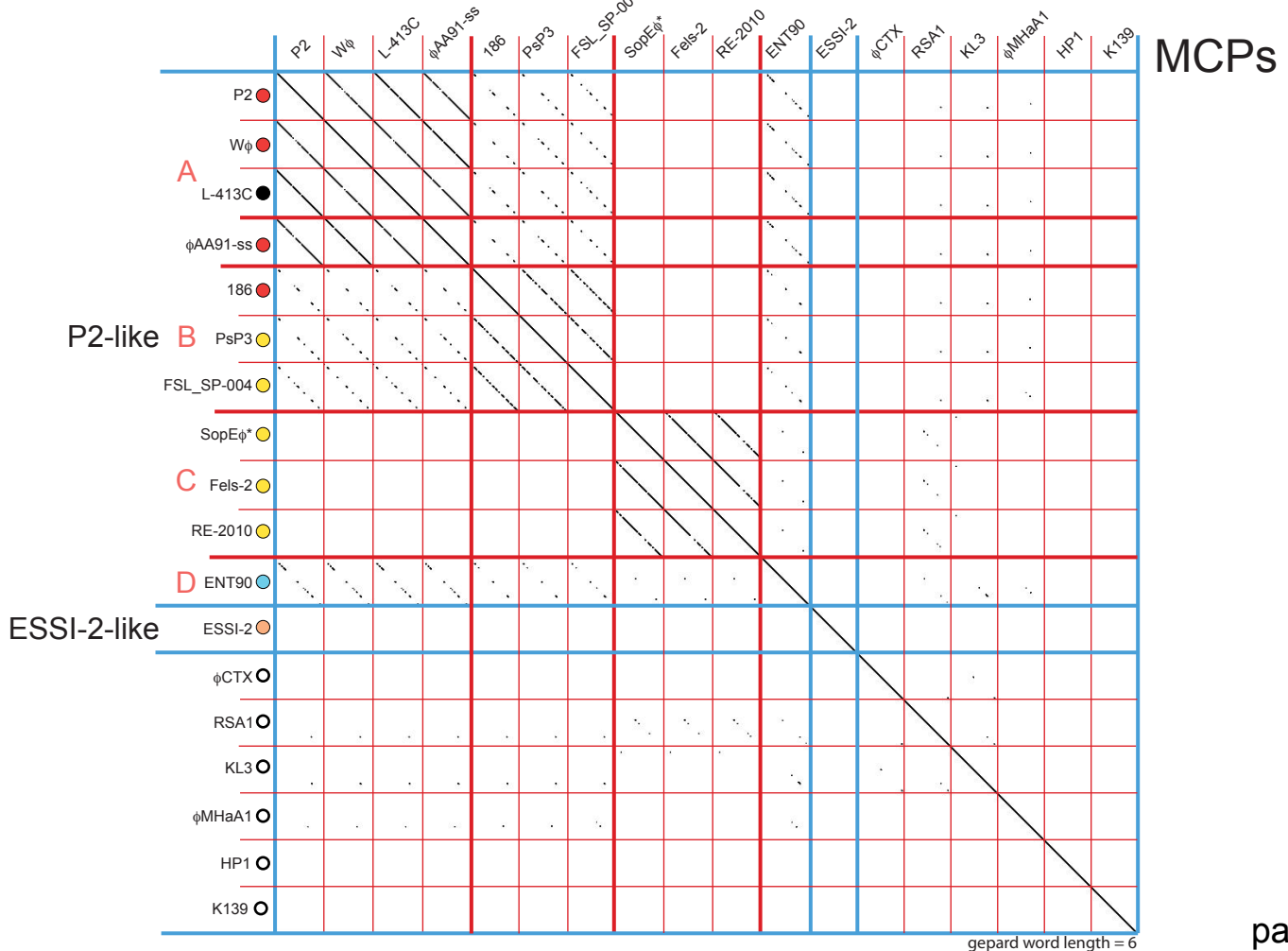
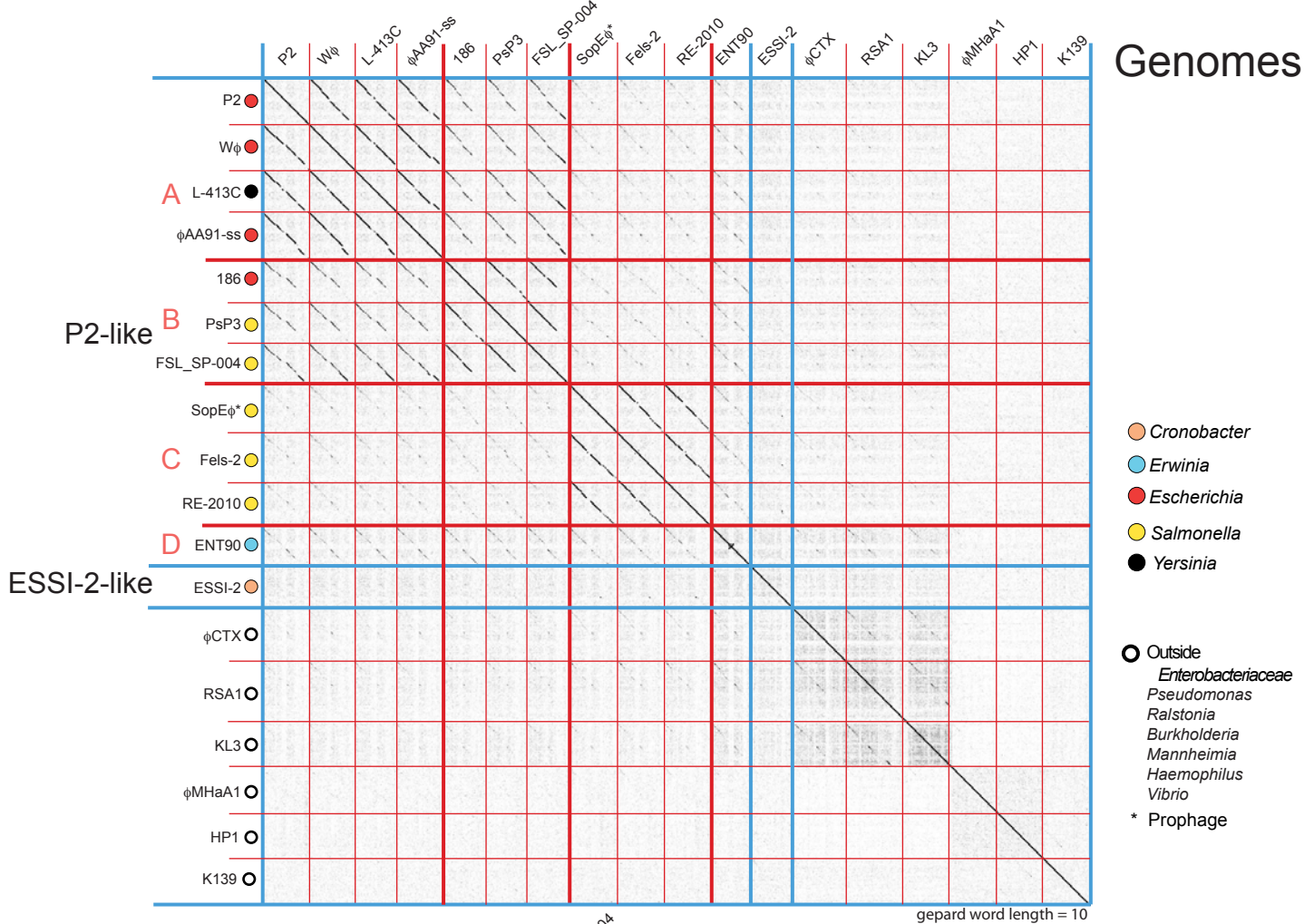
Gepard word length = 5

ϵ 15-like Cluster

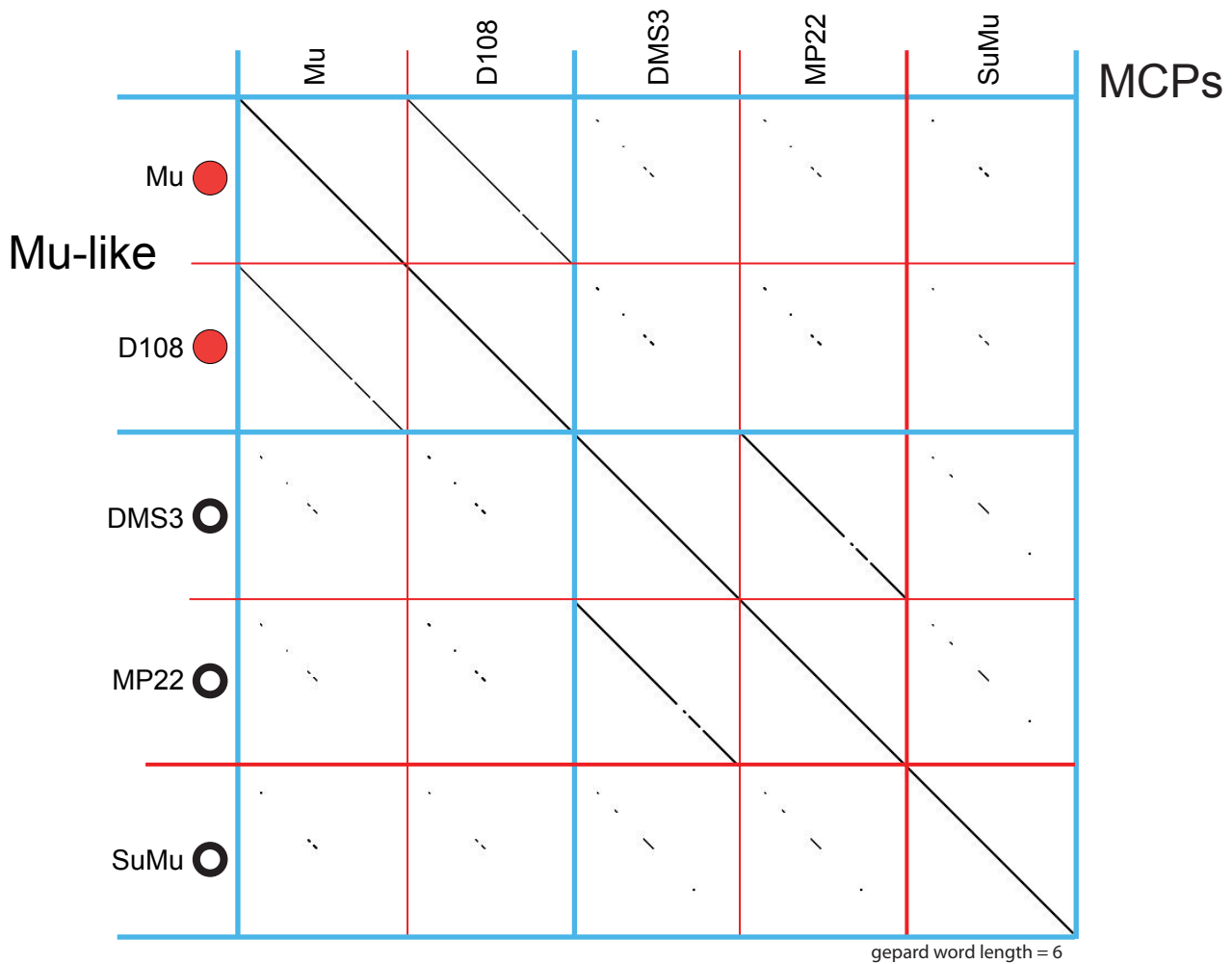
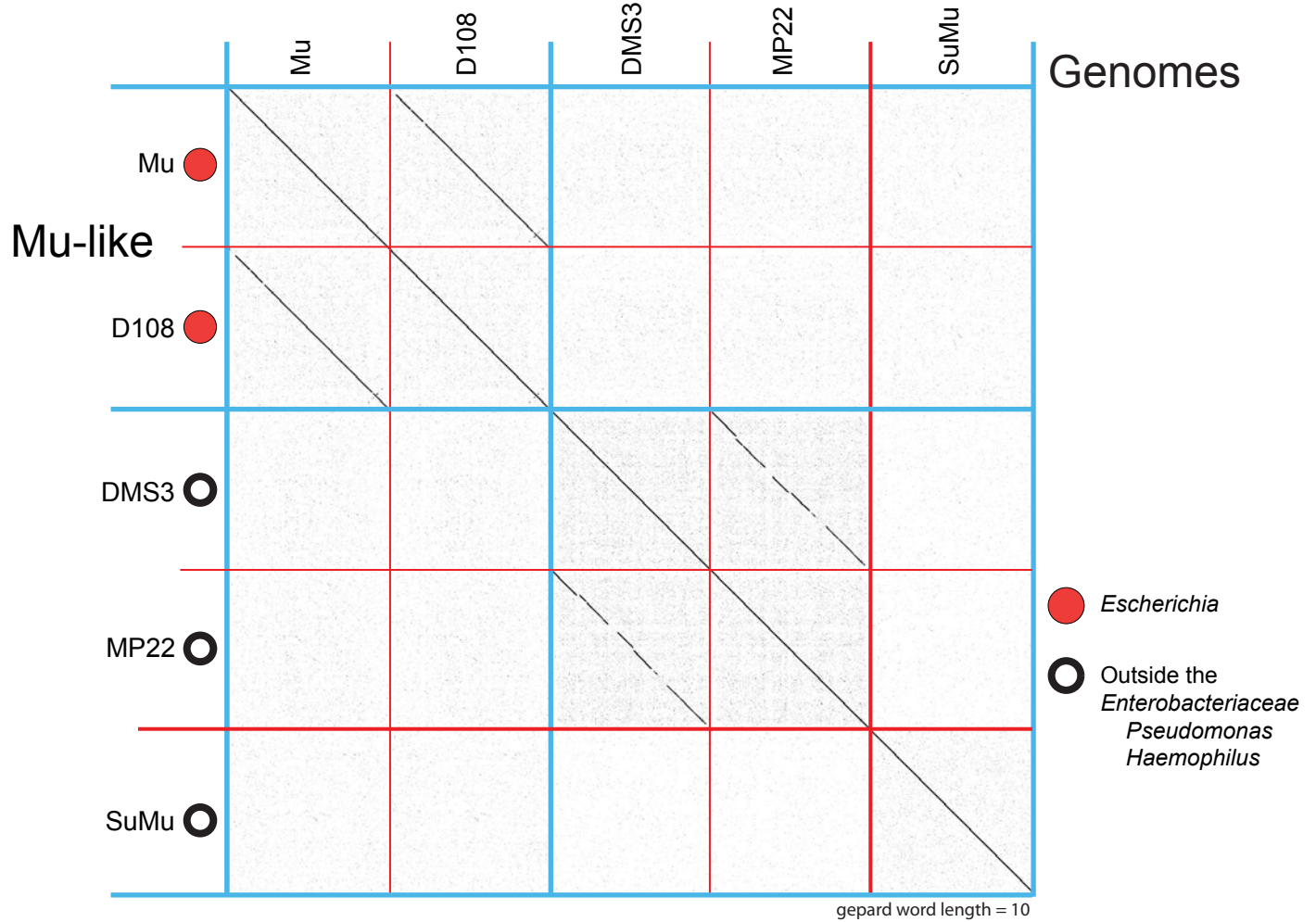
P1 and SSU5-like Clusters



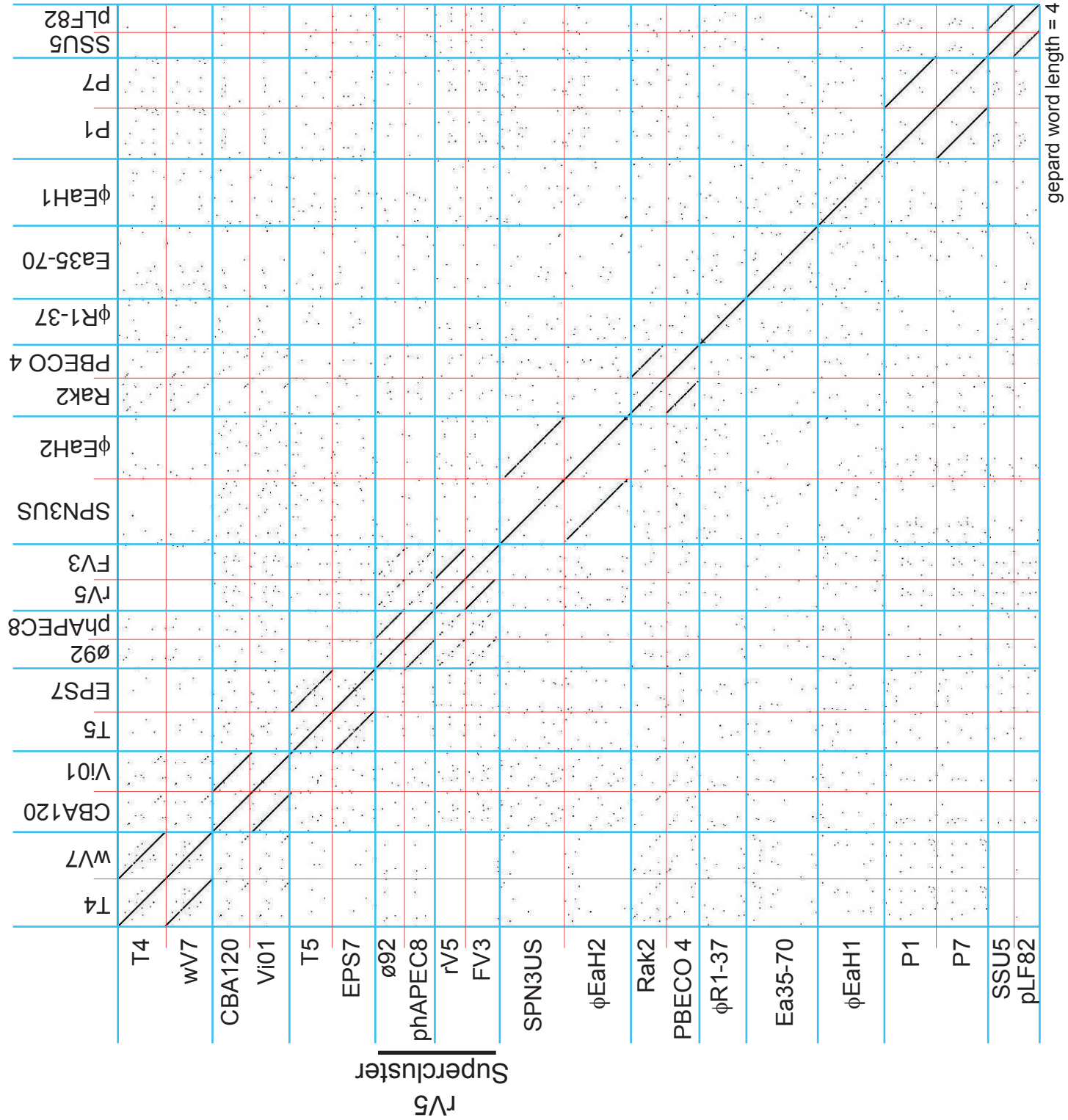
P2 Supercluster



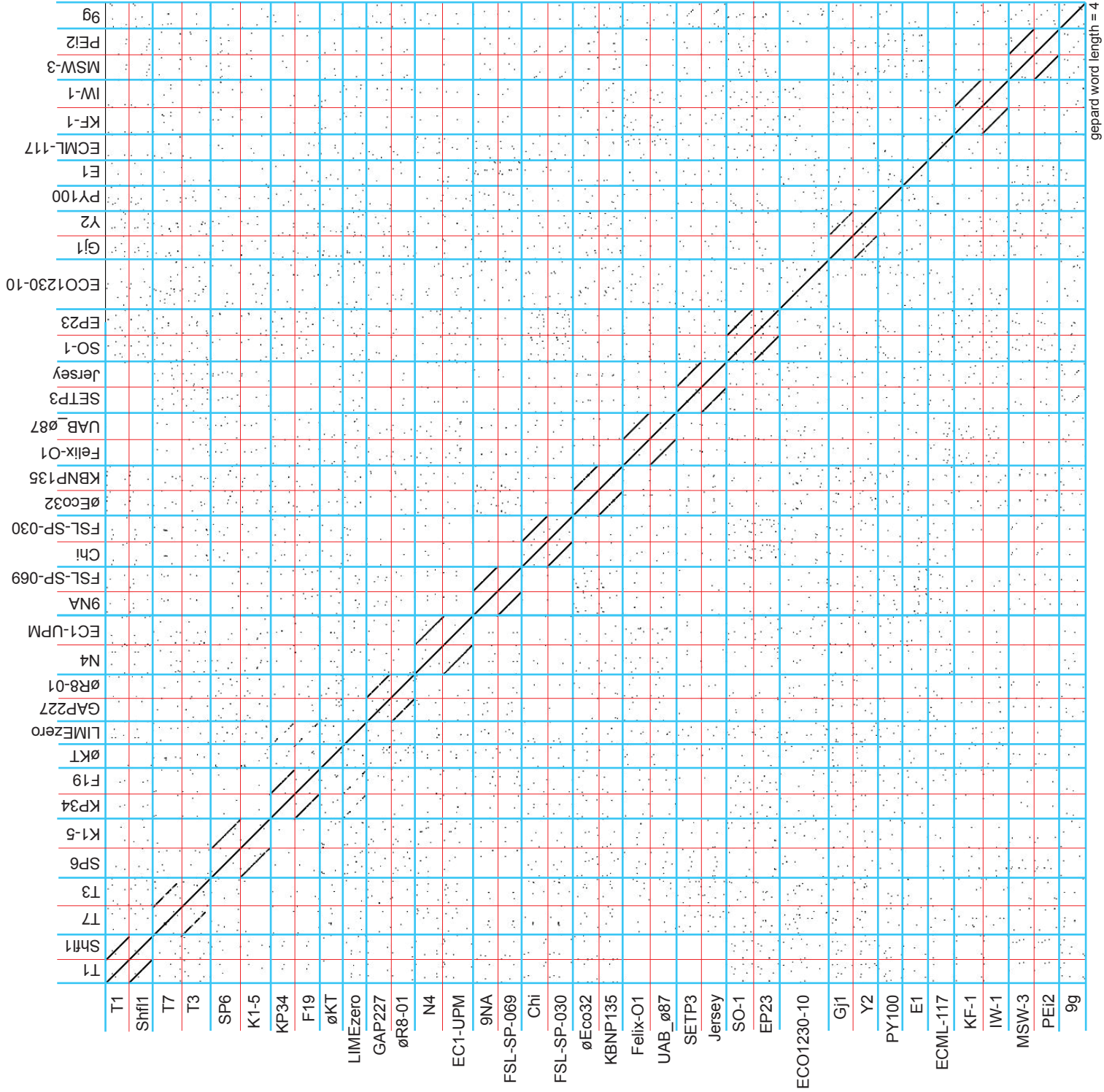
Mu-like Cluster



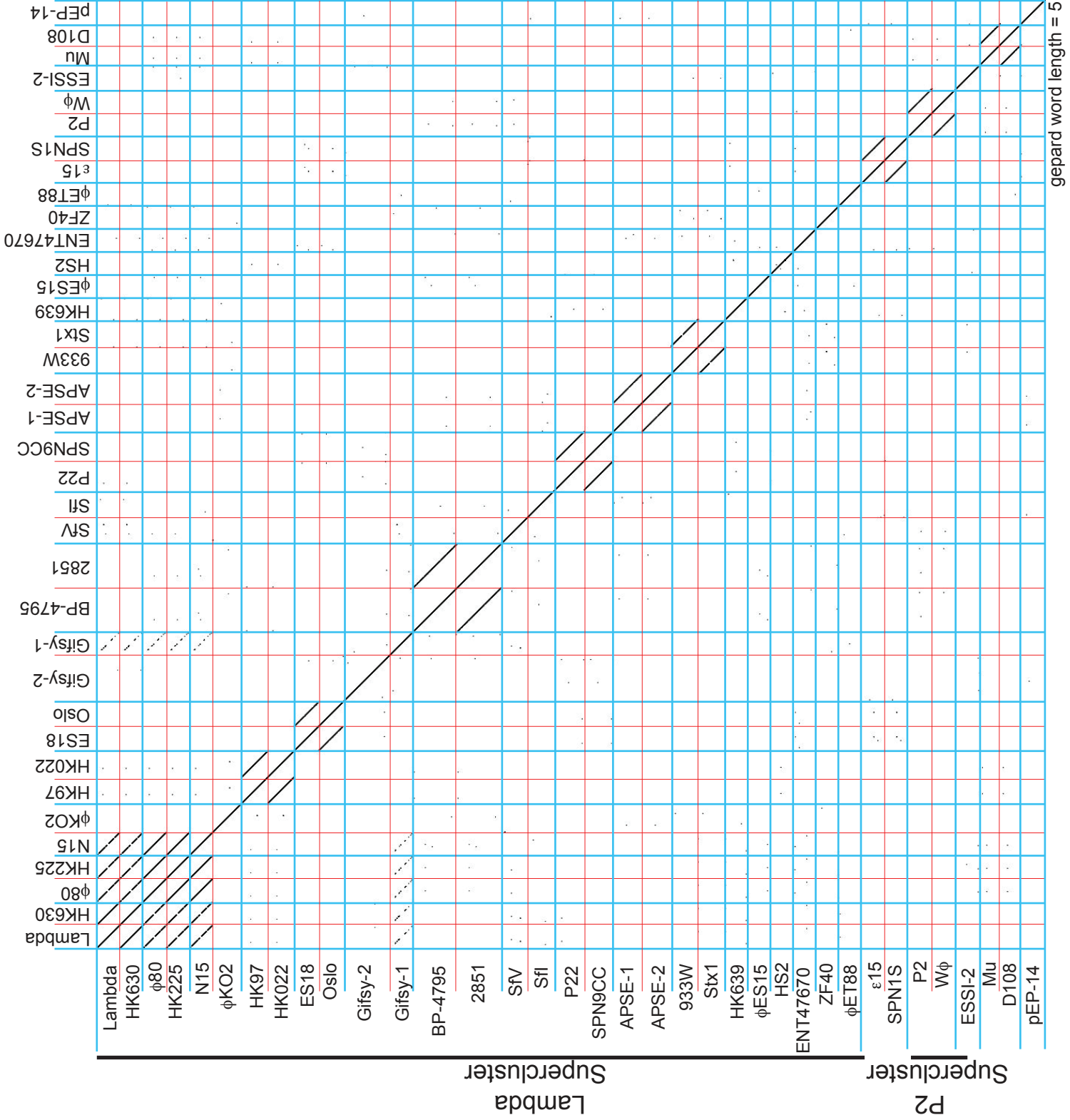
MCPs of Representative Phages With Genomes Larger Than 90 kbp



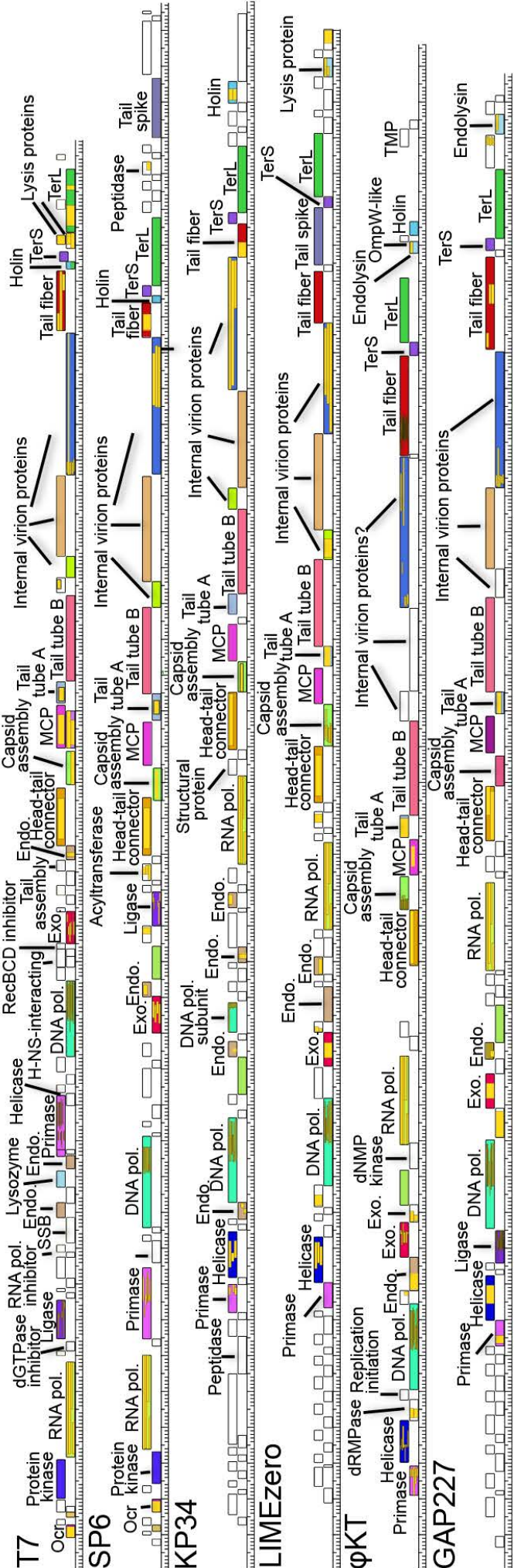
MCPs of Representative Lytic Phages With Genomes Less Than 90 kbp



MCPs of Representative Temperate Phages With Genomes Less Than 90 kbp



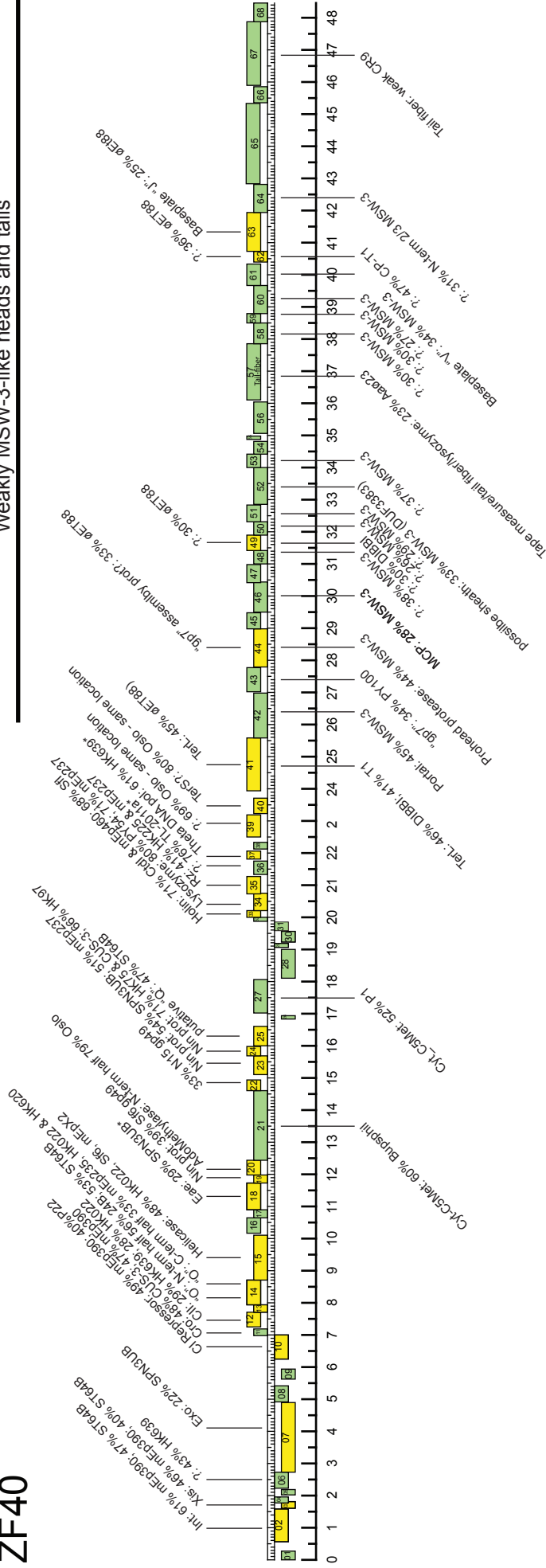
Genomic maps of a representative phage from each of the six T7 supercluster clusters.



Genes of phages ZF40 and øE188 that are homologous and syntenic with phage lambda cluster genomes

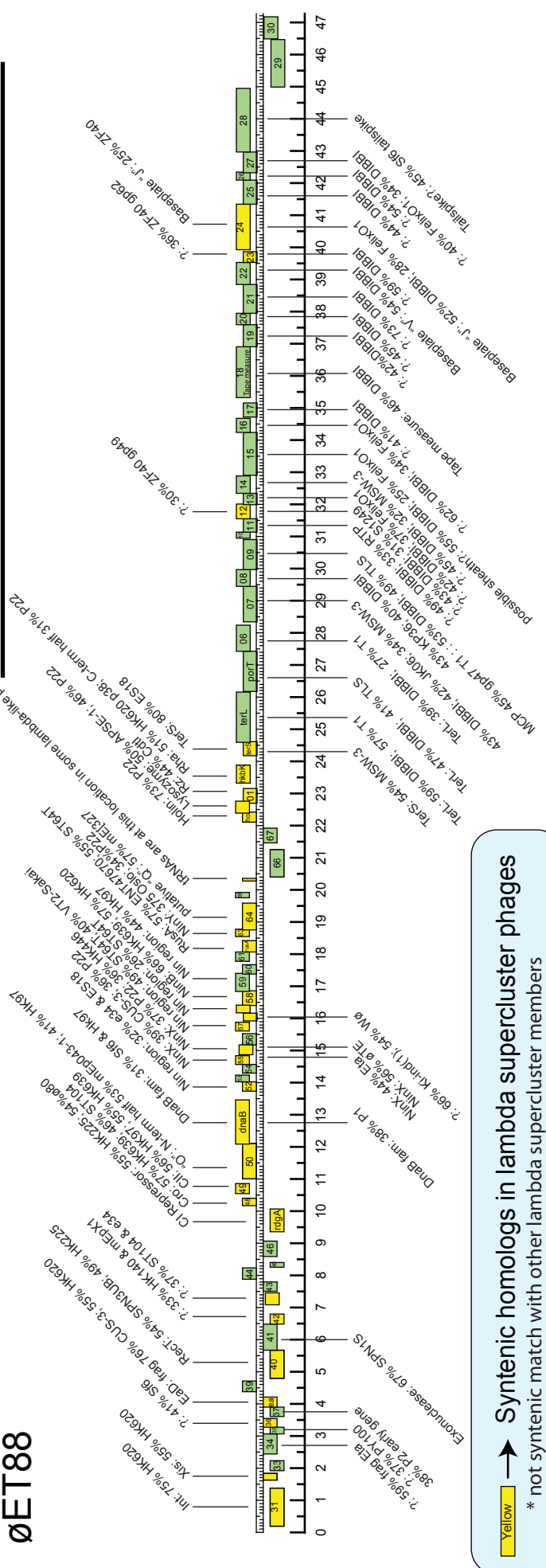
ZF40

Weakly MSW-3-like heads and tails

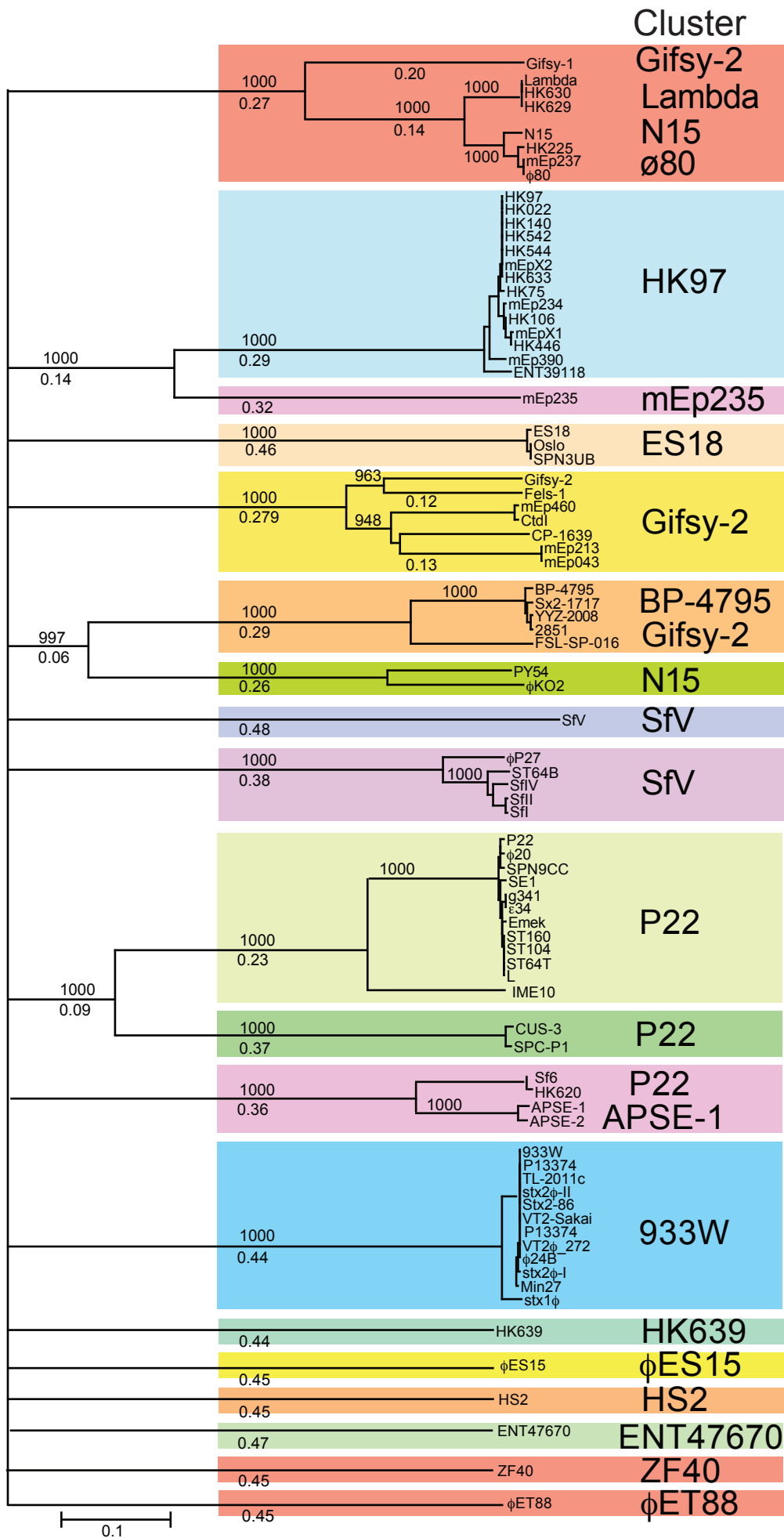


øE188

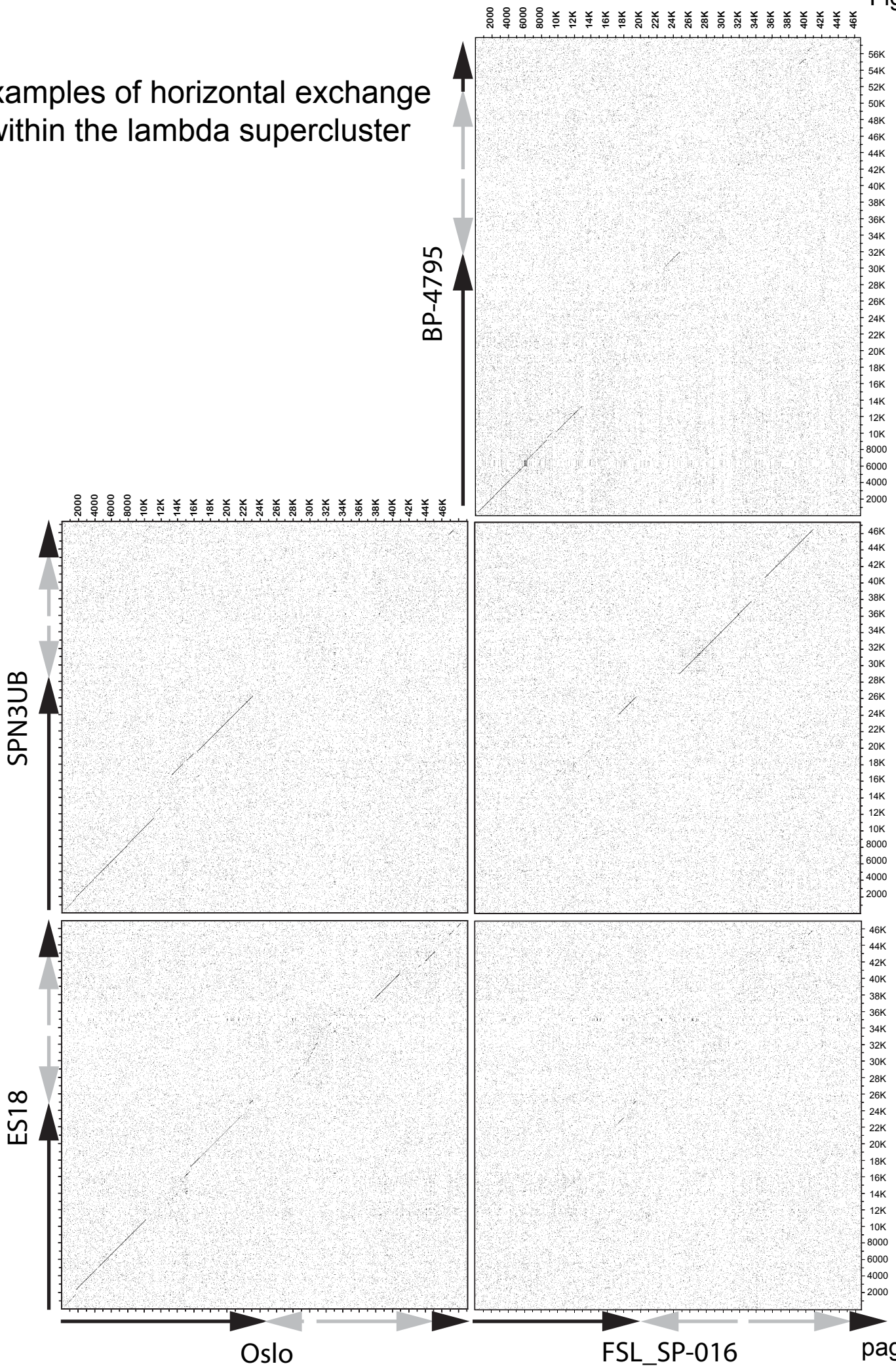
Xanthomonas phage DIBBI-like heads and tails



Syntenic homologs in lambda supercluster phages
 * not syntenic match with other lambda supercluster members



Examples of horizontal exchange within the lambda supercluster



Examples of mosaicism within the HK97-like cluster

