

Supplementary Figure 1. Phylogenetic tree based on 16S rRNA gene sequences of the genus *Azospirillum*. The tree was constructed using the neighbor-joining method. The bar indicates 0.01 nucleotide substitutions per 100 bp. Strain B510 is shown in bold letters. GenBank accession numbers for the nucleotide sequences are provided in parentheses.

Supplementary Figure 2. Gene maps of the seven replicons of *Azospirillum* sp. B510. Gene arrangements are shown on the seven gene maps, at the same scale. Green bars indicate 3-kb-long segments, with numerals in kb on the ends of the bars. On both sides of the scale bars, boxes with gene identification codes indicate the potential genes assigned based on computer predictions. Each arrowhead indicates the direction of translation. Predicted protein genes for which functions could be evaluated by similarity search are classified into 25 functional categories, as indicated by different color codes bottom of the figures. RNA-encoding genes are represented by black bars and boxes with red letters. Nine of the rRNA gene clusters are numbered. The sequences and gene information presented in this paper are available in RhizoBase (<http://genome.kazusa.or.jp/rhizobase>).

Supplementary Figure 3. Cumulative plot of KOPS motifs in the B510 genome. The cumulative distribution of the GGGNAGGG sequence was calculated for the seven replicons, and is plotted on the Y axes. We searched for the KOPS motif (F, forward strand; R, reverse strand) in 1-kb windows.

Supplementary Figure 4. The CRISPR locus on the B510 genome. A self-comparison of the B510 genome by GenomeMatcher (A and B) and the structure of CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) systems (C). (A) Overall self-comparison of the B510 genome. (B) Magnification of genomic regions with a high density of repeated sequences. (C) CRISPRs in the B510 genome were composed of 70 different spacers (of 36-42 bp) and an identical repeat element (of 37 bp) containing an 8-bp palindrome (5'-CCTGGGCG) at the coordinates 3,158,606-3,163,849 on the chromosome, and are preceded by two *cas* genes and a lead sequence (LDR).

Supplementary Figure 5. Sequence alignment of the repeated units in the CRISPR system on the B510 chromosome. A sequence unit for CRISPR consists of a repeat element and a spacer. Yellow shading indicates repeated sequences, and the other colors indicate spacers with distinct sequences. The number above each nucleotide sequence indicates the starting position of the repeat unit on the B510 chromosome, in bp.

Supplementary Figure 6. Predicted tRNA structures of *Azospirillum* sp. B510. The nucleotide sequences of 79 tRNA genes and their structural domains are shown. The loci on the replicons are indicated in parentheses.

Supplementary Figure 7. Comparison of the configuration of rRNA gene clusters (*rrns*). The identification of the *rrn* cluster is depicted in “Supplementary Figure 2”. Boxes on horizontal lines represent arrangements of the RNA genes in the cluster. The broken lines show truncated parts of the *rrn* cluster. The *rrn* clusters are classified into three groups based on variations in their arrangements.

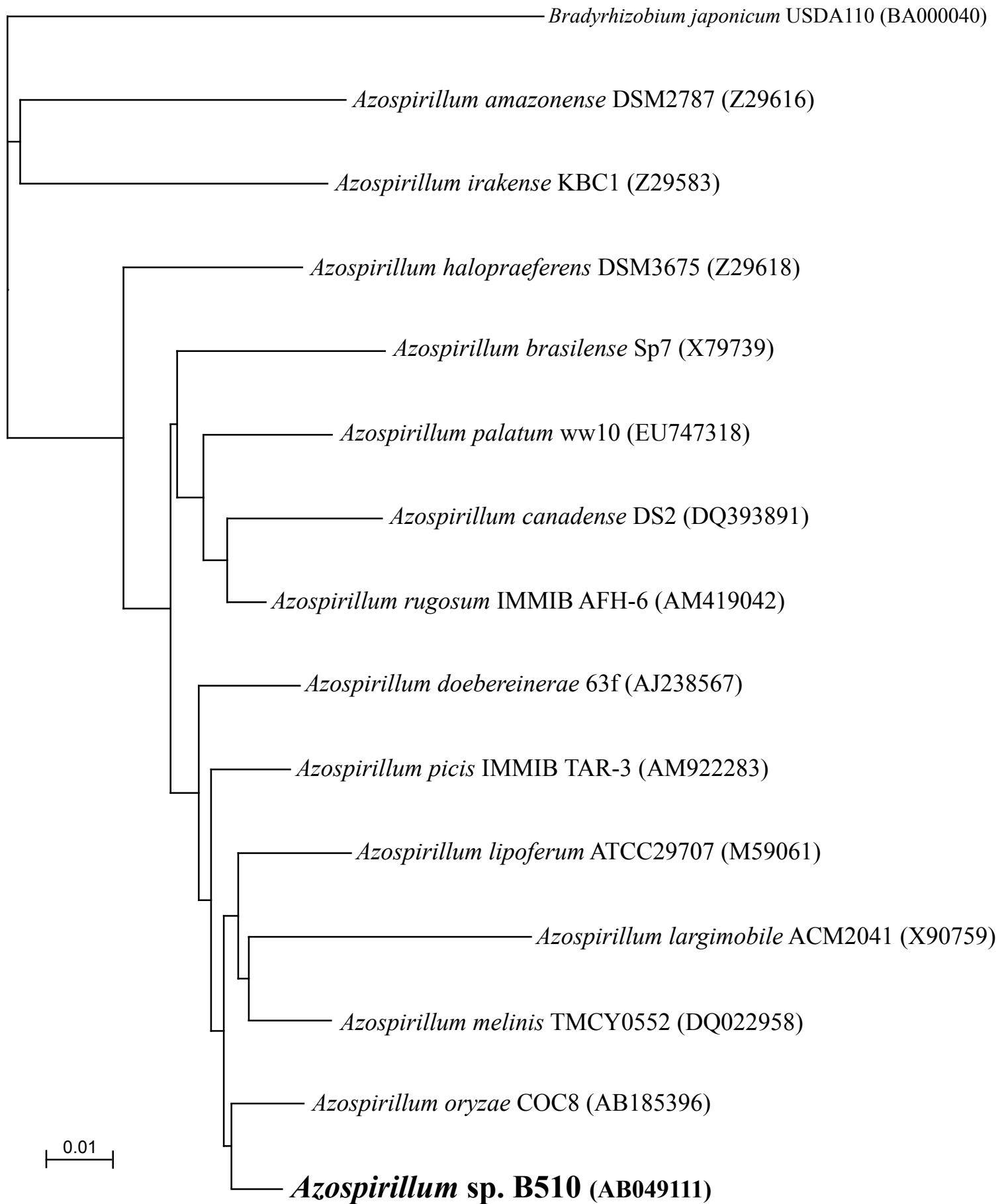
Supplementary Figure 8. The nucleotide sequences of *acdS* and *acdR* in *Azospirillum* sp. B510. The *acdS* and *acdR* genes are indicated with red and blue arrows, respectively, and the direction of the arrows denotes the direction of transcription. The box indicates a putative LRP protein-binding site (A/T rich region) with a palindrome (black arrow), which is similar to that of *Enterobacter cloacae* (Glichko and Glick, 2000).

Glichko, V. P., and Glick, B. R. 2000, Identification of DNA sequences that regulate the expression of the *Enterobacter cloacae* UW4 1-aminocyclopropane-1-carboxylic acid deaminase gene. *Can. J. Microbiol.*, **46**, 1159-1165.

Supplementary Figure 9. Phylogenetic tree of malic enzymes produced by the plant-associated bacteria. This tree was built via the neighbor-joining method. Gene products for the malic enzymes are shown with IDs from the genomic data, as follows: AZL000200, AZLa05660, AZLa08520: *Azospirillum* sp. B510; AZC_0119, AZC_3656: *Azorhizobium caulinodans* ORS571; BBta_3876, BBta_5810: *Bradyrhizobium* sp.

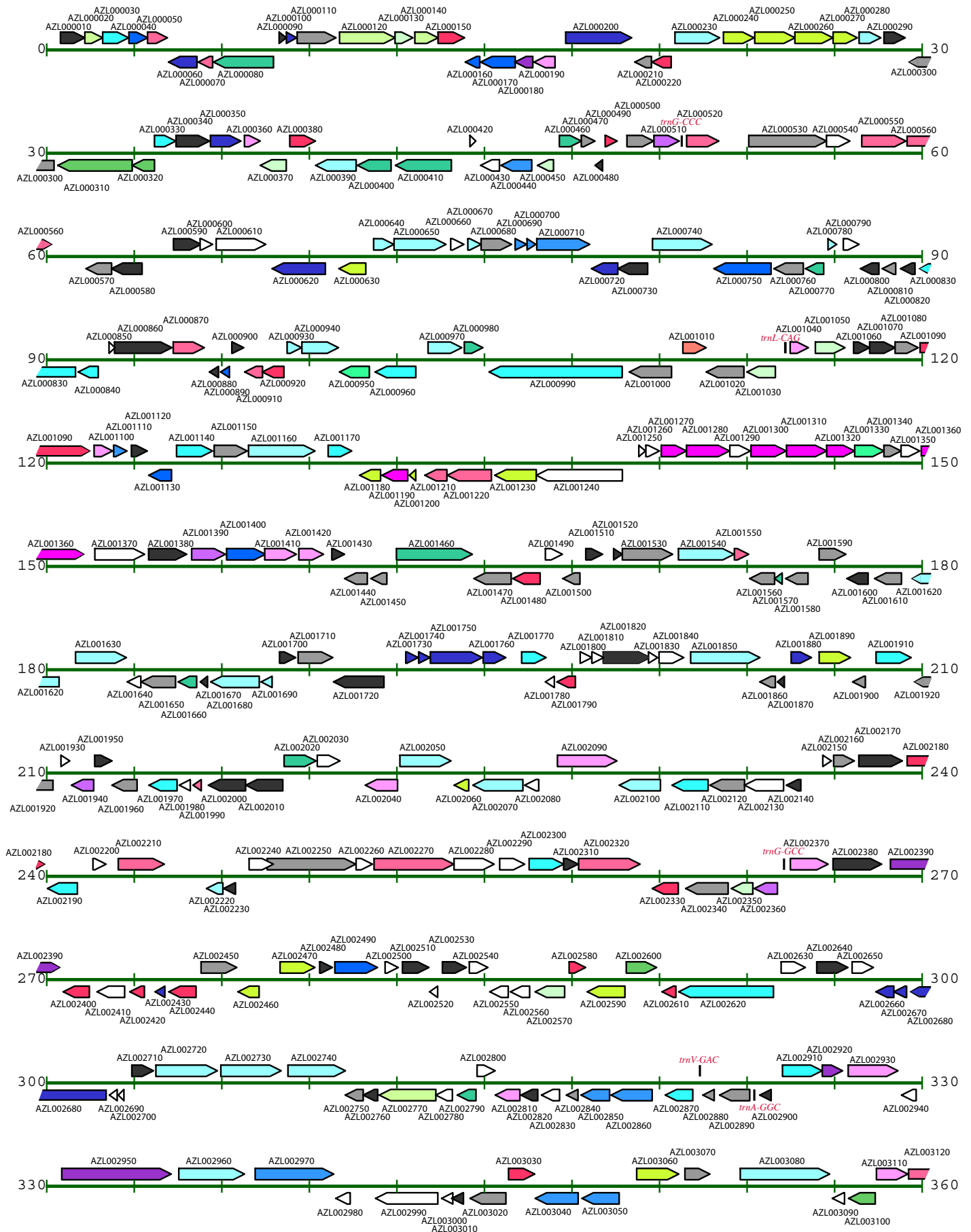
BTai1; BRADO3367, BRADO5322: *Bradyrhizobium* sp. ORS278; GKPORF_B0937, GKPORF_B1530, GKPORF_B2116: *Klebsiella pneumoniae* Kp342; NGR_c00300, NGR_c16880: *Rhizobium* sp. NGR234; RHE_CH00389, RHE_CH02355: *Rhizobium etli* CFN42; RL0407, RL2671: *Rhizobium leguminosarum* bv. *viciae* 3841; SMc00169, SMc01126: *Sinorhizobium meliloti* 1021; azo0821, azo3211: *Azoarcus* sp. BH72; bll6469, blr3726, blr4145: *Bradyrhizobium japonicum* USDA110; and mlr0809, mlr5329: *Mesorhizobium loti* MAFF303099.

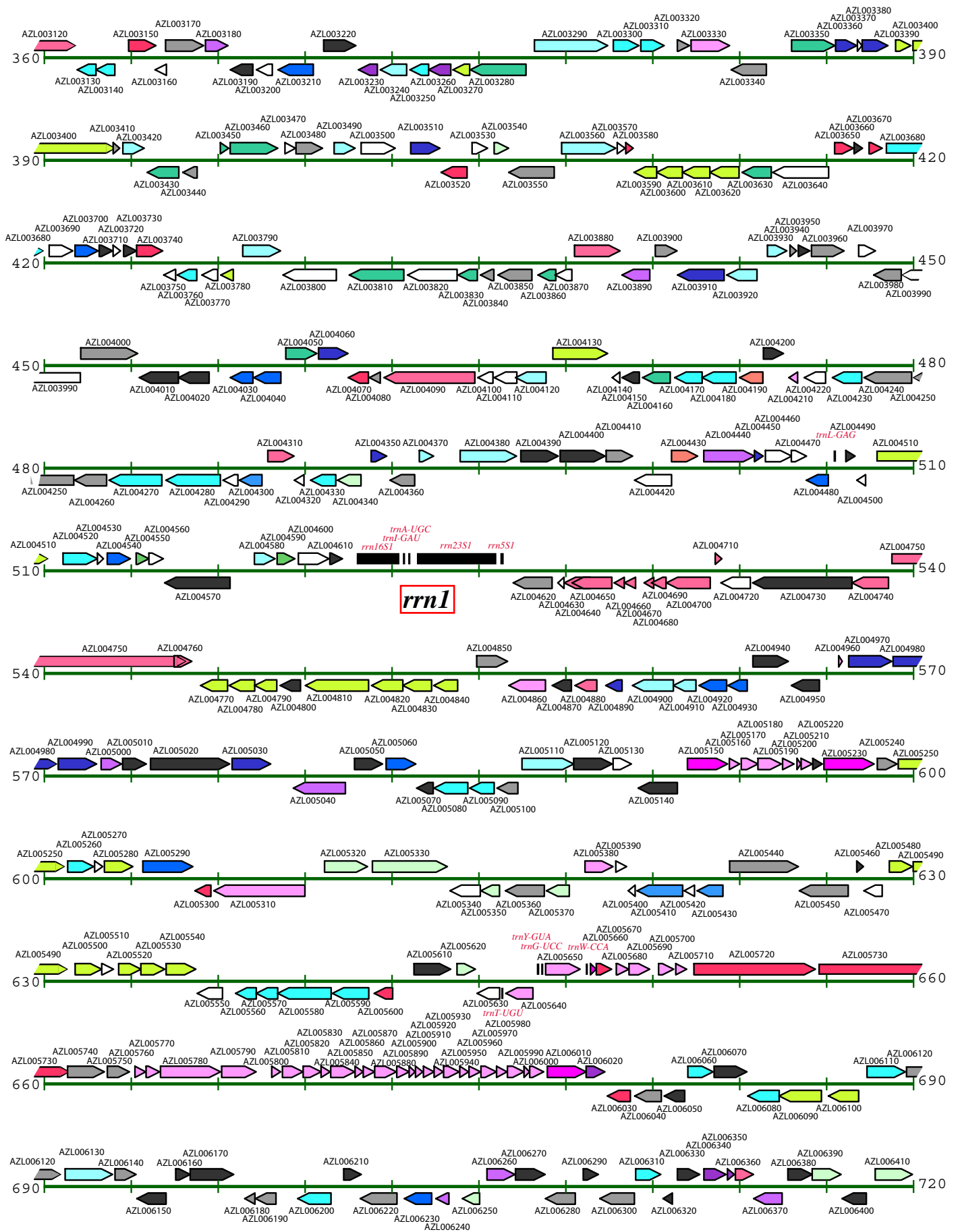
Supplementary Figure 10. Comparison of the pRhico-plasmid and B510-genome sequences. The pRhico plasmid is presented at the top of the figure, with contigs 1 to 5 arranged in numerical order. The B510 genome is shown at the bottom of the figures, with genomic elements arranged in the following order: chromosome, pAB510a, pAB510b, pAB510c, pAB510d, pAB510e, and pAB510f. The B510 chromosome and plasmid sequences are drawn laid end-to-end in the figure. These sequences were analyzed using the blastn function of the GenomeMatcher V.1.270. The BLAST E-value of 10^{-10} was considered significant for the assignments. A red line linking the top to the bottom represents a match between pRhico and the region encoding the replication protein for pAB510f. Blue lines indicate genes involved in cell envelope biosynthesis and outer membrane constitution. Yellow lines link regions encoding tRNA. Green bars indicate a putative transposase encoded in IS. Pale brown bars indicate other matches.

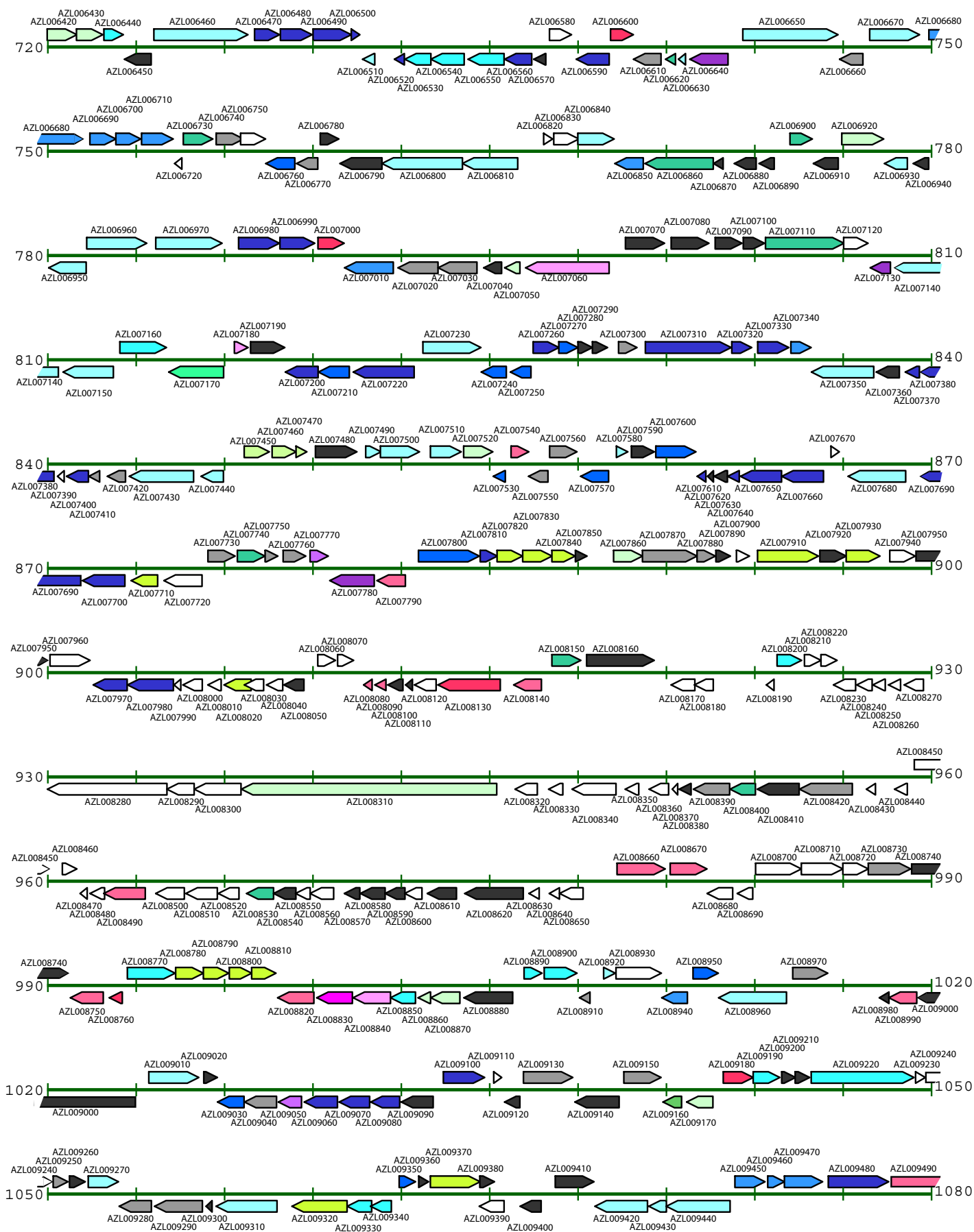


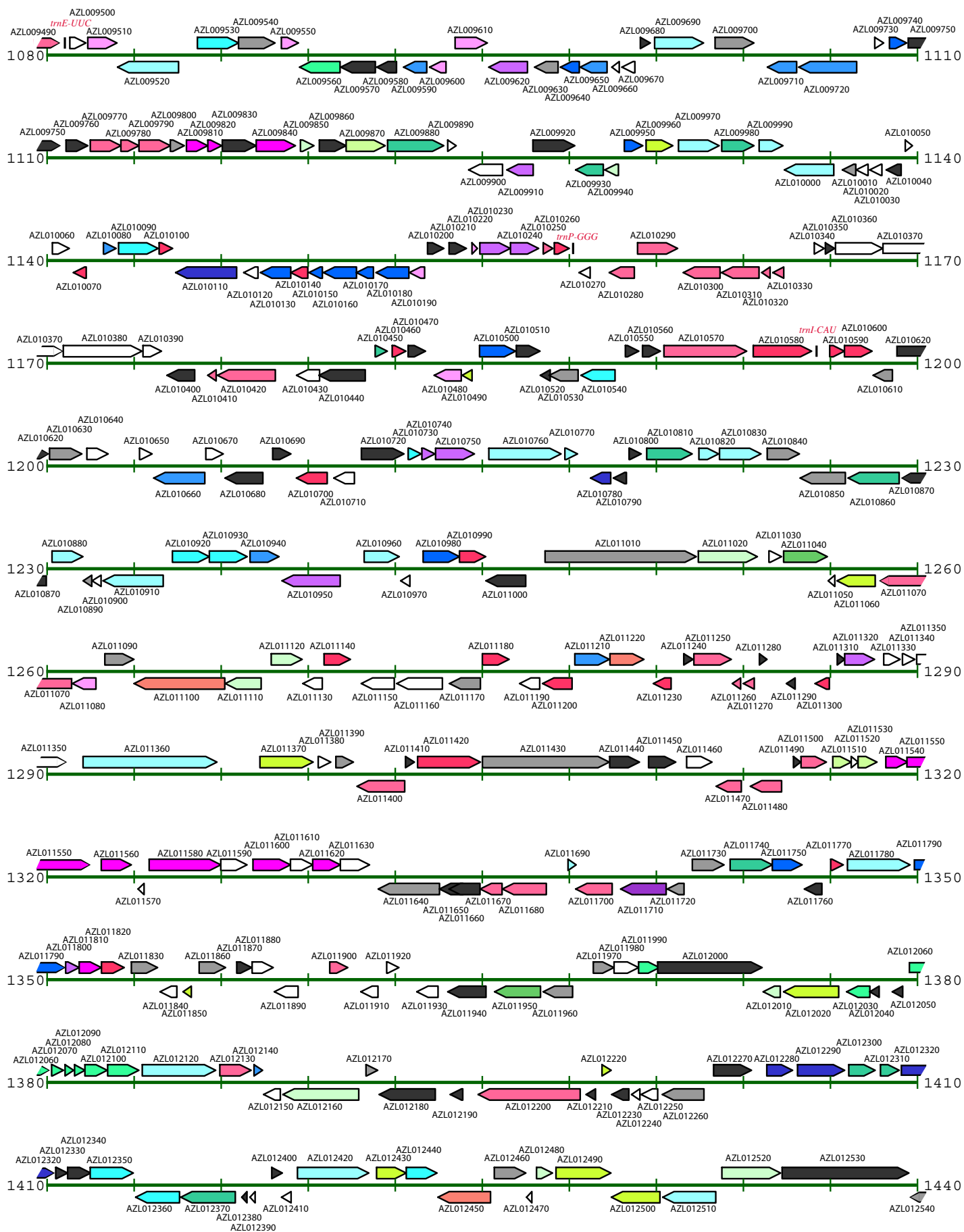
Supplementary Figure 1

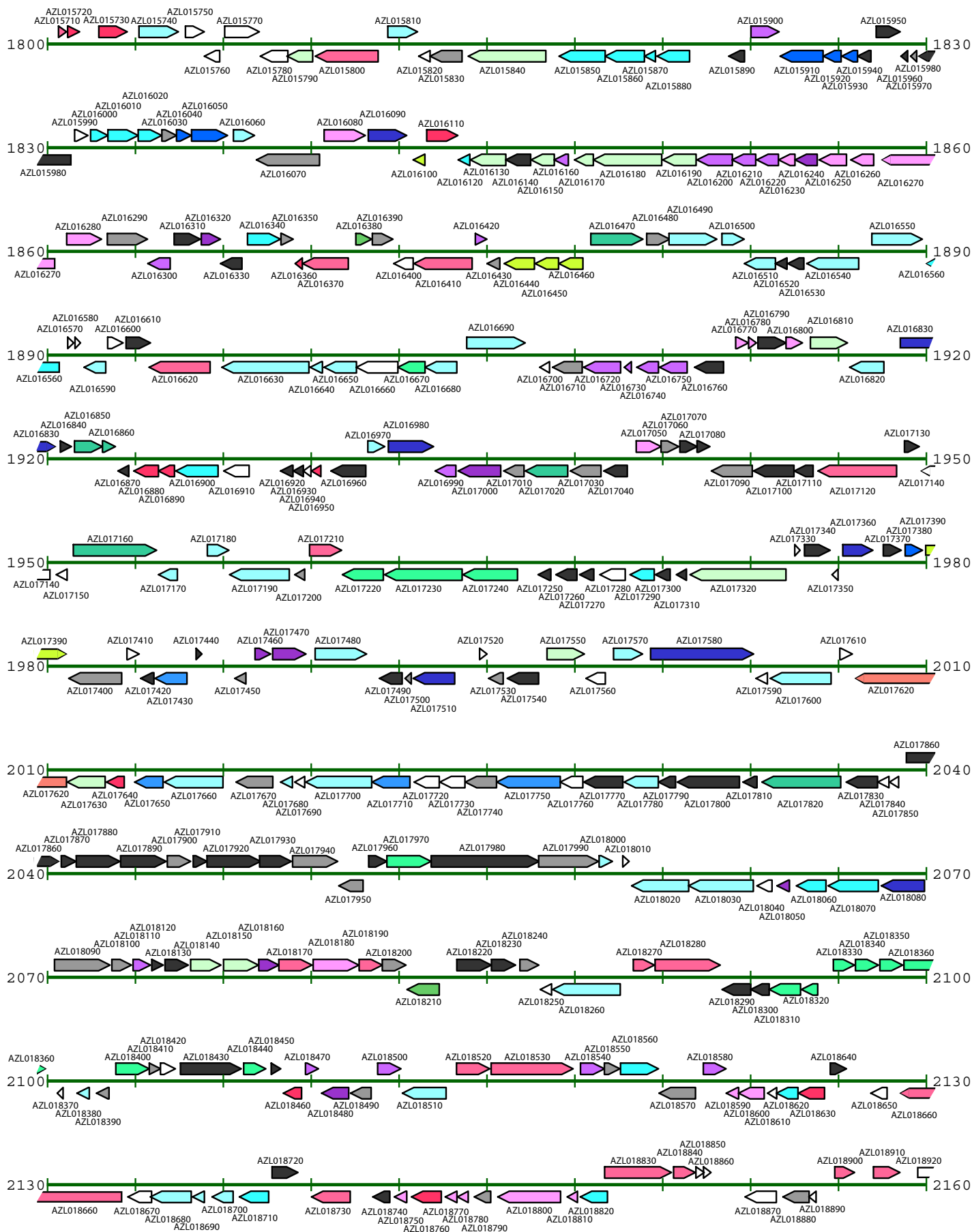
chromosome

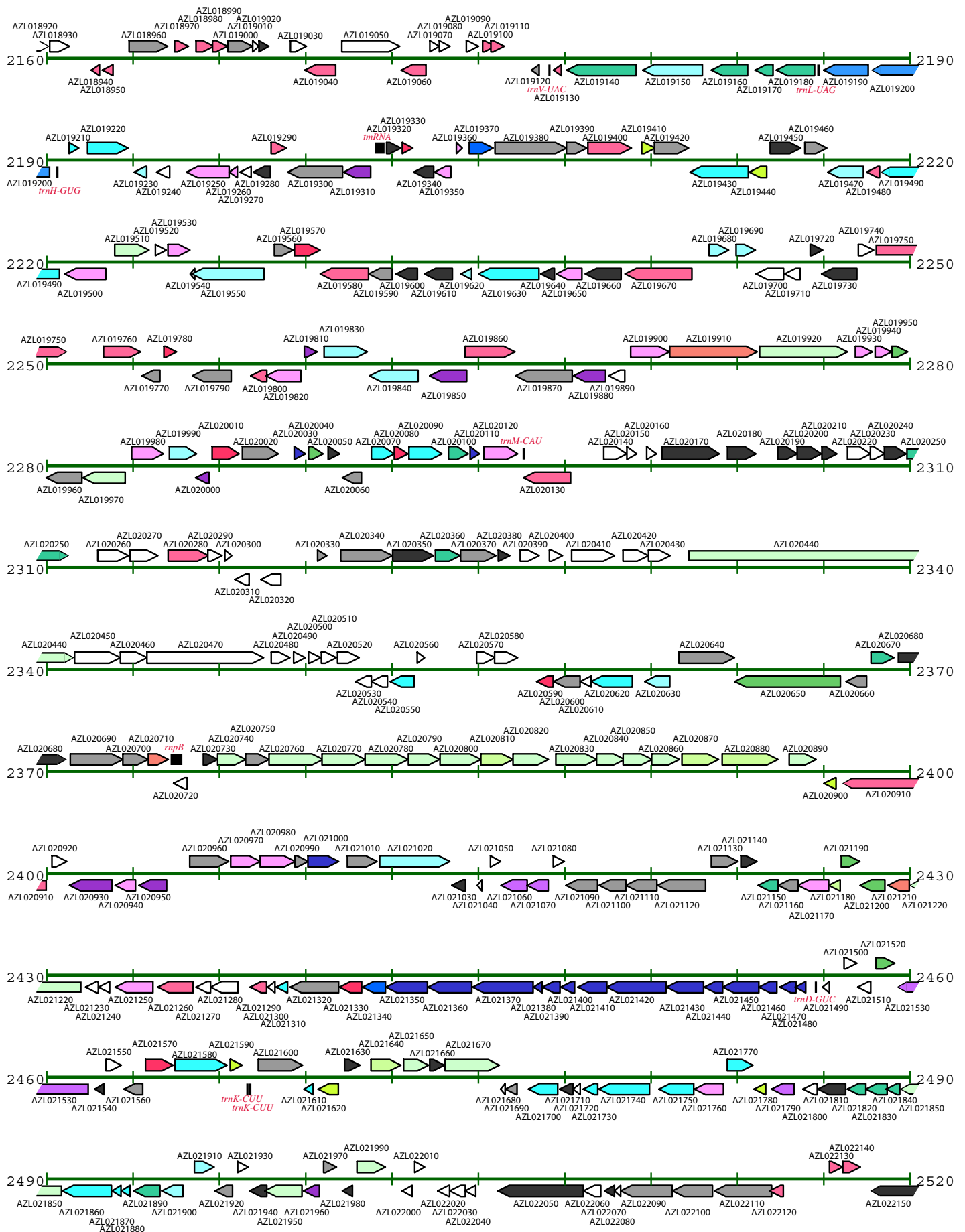


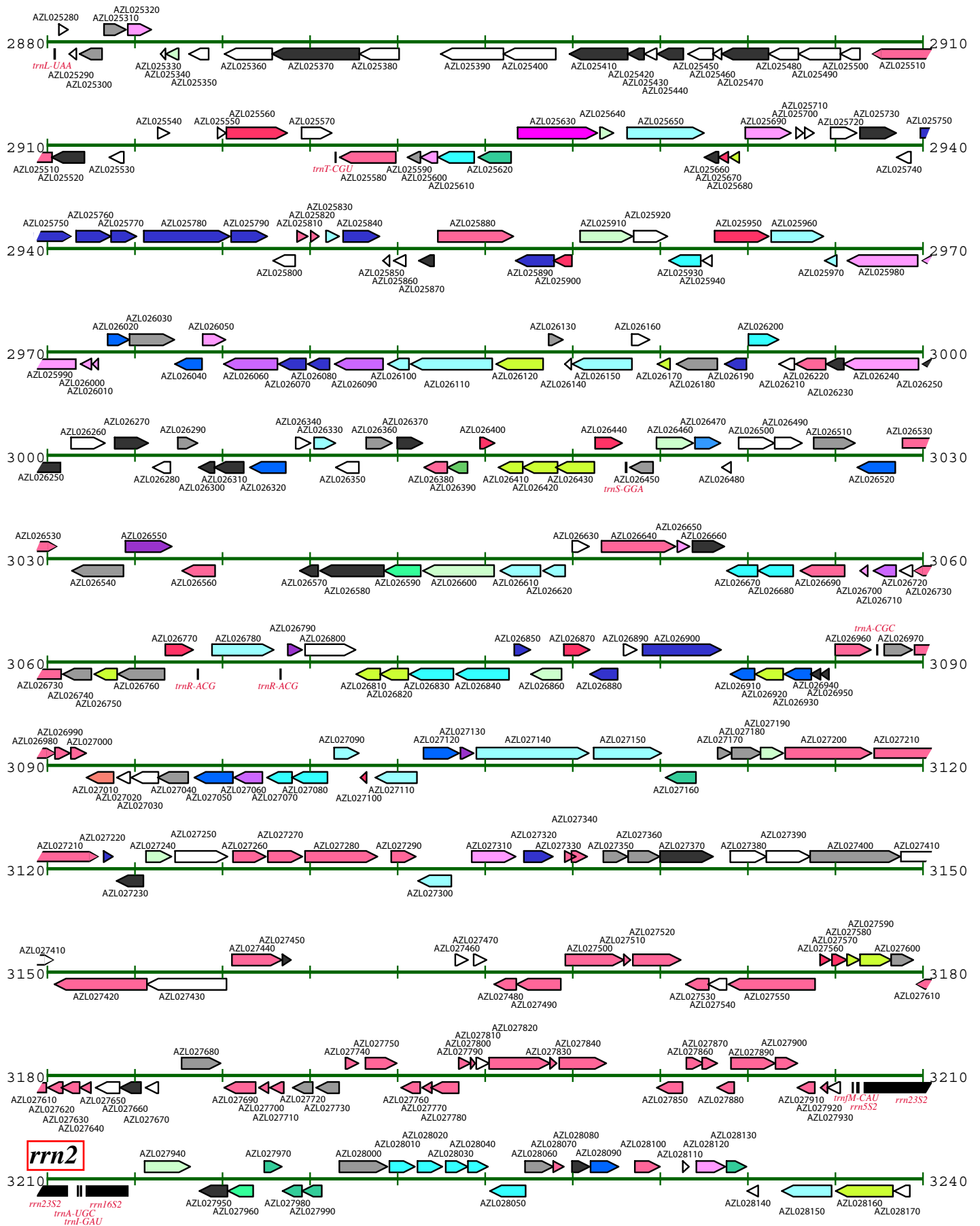


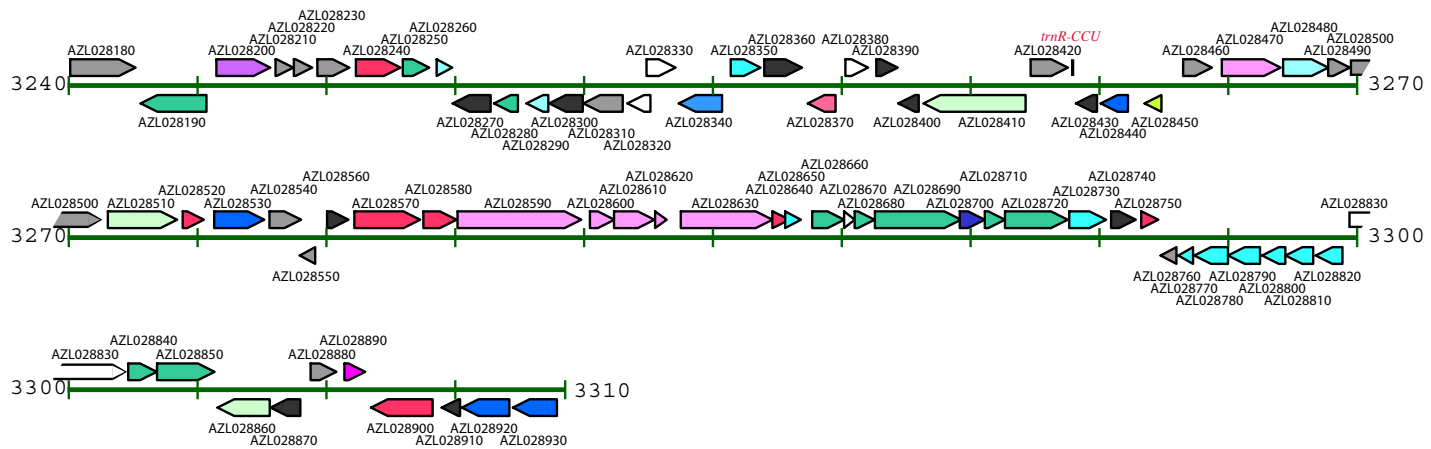




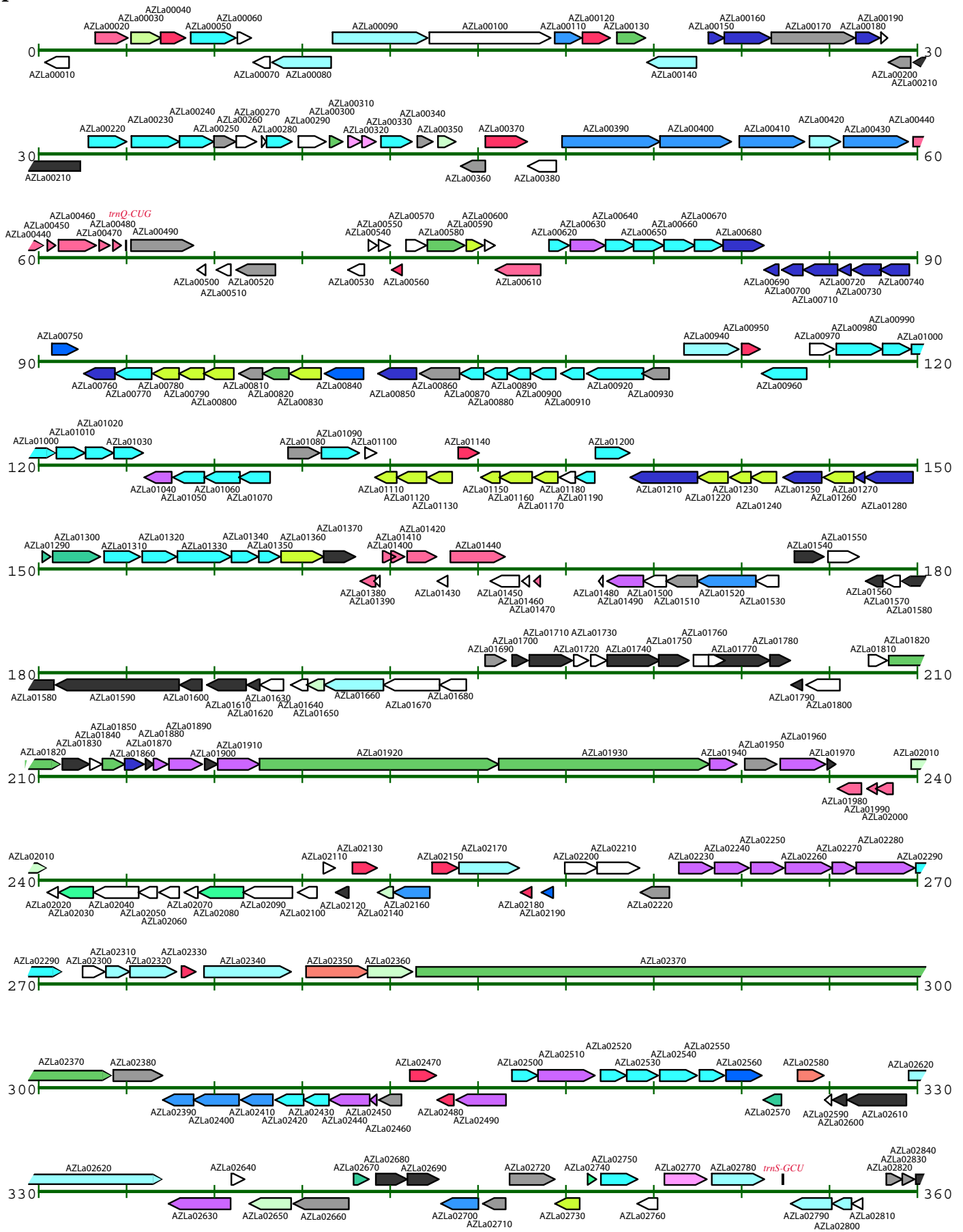


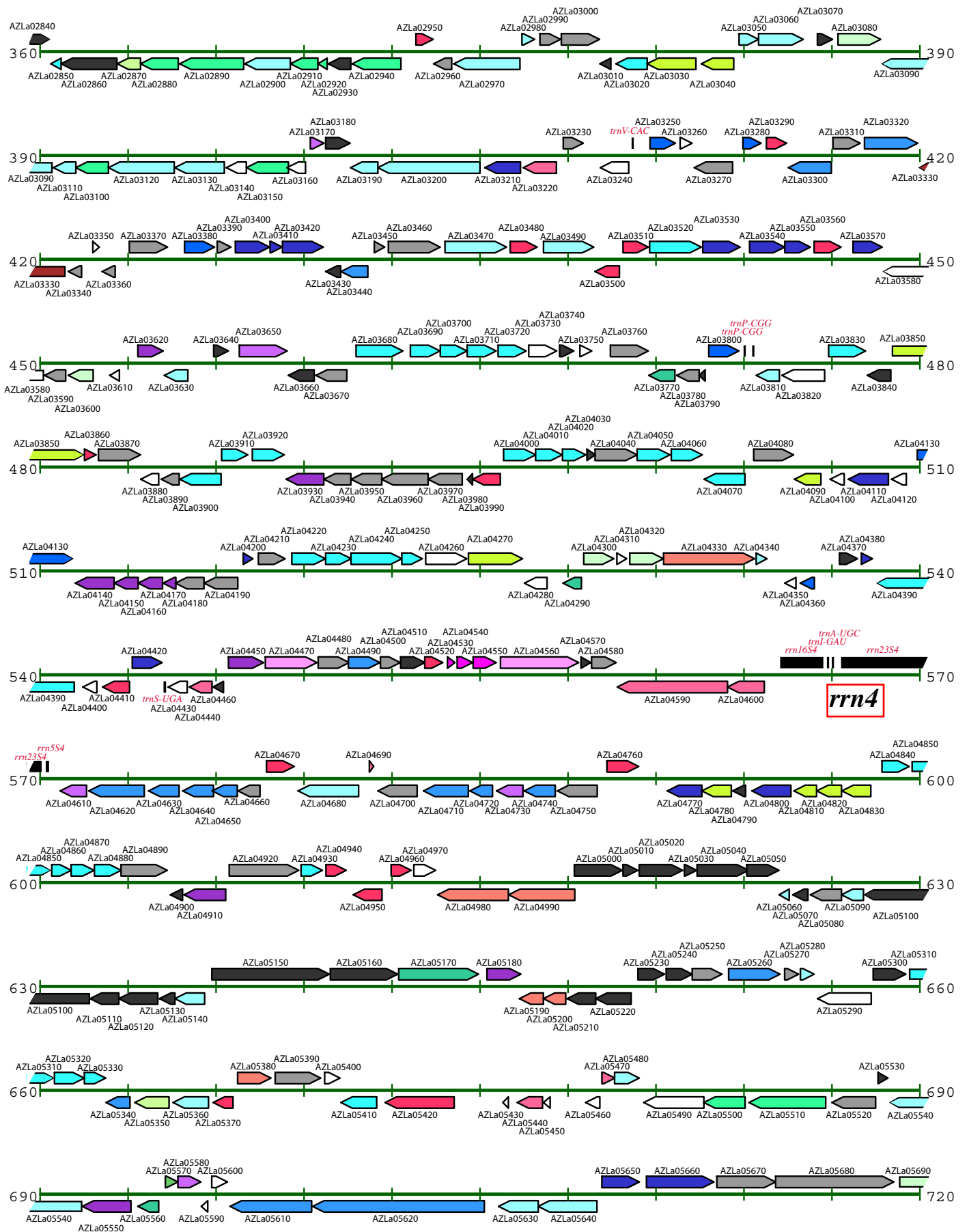


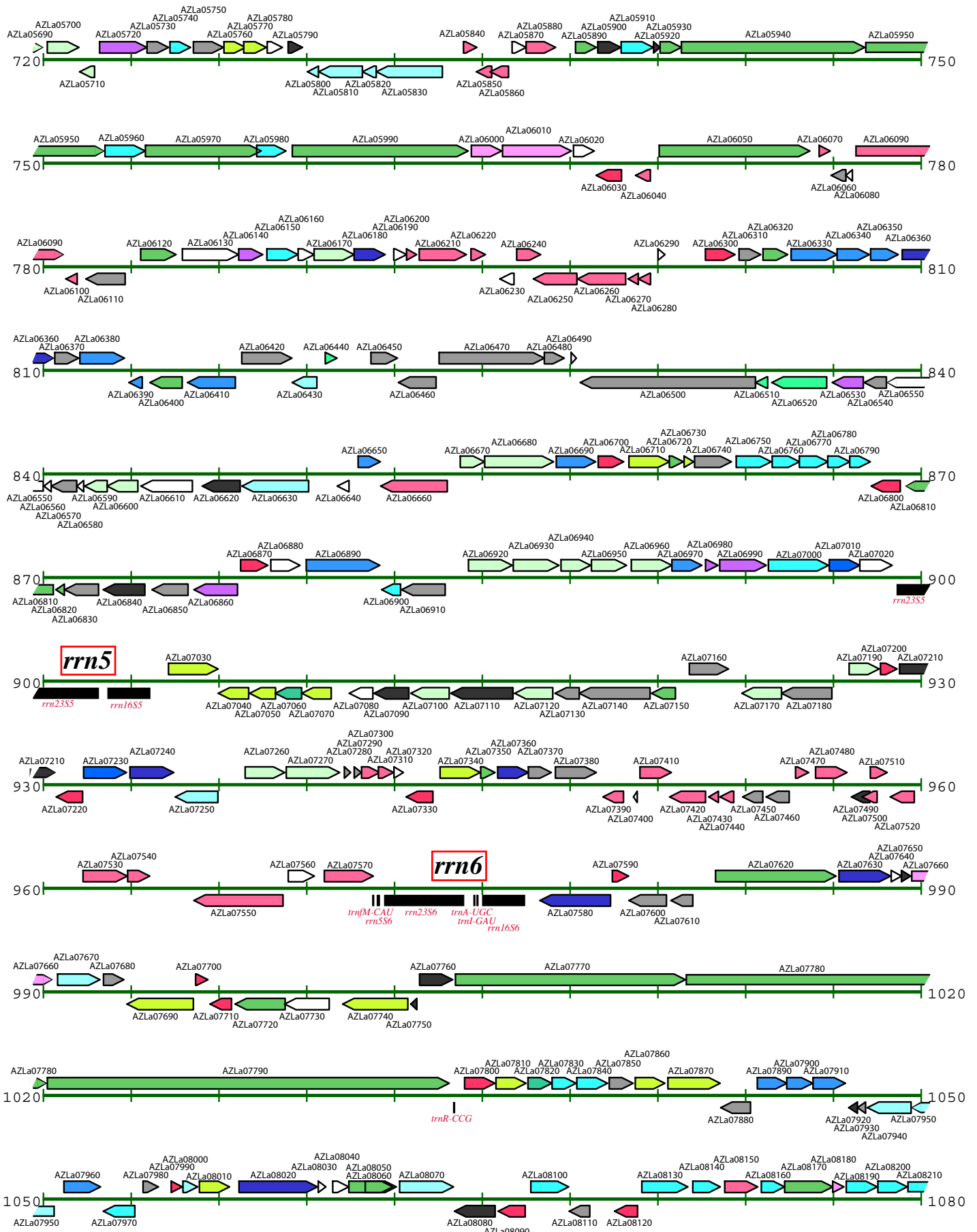


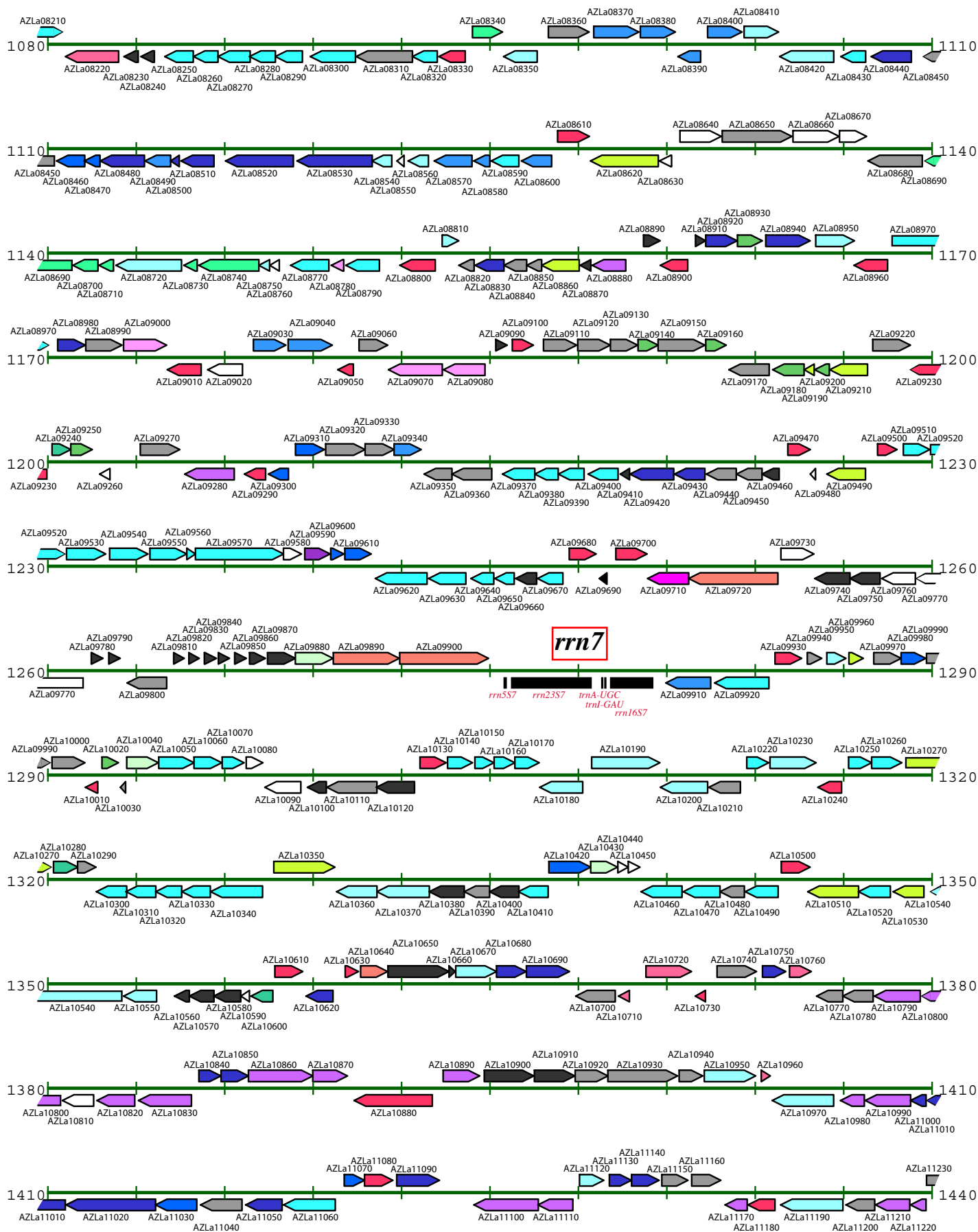


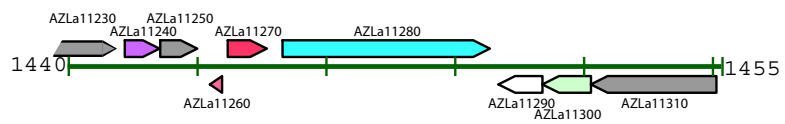
pAB510a



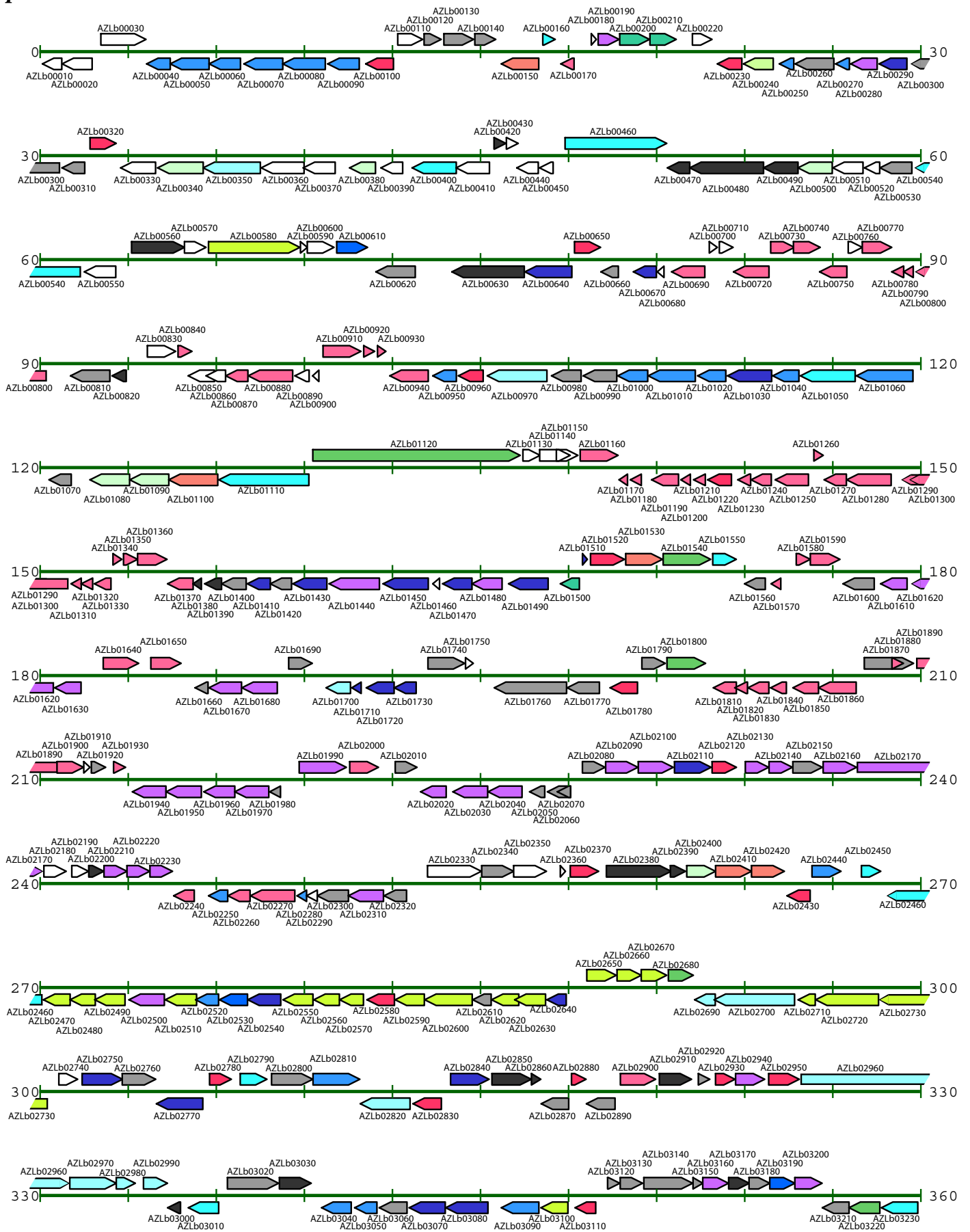


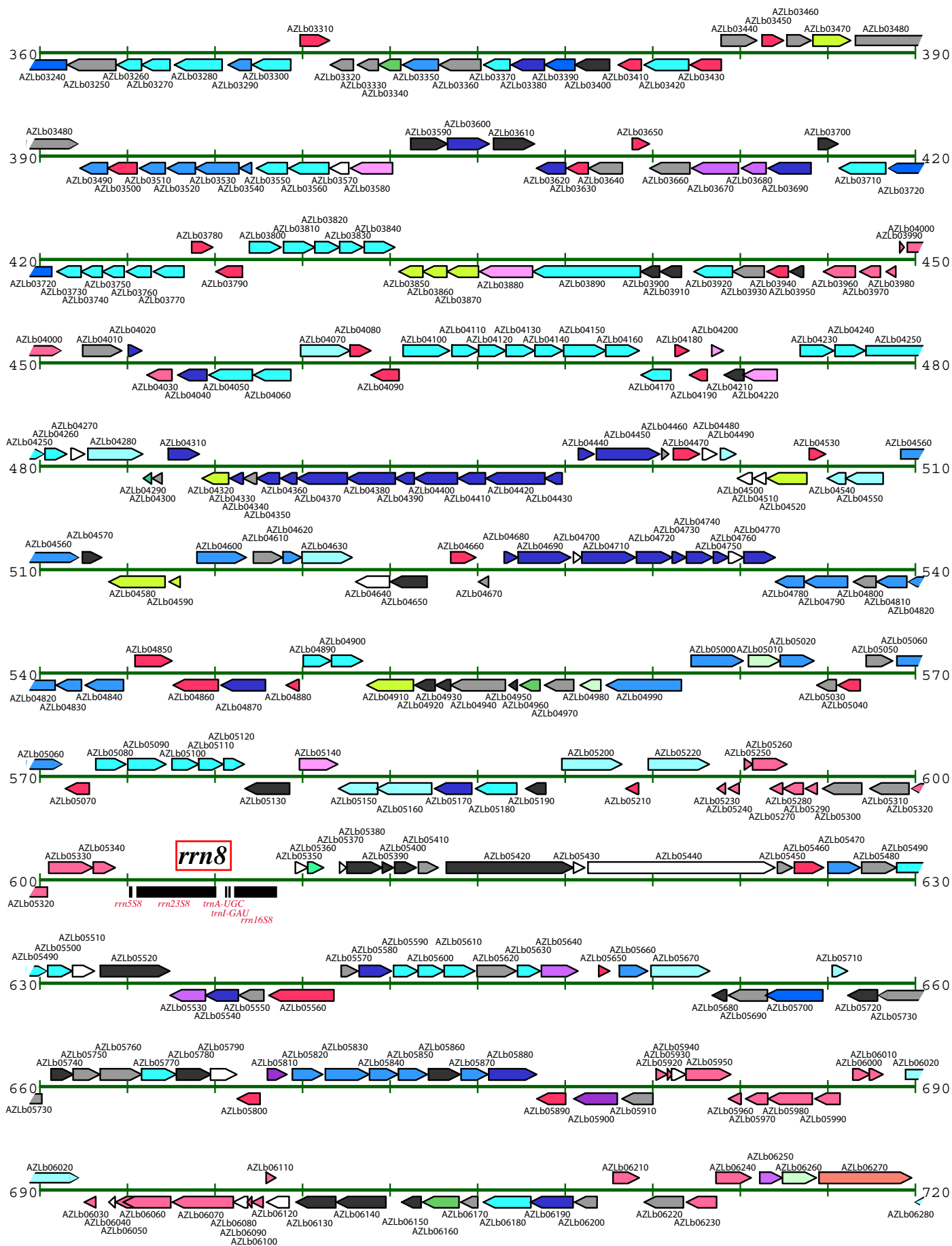


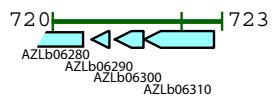




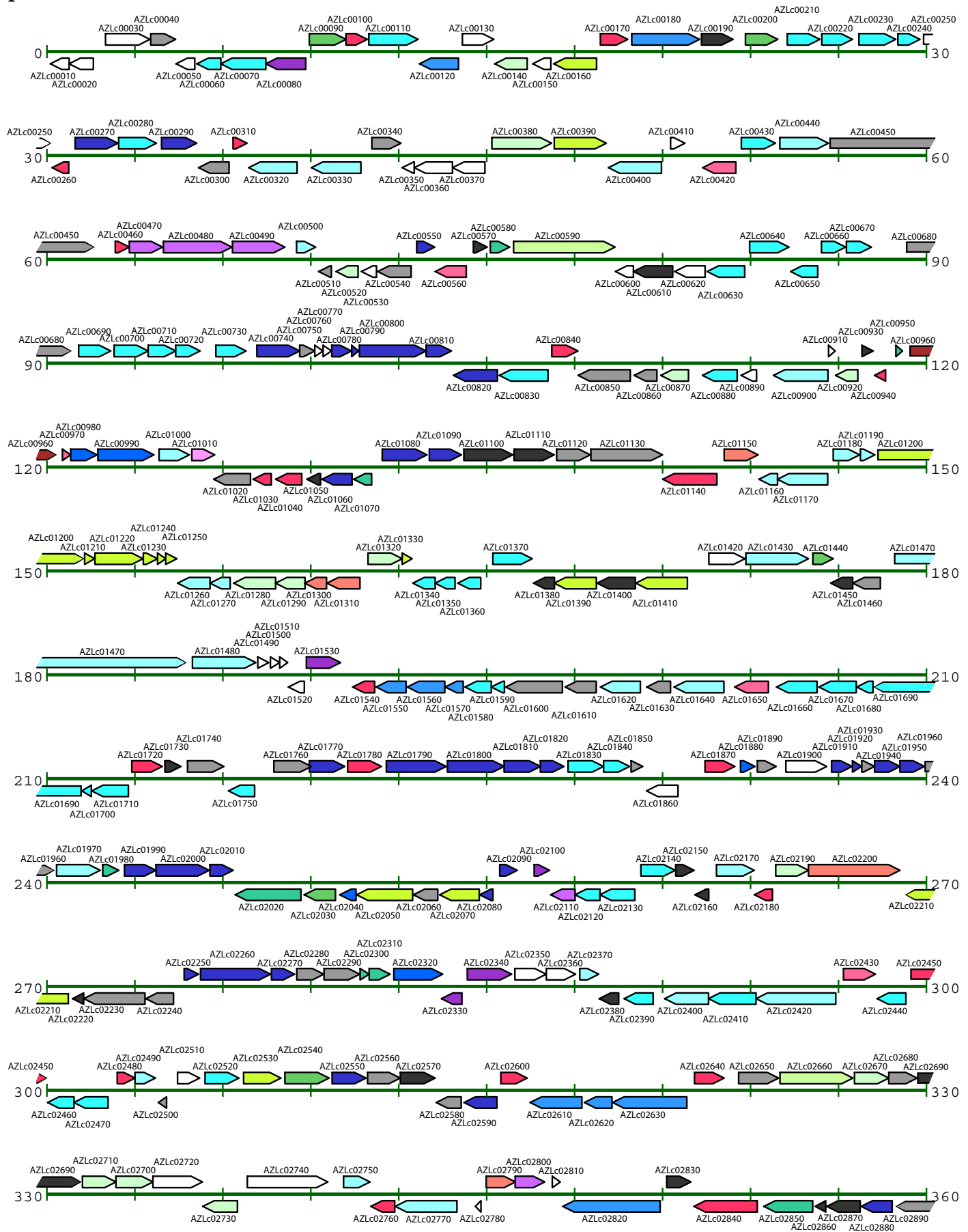
pAB510b

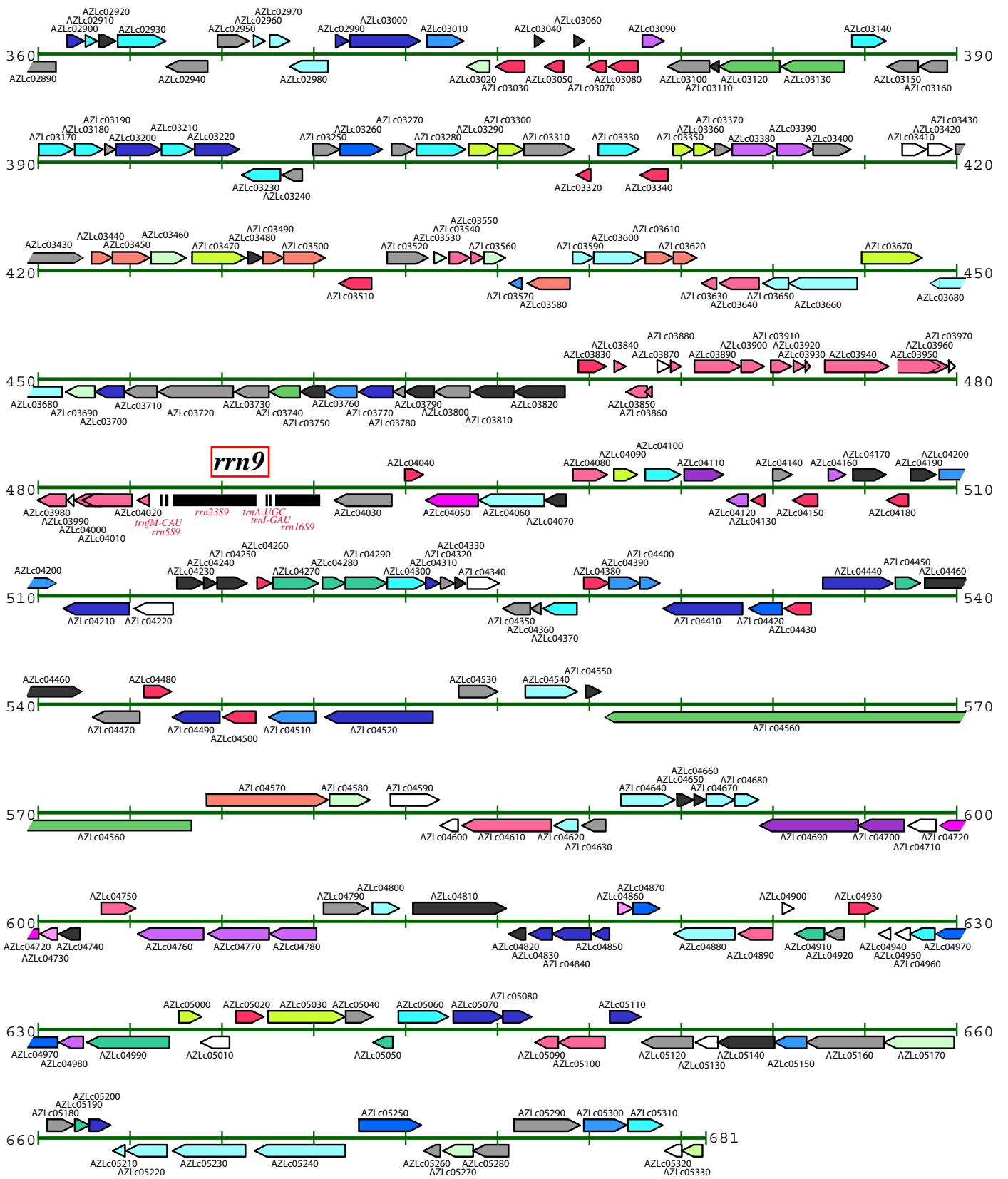




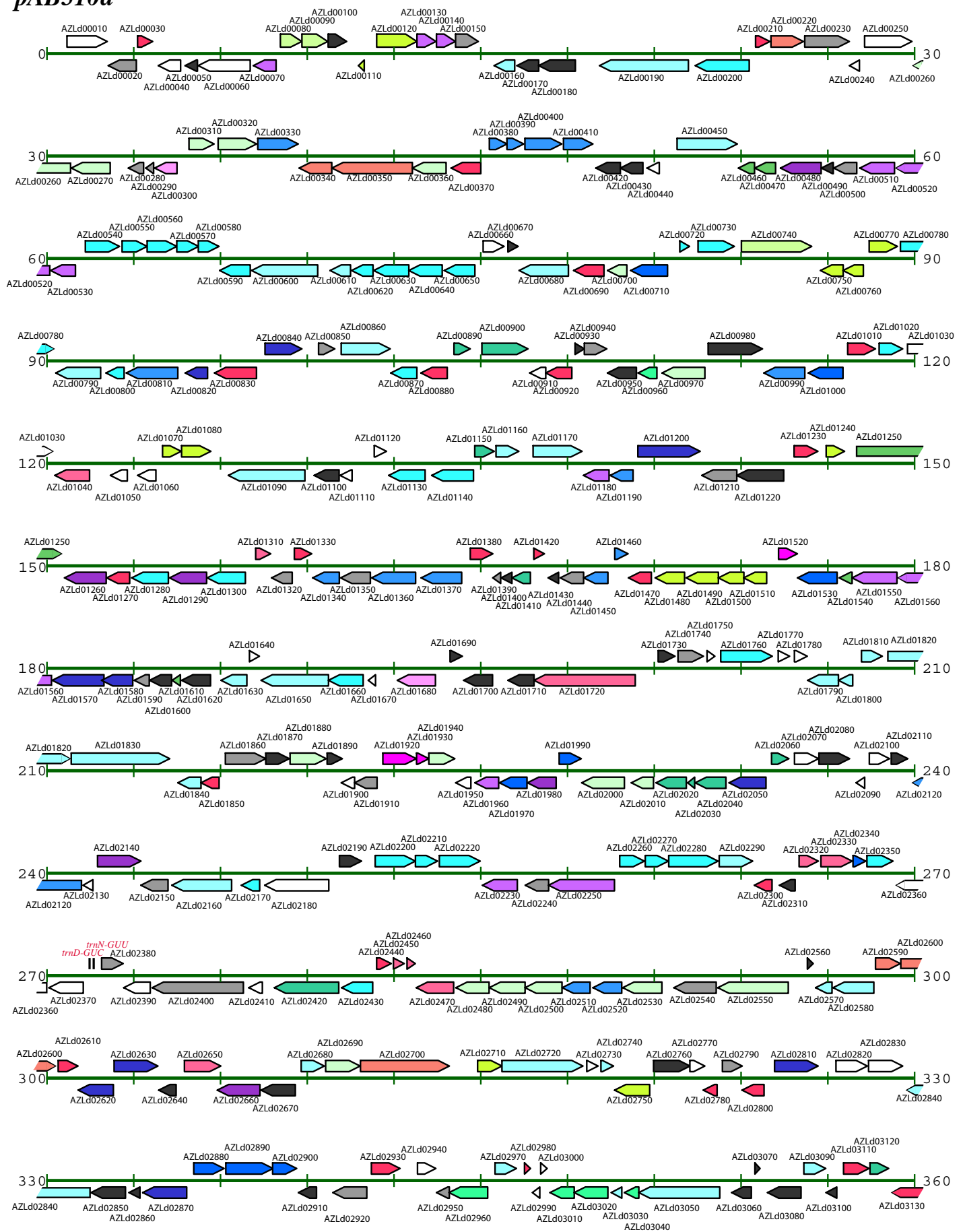


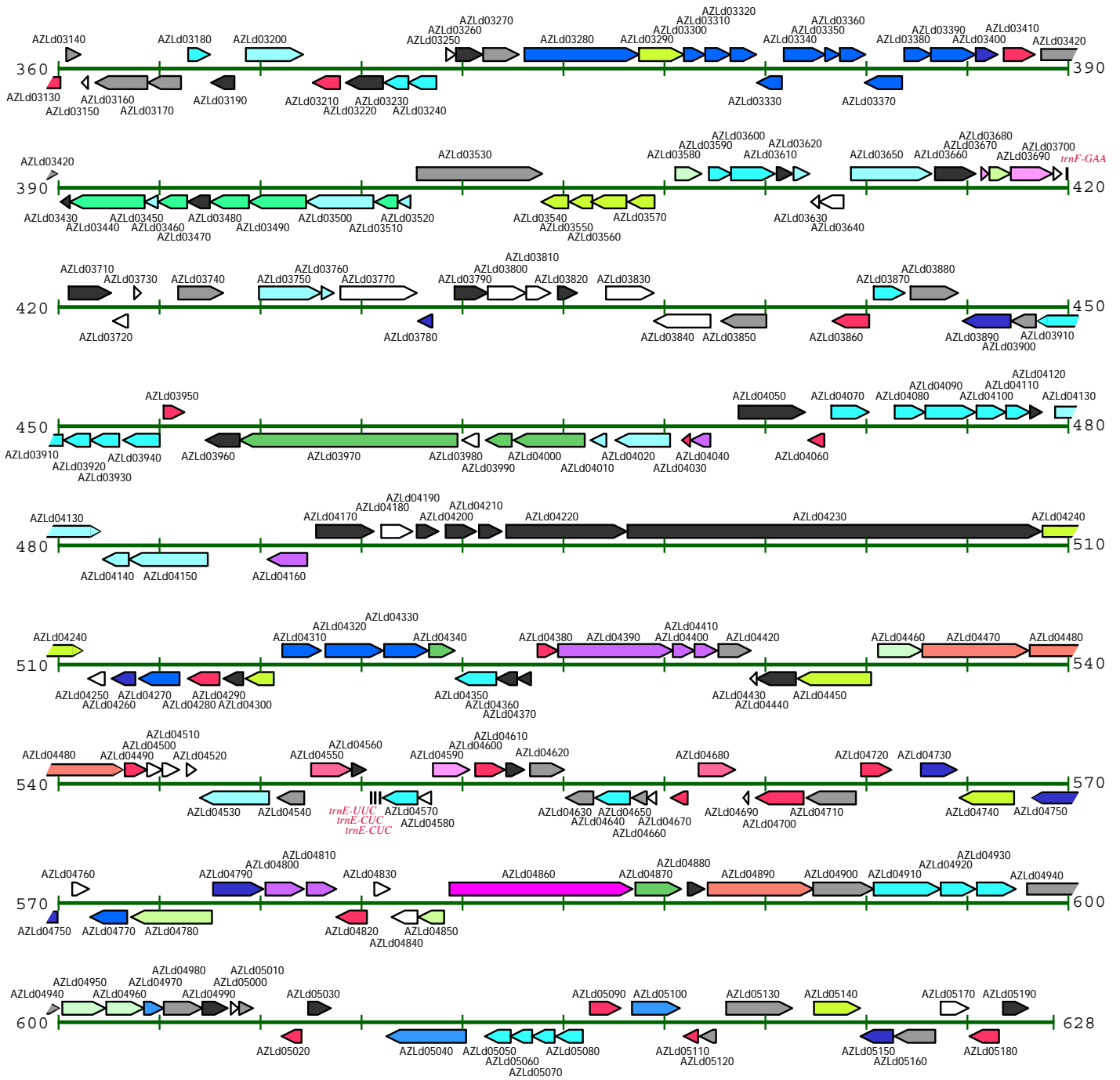
pAB510c



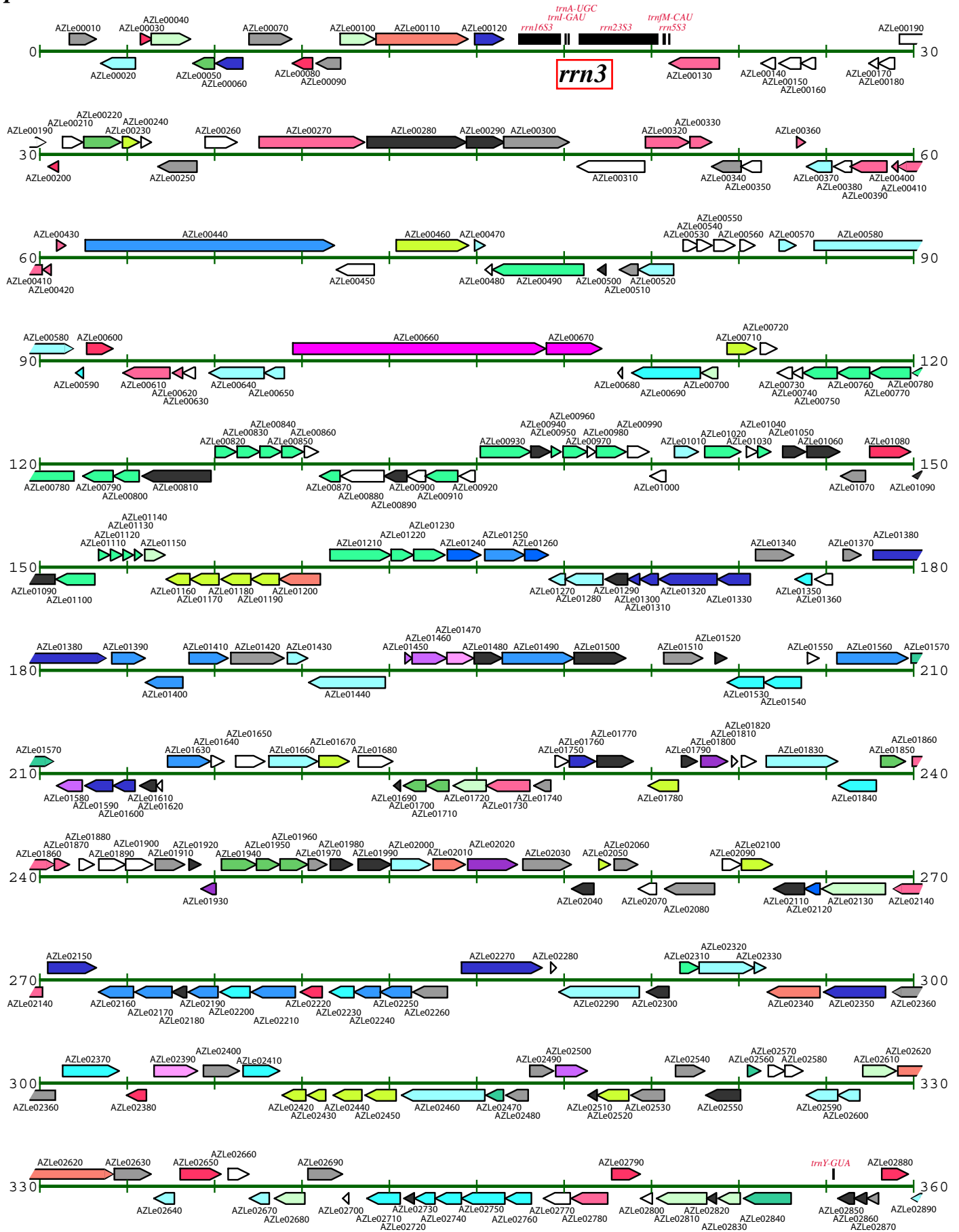


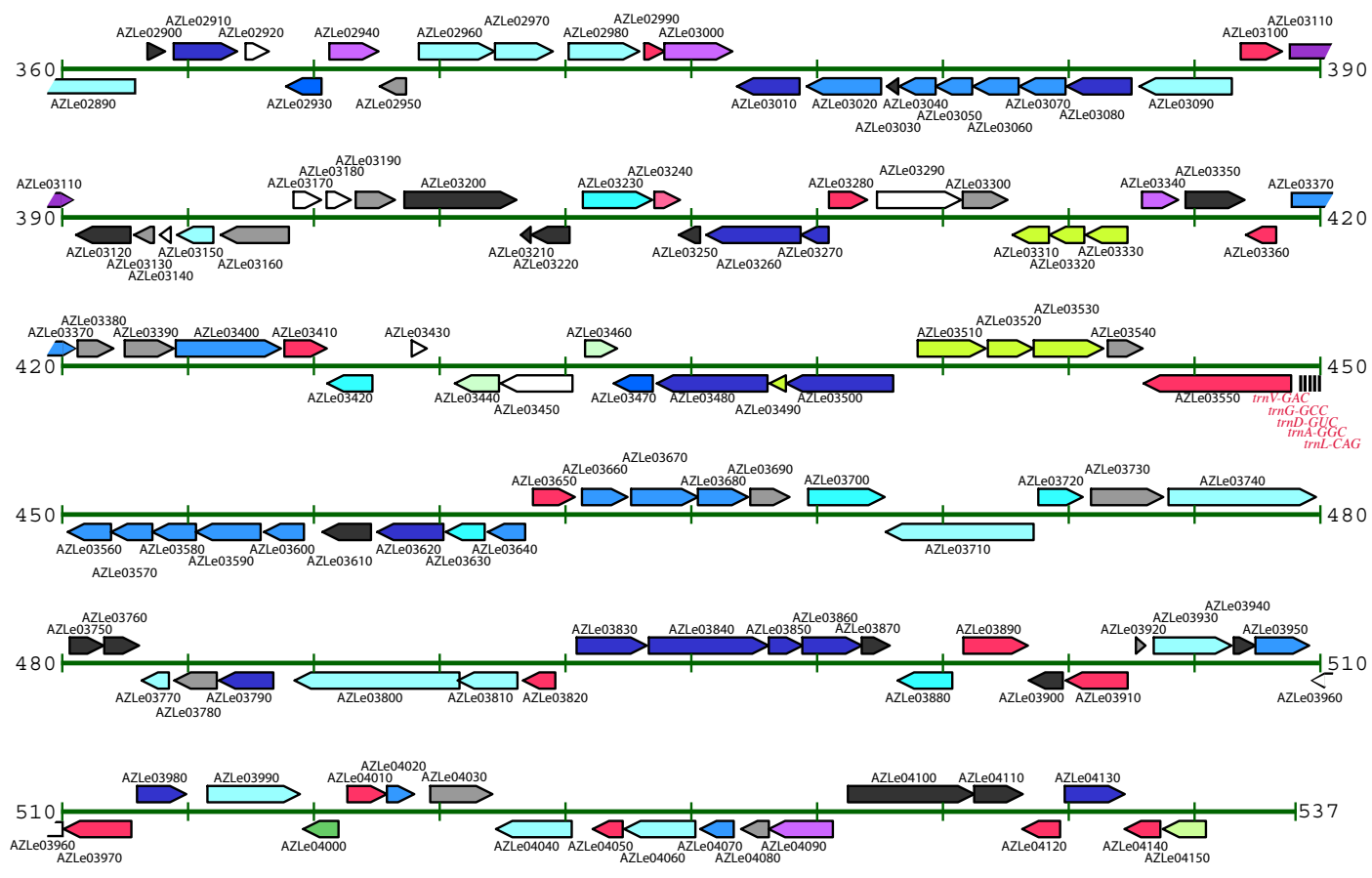
pAB510d



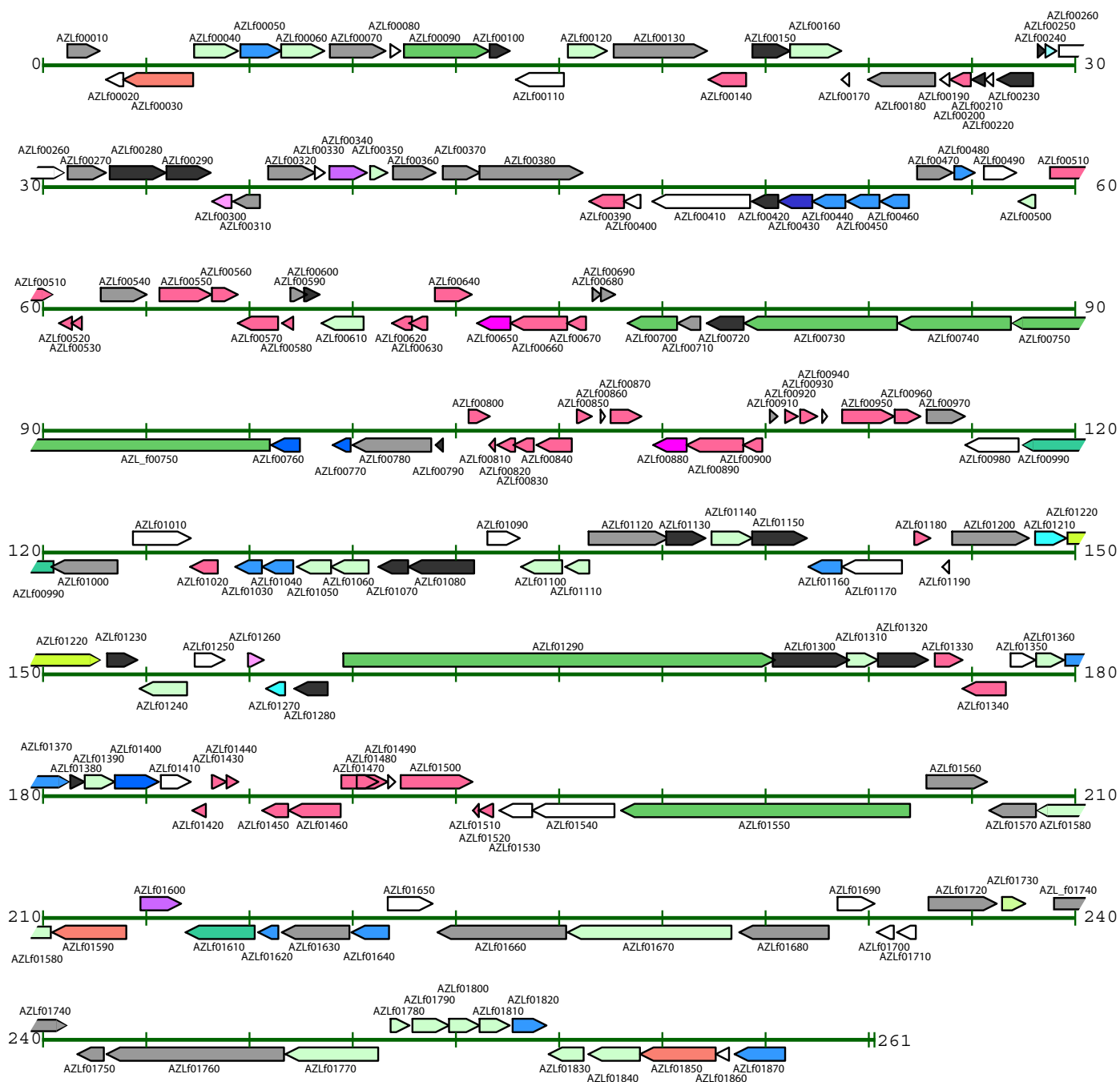


pAB510e





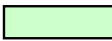



















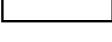


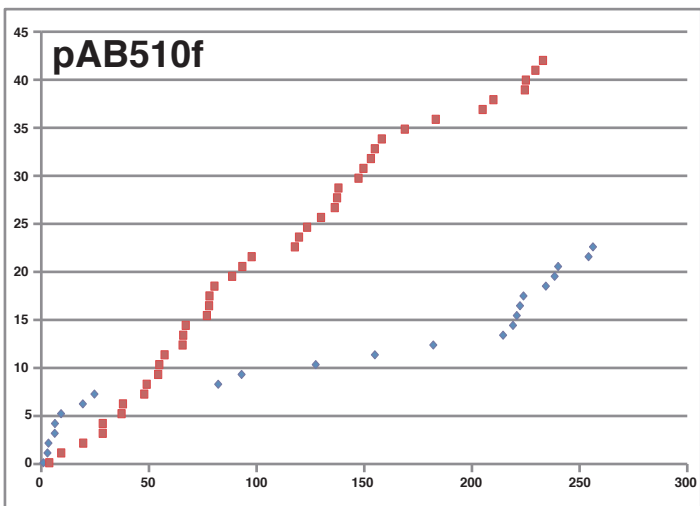
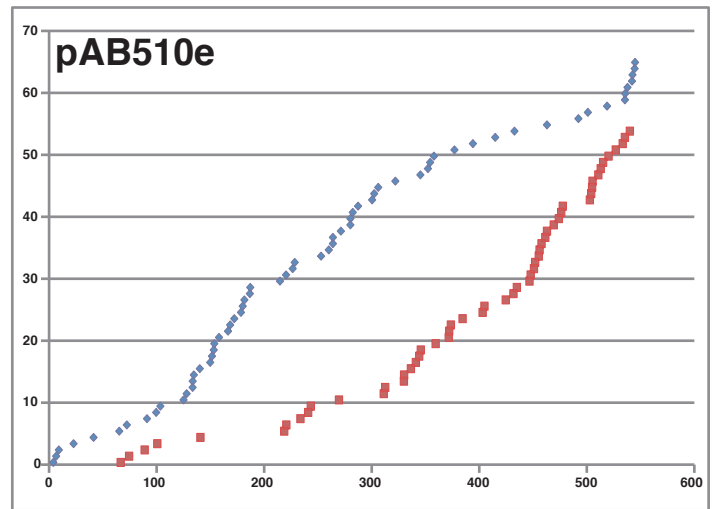
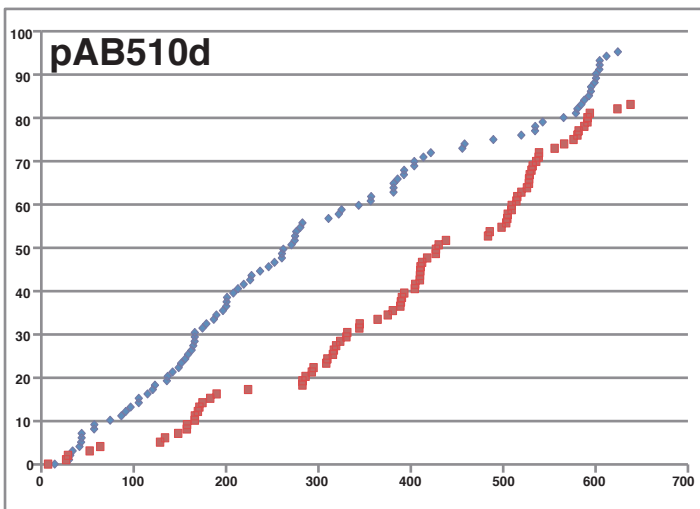
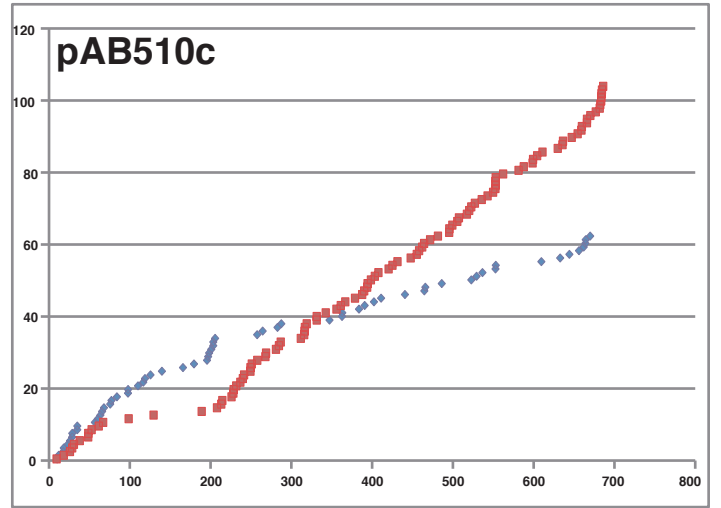
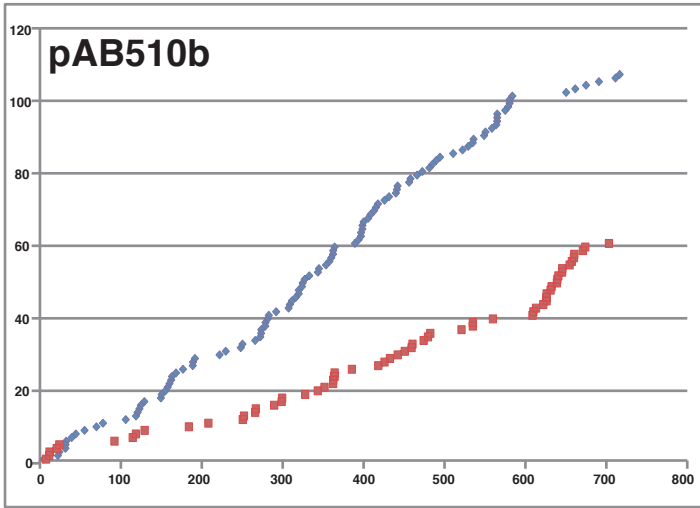
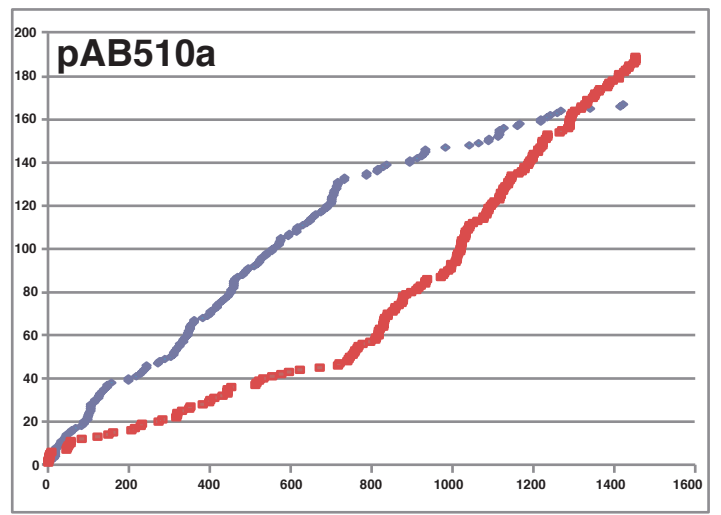
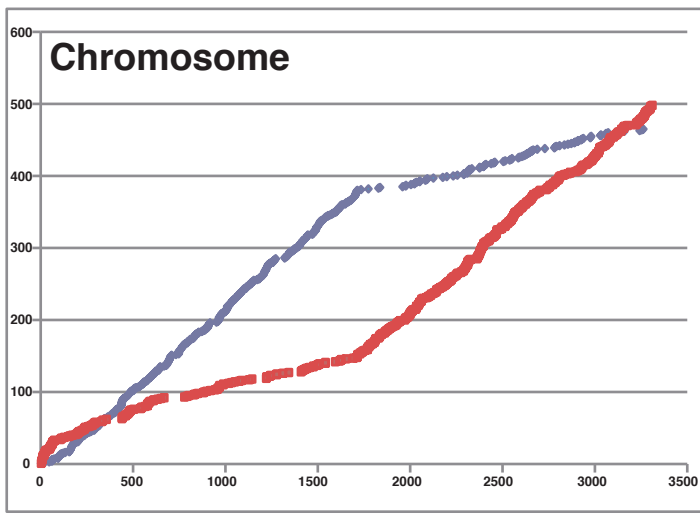


pAB510f



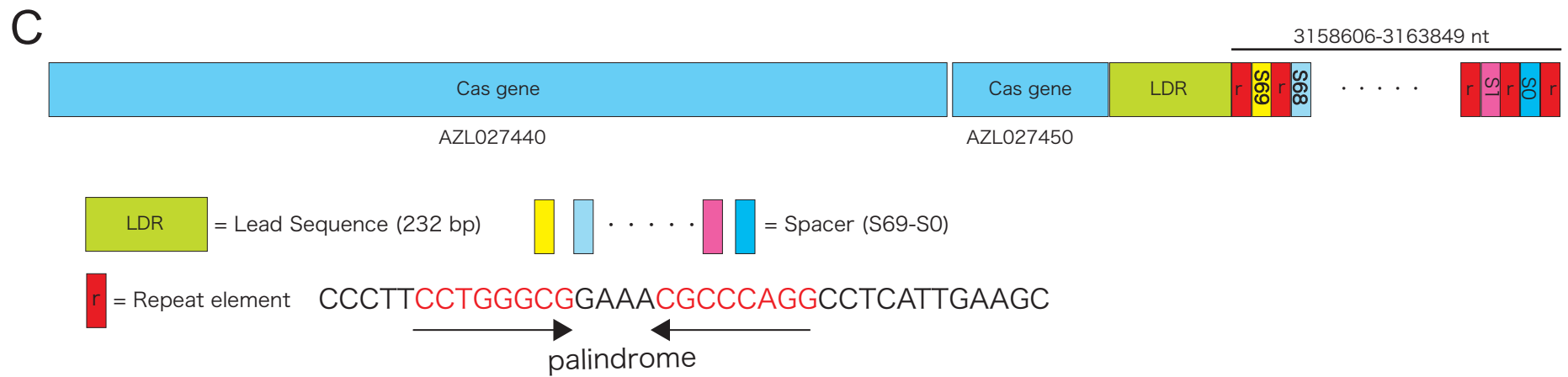
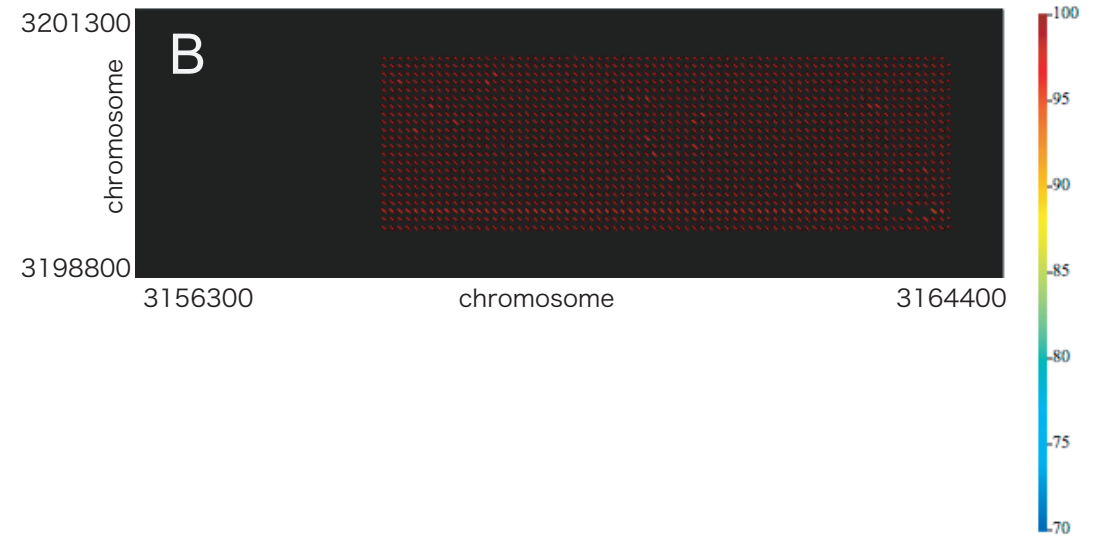
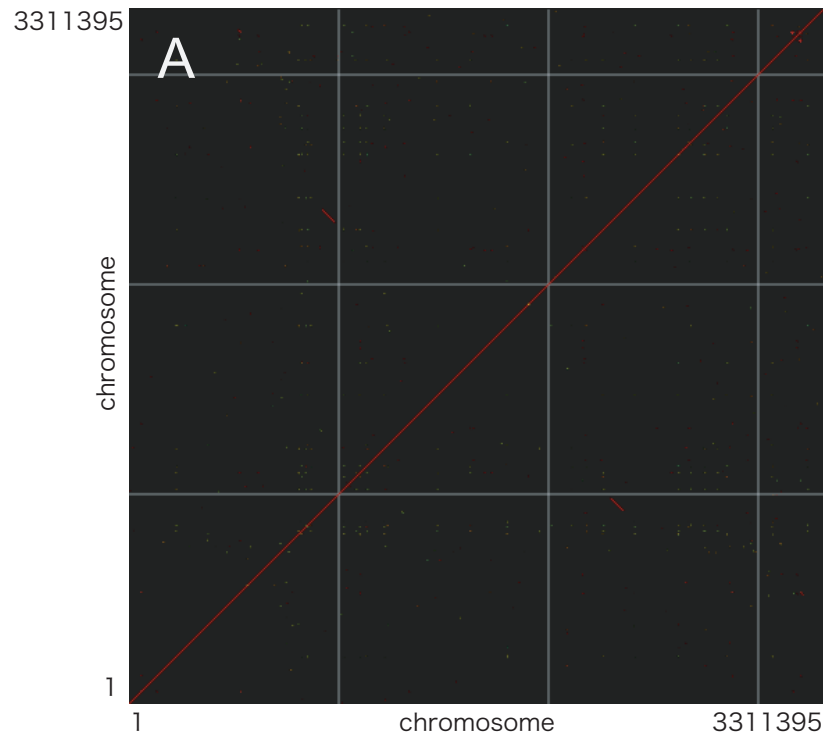
Color codes for functional categories

	Translation, ribosomal structure and biogenesis		Signal transduction mechanisms		Amino acid transport and metabolism
	RNA / RNA processing and modification		Cell wall /membrane/envelope biogenesis		Nucleotide transport and metabolism
	Transcription		Cell motility		Coenzyme transport and metabolism
	Replication, recombination and repair		Extracellular structures		Lipid transport and metabolism
	Chromatin structure and dynamics		Intracellular trafficking, secretion, and vesicular transport		Inorganic ion transport and metabolism
	Cell cycle control, cell division, chromosome partitioning		Posttranslational modification, protein turnover, chaperones		Secondary metabolites biosynthesis, transport and catabolism
	Nuclear structure		Energy production and conversion		General function prediction only
	Defense mechanisms		Carbohydrate transport and metabolism		Function unknown
					Not in COGs



X: Nucleotides from putative origin (Kbp)
 Y: Cumulative number of KOPS motifs
 ◆ KOPS (F)
 ■ KOPS (R)

Supplementary Figure 3



Supplementary Figure 4.

REPEAT
(Arrow: palindrome)

SPACER

3158606	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TACATCTGCGACATCGTTGCGCTTCGCGGTGGCCCGT	
3158680	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGCAAGGATCGAGTTGCTAGCTCCGCGAAGGAGG	3158679
3158754	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGCAAGGATCGAGTTGCTAGCTCCGCGAAGGAGG	3158753
3158830	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGCAAGGATCGAGTTGCTAGCTCCGCGAAGGAGG	3158829
3158901	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TCAGCGTCCTTAGGATCGTCCGTGAGAAGGTCGGCCAT	3158900
3158976	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGCAGGATGGACCCAACCTCGAACTGGAAGAGGCCG	3158975
3159049	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGTCTGCCAAGTCCATGCGGGCGAAAGAGACGGAT	3159048
3159123	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GTTCGGAAATATTGATCCGACCCCATTCGGATCAT	3159122
3159196	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGGCCCGCCGCGTGC TGGGCGCCTCCAATGTCGGACC	3159195
3159270	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GGCGTCTGCCGTCTCCAGCCACTTCGCCAGCACTCC	3159269
3159343	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GGGTGAAGCGGGGTGCTGGTAGGGGATGGGTGGTAG	3159342
3159418	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CCCGCGGCACCGGATCGAGGGGCCGCGGCGCAACGGC	3159417
3159492	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	ATCACAATGAACGGCGAACCACCTCGCCCACCCTTG	3159491
3159565	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGGGCGCGGGGAGGGGGAGCGCCCCCATTCCTTCAGCC	3159564
3159641	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GCCATGTCGCCGTCGTGCTGTGGTGTCAATCGTGTGCCACTT	3159640
3159720	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC		3159719
3159794	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TCAGTCGGTGCGGATACCGCCTTGGCGATCACAGAGACGG	3159793
3159871			3159870

CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3159944
3159945	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160018
3160019	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160090
3160091	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160164
3160165	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160238
3160239	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160313
3160314	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160391
3160392	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160464
3160465	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160538
3160539	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160612
3160613	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160685
3160686	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160760
3160761	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160834
3160835	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160907
3160908	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3160982
3160983	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3161057
3161058	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3161130
3161131	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3161204
3161205	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3161277
3161278	
CCCTTCCTGGGCGGAAACGCCCAGGCCTCATTGAAGC	3161354
3161355	

CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	3161426
3161427	
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGAGGGAAACCACCAGCCGAGCAACGTCGCTCGGCTTG
3161503	3161502
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GACGCATGAAGTTCGTACCGGTGGCGTCCCGCATATC
3161579	3161578
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CCCTGCCACCCTGATCTGCAGCTTGTGCATGTCGAGCCG
3161655	3161654
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGCCGCTCTATGAGTTCGTGCCACGCGGTCCATCAGC
3161731	3161730
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGGCCATGGTCTATGACAGGAACGCCCTGGCCCTGCA
3161807	3161806
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	ATAAGAGGCTCTGGGCGGAAACGCCCTGGCCCTGCA
3161885	3161884
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GCCATTCCCTTGTGCTCTGGTCAAATGGGCAA
3161957	3161956
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGCCATACAGGCCAGGACTTCAGACCTTCTGGGAAATC
3162034	3162033
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GACTCCCAACGGACGCTGACGAAATGGGCGCACGCGC
3162110	3162109
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TTGAACACTGCGGCGGAACACCCCATCGTGGTAACG
3162183	3162182
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AATGATATGGCTCTCTCGGGCTTCGTCAATGTAGACCC
3162258	3162257
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGCAAGAATATGTGGATCGAAACATCCTCAATCTGAA
3162334	3162333
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AACTTCACGATCTCAATGGCTGTTTTGTCAATCAATAA
3162408	3162407
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGCAGTCCAGCGGATGCGGCGGTGGTTGGACAACCGC
3162482	3162481
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGCCCCAGGTAGCCCAGCTTGTAGGTCTGGCTGGCTTC
3162557	3162556
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGTCCGACAGGTGGCGCCCGACAGGTGGGCGCGC
3162630	3162629
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CCGGGTCAGGTTAGCCATGAGACGGCCATCCGGCA
3162703	3162702
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GACACAGACAGCAACAAGGCCACCTCATCGGGTGC
3162775	3162774
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	SGTNTTGGCTCAGTTCCCAACGCTGCGGCATAATC
3162848	3162847

CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TCTACCGGCAGCAATGCCGGCAGATCCTGATCGGGG	3162920
3162921		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGAATGCTGGGGCATCATGTCCGGCCGTATCCCGA	3162993
3162994		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CCGTGGAAAGTGACCGATCTGCGTTTGCTATGAACTGAG	3163069
3163070		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGTCAACCCCAATCAGCCAGACCCTGGGTGCGGGCT	3163142
3163143		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGAATGCGCATTGTGTTCTCCATTTCTCAGGCGGGCC	3163217
3163218		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC		3163291
3163292		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AAATGGAGCGTCAGGCGCTCCAGGGGGAGCGGTCCGA	3163366
3163367		
CTCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CACAAGGCGATAGCGGCGCCCCAGCTCGGCCCGCG	3163438
3163439		
CTCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TTGCTTTCGCGCCCTGCGCCGTCATCCTTTGGGGCG	3163512
3163513		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AAATGCAGCGTCAGGCGCTCCAGGGGGAGCCGTCGGA	3163587
3163588		
CTCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CACAAGGCGATAGCGGCGCCCCAGCTCGGCCCGCG	3163659
3163660		
CTCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AAGGCAGTAGAGGCGCCGGCGGTCCGCTCCGCCGGAACAG	3163736
3163737		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	ACCTCCACCAGGTGTACTGCCGCAAAGTCTCGCCACC	3163812
3163813		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC		

Supplementary Figure 5

Aminoacyl Stem	D domain		Anticodon domain		Variable region	TW domain		Aminoacyl stem
	stem	loop	stem	loop		stem	loop	
AZL_ra01	Type: Gln		Anticodon: CTG (62973-63046)					
TGGGGGA TC	GTC	AGCGGTA	GGAC A	GCGGA	CTCTGAC	TCCGC	CAGC	CTAGG TTCGAAT CCTAG TCCCCA ACCA
AZL_ra02	Type: Ser		Anticodon: GCT (355389-355481)					
GGAGGGG TG	GCCG	AGTGGCTGA	AGGC A	ACGGT	TTGCTAA	ACCGT	CATAGGGTTTAAAGCCCTATC	GTGGG TTCGAAT CCCAT CCCCTCC GCCA
AZL_ra03	Type: Val		Anticodon: CAC (410196-410270)					
GGCGCGG TA	GCTC	AGCGGGA	GAGC A	TTCGC	TTCACAC	GCGAA	GGGTC	ACAGG TTCAACT CCTGT CGCGCCC ACCA
AZL_ra04	Type: Pro		Anticodon: CGG (474014-474090)					
CGGAGAG TA	GCTC	AGCCTGGTA	GAGC A	CTGCT	TTCGGGA	GGCAG	GGGCC	GGAGG TTCGAAT CCTCT CTCTCCG ACCA
AZL_ra05	Type: Pro		Anticodon: CGG (474308-474384)					
CGGAGAG TA	GCTC	AGCCTGGTA	GAGC A	CTGCT	TTCGGGA	GGCAG	GGGCC	GGAGG TTCGAAT CCTCT CTCTCCG ACCA
AZL_ra06	Type: Ser		Anticodon: TGA (544324-544235)					
GGATGGG TG	GCCG	AGTGGTTTA	AGGC A	GCGGT	CTTGAAA	ACCCG	CGTAGGTGCGAGCCTACC	GTGGG TTCGAAT CCCAC CCCATCC GCCA
AZL_ra08	Type: Ile		Anticodon: GAT (566855-566931)					
GGGCTAG TA	GCTC	AGTTGGTTA	GAGC G	CGCGC	TTGATAA	GCGTG	AGGTC	GGAGG TTCAAAT CCTCC CTGGCCC ACCA
AZL_ra09	Type: Ala		Anticodon: TGC (566961-567036)					
GGGGCGG TA	GCTC	AGTTGGGA	GAGC G	CCTGC	TTTCAA	GCAGG	AGGTC	GTCGG TTCGATC CCGTC TGCCCTCC ACCA
AZL_ra14	Type: fMet		Anticodon: CAT (971336-971260)					
CGCGGGG TG	GAGC	AGCCCGGTA	GCTC G	TCAGG	CTCATAA	CCTGA	AGGCC	GCAGG TTCAAAT CCTGC CCCCAC ACCA
AZL_ra17	Type: Ala		Anticodon: TGC (974794-974719)					
GGGGCCA TA	GCTC	AGTTGGGA	GAGC G	CCTGC	TTTCAA	GCAGG	AGGTC	GTCGG TTCGATC CCGTC TGCCCTCC ACCA
AZL_ra18	Type: Ile		Anticodon: GAT (974899-974823)					
GGGCTAG TA	GCTC	AGTTGGTTA	GAGC G	CGCGC	TTGATAA	GCGTG	AGGTC	GGAGG TTCAAAT CCTCC CTGGCC ACCA
AZL_ra20	Type: Arg		Anticodon: CCG (1034101-1034025)					
GGACCCG TA	GCTC	AGCTGGATA	GAGC G	CTGCC	CTCCGAA	GGCAG	AGGTC	ATGAG TTCGAAT CTCGT CGGGTC GCCA
AZL_ra23	Type: Ala		Anticodon: TGC (1278873-1278798)					
GGGGCCA TA	GCTC	AGTTGGGA	GAGC G	CCTGC	TTTCAA	GCAGG	AGGTC	GTCGG TTCGATC CCGTC TGCCCTCC ACCA
AZL_ra24	Type: Ile		Anticodon: GAT (1278979-1278903)					
GGGCTAG TA	GCTC	AGTTGGTTA	GAGC G	CGCGC	TTGATAA	GCGTG	AGGTC	GGAGG TTCAAAT CCTCC CTGGCCC ACCA
AZL_rb03	Type: Ala		Anticodon: TGC (606446-606371)					
GGGGCCA TA	GCTC	AGTTGGGA	GAGC G	CCTGC	TTTCAA	GCAGG	AGGTC	GTCGG TTCGATC CCGTC TGCCCTCC ACCA
AZL_rb04	Type: Ile		Anticodon: GAT (606563-606487)					
GGGCTAG TA	GCTC	AGTTGGTTA	GAGC G	CGCGC	TTGATAA	GCGTG	AGGTC	GGAGG TTCAAAT CCTCC CTGGCCC ACCA
AZL_rc01	Type: fMet		Anticodon: CAT (484080-484004)					
CGCGGGG TG	GAGC	AGCCCGGTA	GCTC G	TCAGG	CTCATAA	CCTGA	AGGCC	GCAGG TTCAAAT CCTGC CCCCAC ACCA
AZL_rc04	Type: Ala		Anticodon: TGC (487527-487452)					
GGGGCCA TA	GCTC	AGTTGGGA	GAGC G	CCTGC	TTTCAA	GCAGG	AGGTC	GTCGG TTCGATC CCGTC TGCCCTCC ACCA
AZL_rc05	Type: Ile		Anticodon: GAT (487644-487568)					
GGGCTAG TA	GCTC	AGTTGGTTA	GAGC G	CGCGC	TTGATAA	GCGTG	AGGTC	GGAGG TTCAAAT CCTCC CTGGCCC ACCA
AZL_rd01	Type: Asn		Anticodon: GTT (271587-271662)					
TCCCGGT TA	GTC	AGAGCCTA	GGAC A	GCGGA	CTGTAAA	TCCGC	TTGTC	GCTGG TTCGAT CACAG ACCGGGA GCCA
AZL_rd02	Type: Asp		Anticodon: TGC (271667-271743)					
GCGGGTG TA	GCTC	AGTTGGTTA	GAGC G	CCGCG	CTGTAC	GCCGG	AGGCC	GCGGG TTCAGT CCGCT CACTCGC GCCA
AZL_rd03	Type: Phe		Anticodon: GAA (419945-420020)					
GCCCCAG TA	GCTC	AGTTGGTA	GAGC A	GGGGA	CTGAAA	TCCCC	GTGTC	GGCGG TTCAAAT CCGTC CCTGGCC ACCA
AZL_rd04	Type: Glu		Anticodon: TTC (549351-549276)					
GTCCCCG TC	GTCT	AGAGCCTA	GGAC A	CCGCC	CTTTAC	GGCGG	CGAC	ACGGG TTCGACT CCCGT CGGGGAC GCCA
AZL_rd05	Type: Glu		Anticodon: CTC (549471-549396)					
GTCCCCA TC	GTCT	AGAGCCTA	GGAC A	CCGCC	CTCTAC	GGCGG	TAAC	AGGGG TTCGACT CCCCT TGGGGAC GCCA
AZL_rd06	Type: Glu		Anticodon: CTC (549614-549539)					
GTCCCCA TC	GTCT	AGAGCCTA	GGAC A	CCGCC	CTCTAC	GGCGG	TAAC	AGGGG TTCGACT CCCCT TGGGGAC GCCA
AZL_re02	Type: Ile		Anticodon: GAT (18041-18117)					
GGGCTAG TA	GCTC	AGTTGGTTA	GAGC G	CGCGC	TTGATAA	GCGTG	AGGTC	GGAGG TTCAAAT CCTCC CTGGCCC ACCA
AZL_re03	Type: Ala		Anticodon: TGC (18147-18222)					
GGGGCGG TA	GCTC	AGTTGGGA	GAGC G	CCTGC	TTTCAA	GCAGG	AGGTC	GTCGG TTCGATC CCGTC TGCCCTCC ACCA
AZL_re06	Type: fMet		Anticodon: CAT (21600-21676)					
CGCGGGG TG	GAGC	AGCCCGGTA	GCTC G	TCAGG	CTCATAA	CCTGA	AGGCC	GCAGG TTCAAAT CCTGC CCCCAC ACCA
AZL_re07	Type: Tyr		Anticodon: GTA (357234-357319)					
GGAAGGG TG	GCCG	AGTGGTTAA	AGGC A	GCAGA	CTGTAAA	TCTGC	CCGCGTACGCGTAC	GCTGG TTCGAAT CCAGC CCCTCC ACCA
AZL_re08	Type: Val		Anticodon: GAC (449596-449522)					
GGTCTCG TA	GCTC	AGCGGGA	GAGC A	CTACG	TTGACAT	CGTAG	GGGTC	ACTGG TTCAACT CCAGT CCGGGAC ACCA
AZL_re09	Type: Gly		Anticodon: GCC (449699-449625)					
GCGGGCG TA	GCTC	AGGGGTA	GAGC A	CAACC	TTGCCAA	GTTTG	GGGTC	GAGGG TTCGAAT CCCTT GCGCCG TCCA
AZL_re10	Type: Asp		Anticodon: GTC (449780-449704)					
GCGGGTG TA	GCTC	AGTTGGTTA	GAGC G	CCGCG	CTGTAC	GCCGG	AGGCC	GCGGG TTCAGT CCGCT CACTCGC GCCA
AZL_re11	Type: Ala		Anticodon: GGC (449929-449854)					
GGGGCCA TA	GCTC	AGTTGGGA	GAGC G	CTACA	ATGGCAT	TGTAG	AGGTC	AGGAG TTCGATC CCTCT TGCCCTCC ACCA
AZL_re12	Type: Leu		Anticodon: CAG (450054-449968)					
GCCTGTG TG	GCGG	AATTGGTAG	ACGC G	CTAGG	TTCAGT	CCTAG	TGGCTGAAAGCCGT	GGGGG TTCGATC CCTCT CAAGCGC ACCA

Supplementary Figure 6.

variation1 (*rrn2, rrn3, rrn6, rrn9*)



variation2 (*rrn1, rrn4, rrn7, rrn8*)



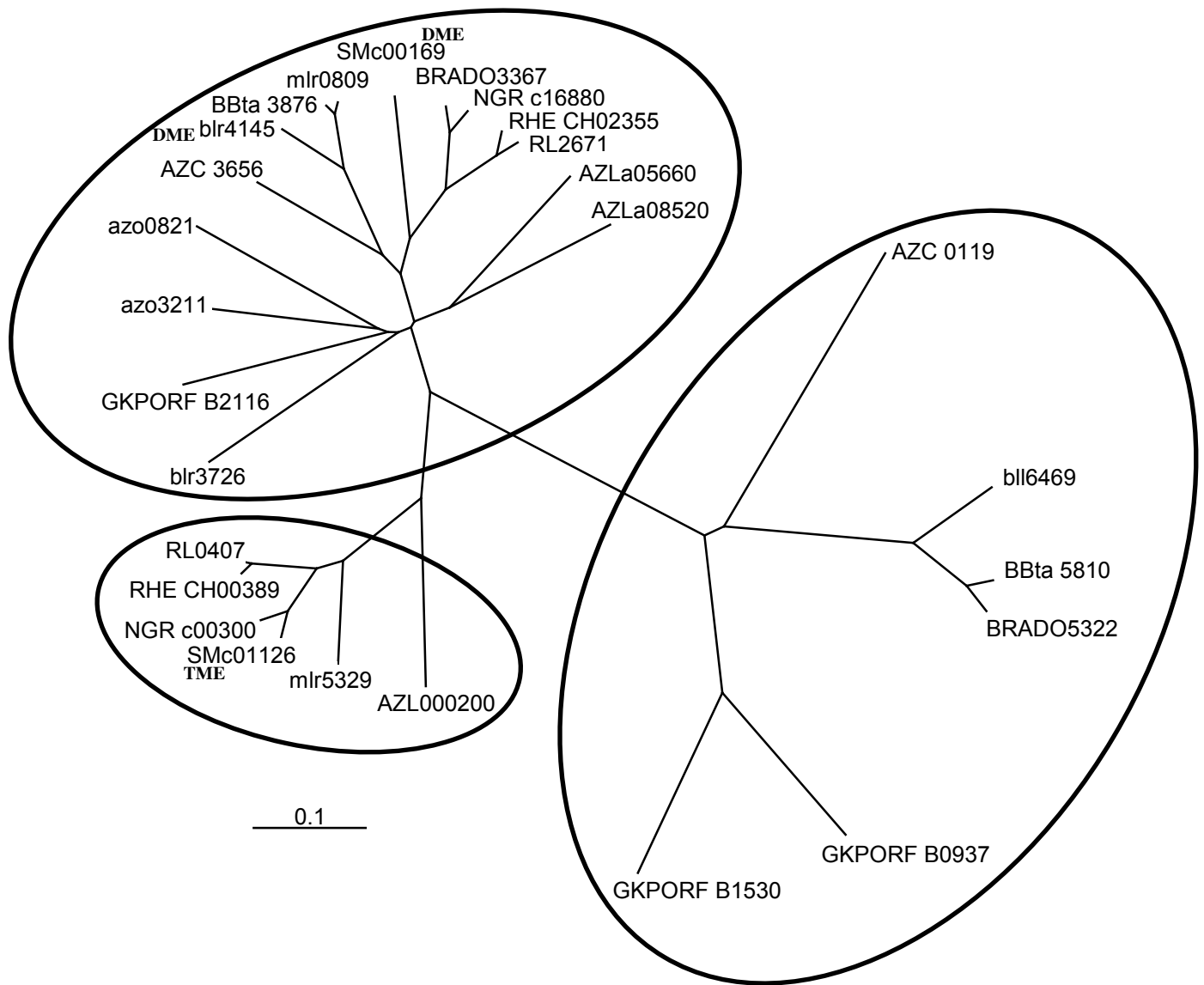
variation3 (*rrn5*)



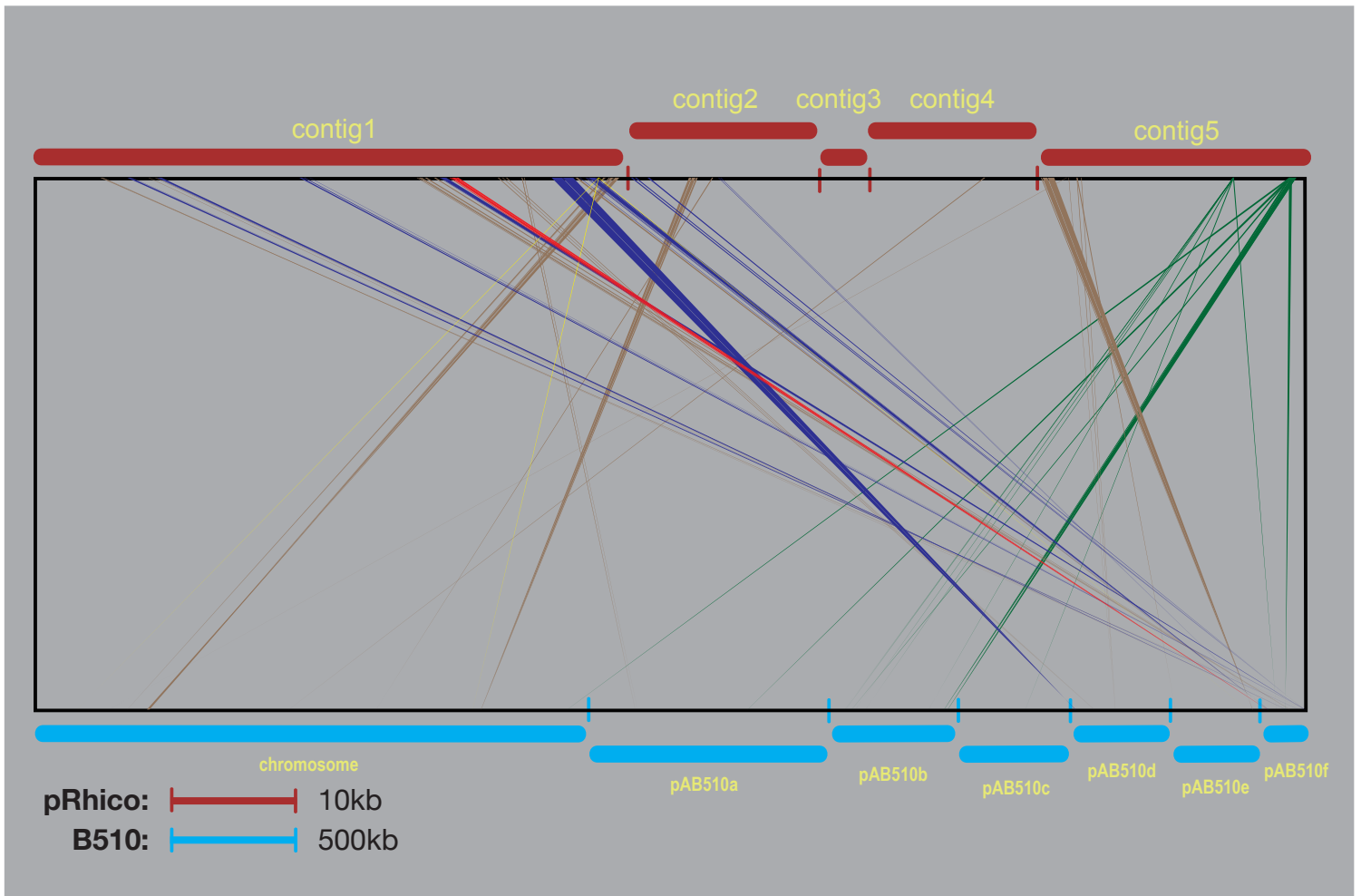
5' (471791)

TCAGAAATCCAGCATCGCGTTGTCGATCACCTCCTTCATGACGAAGAAGGTGCGGTCTGCCGCACGCCGGAAGCGCGATCAGTTGCTGCCCGTGCAG
ACGGTTGAAATCCGCGATGTCGCTGACCCGGATCTTCAGGAAGAAATCGAAGTCCCCGGCGACGAGATGACAGTCCAGCACGAAGGGCAGCTTGGGAT
CGCCTGCTCGAACCGCGCGAAGCTTCCGGCGTCGAGCGGTCCAACCTCGACCCCGACCATACCAGCGTGCCCGATCCACCTTCGCCGGCTCGATCTC
GGCCCGGACGGAACGGATGTAGCCCTCGGCGAAAAGCCGTTGGGTGTGCCGGTGGCAGGTGGCCGGGCTGAGATTACCCGCCGCCAGATCGGCGTT
CGTCATCCGGCCATCCCTTTGAAGAAGCTGGAGAATCTTGGGTGCGCCGGGCTGGGCCCTCATCGGCGGTGGTTTCTCATTTTCATCCATGAATT
TGGGATGGATAGCGCACAAAATGGCATTTATCAGCCTGAAACAGCAATCCGATAGCGAAGTTTGAGAGCACCTTTCATTTCTACGGTGATACCGTCCG
TTCACTTGAGACAGTGGAGCCCCGCCATGCTGCGCCTCGATCGCTTCGAACGTTATCCGCTCACCTTCGGCCCGACTCCCATCGAACATCTGCCCCGCC
TGACCGAGGCGCTCGGCGGCAAGGTCGAGATCTACGCCAAGCGCGACGACTGCAACTCCGGCCTCGCCATGGGCGGCAACAAGCTGCGCAAGCTGGAAT
ACATCGTTCCCGACGCCATCGCGGCCGGCGCCGACACGCTGGTGTCCATCGGCGGCGTGCAAGTCAACACACCCGCATGGTGGCGGCCACCGCCGCA
AGCTCGGCATGAAGTGCCTGGTGCAGGAGAGCTGGGTTCCGCACGAGGACGCCGTCTATGACCGCGTCGGCAACATCCTGCTGACCCGCTGATGG
GCGCCGACAGCCGCATCGTTTCCGACGGCTTCGACATCGGCATCCGCAAGAGCTGGGAGGACGCCATCCAGTCGGTGAAGGATGTCGGCGGCAAGCCCT
ACGGCATTCCGGCCGGCGGCTCCGTCCACAAATACGGCGGTCTCGGCTATGTCGGCTTCGCCGAGGAGGTCCGCAAGCAGGAGGCCGAGCTCGGCTTCA
AGTTCGACTACATCGTCTGCTGCTGACCGGCTCGACCCAGGGCGGCATGATCGTTCGGCTTCGCCGCCGACGACCGCGCCGACCGCGTGATCGGCA
TCGACGCCTCCGGCACGCCGGAACAGACCCGCGCCAGGTCCGCCAGATCGTCGACAACACGGCCGAGCTGGTTCGAACTCGGCCGCAAGGTCCGCGACG
ACGAGATCGTCATCCTGGAAGACTATGCCTATCCGGCCTATGGCGTGCCAGCGCCGAGACCAACGAGGCGATCCGCCTGGCCGCCCGCACCAGGCGCA
TGATCACCGACCCGGTCTATGAGGGCAAGTCGATGCAGGGCATGATCGACCTCGTGAAGAAGGGCTGGTTCCCCGAAGGCTCGAAGGTCCTCTACGCC
ATCTCGGCGGCGCCCCGGCGCTGAACGGCTACAGCTACACCTACCGGAACGGCTGA-3' (470631)

Supplementary Figure 8



Supplementary Figure 9



Supplementary Figure 10