

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.

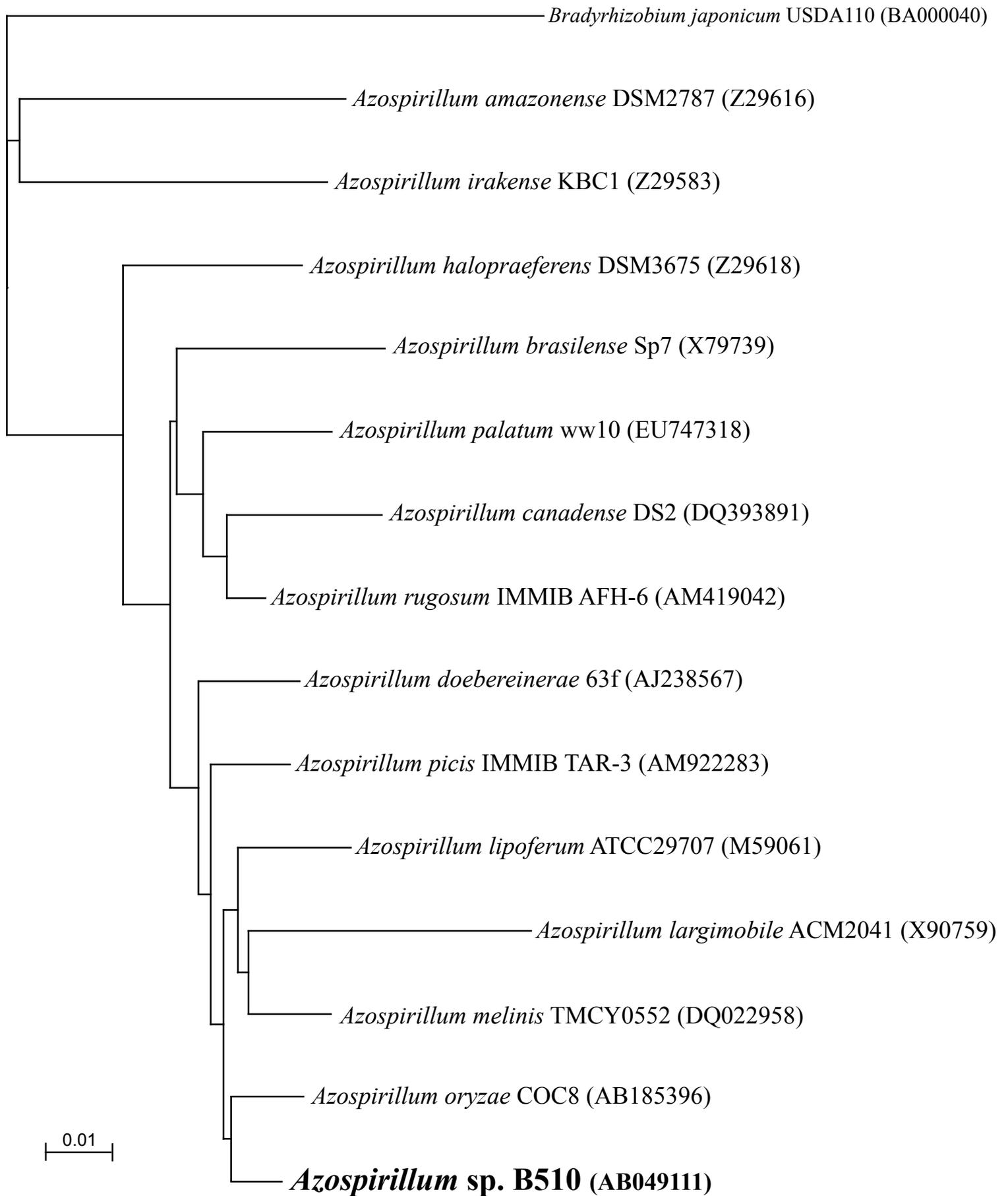
Supplementary Figure 8. The nucleotide sequences of *acdS* and *acdR* in *Azospirillum* sp. B510. The *acdS* and *acdR* genes are indicated with red and blue arrows, respectively, and the direction of the arrows denotes the direction of transcription. The box indicates a putative LRP protein-binding 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.

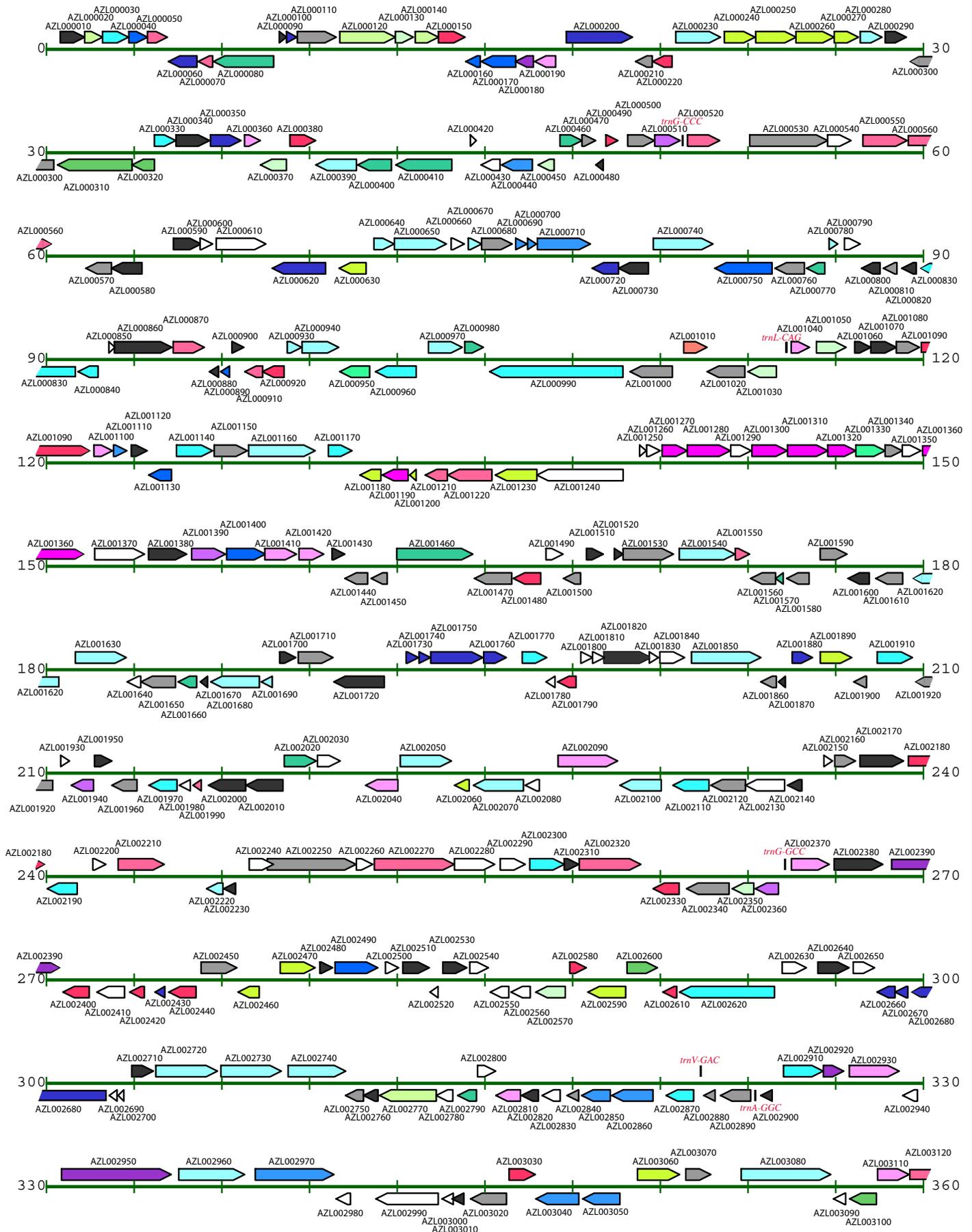
BTai1; BRADO3367, BRADO5322: *Bradyrhizobium* sp. ORS278; GKPORF_B0937, GKPORF_B1530, GKPORF_B2116: *Klebsiella pneumoniae* Kp342; NGR_c00300, NGR_c16880: *Rhizobium* sp. NGR234; RHE_CH00389, RHE_CH02355: *Rhizobium etli* CFN42; RL0407, RL2671: *Rhizobium leguminosarum* bv. *viciae* 3841; SMc00169, SMc01126: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

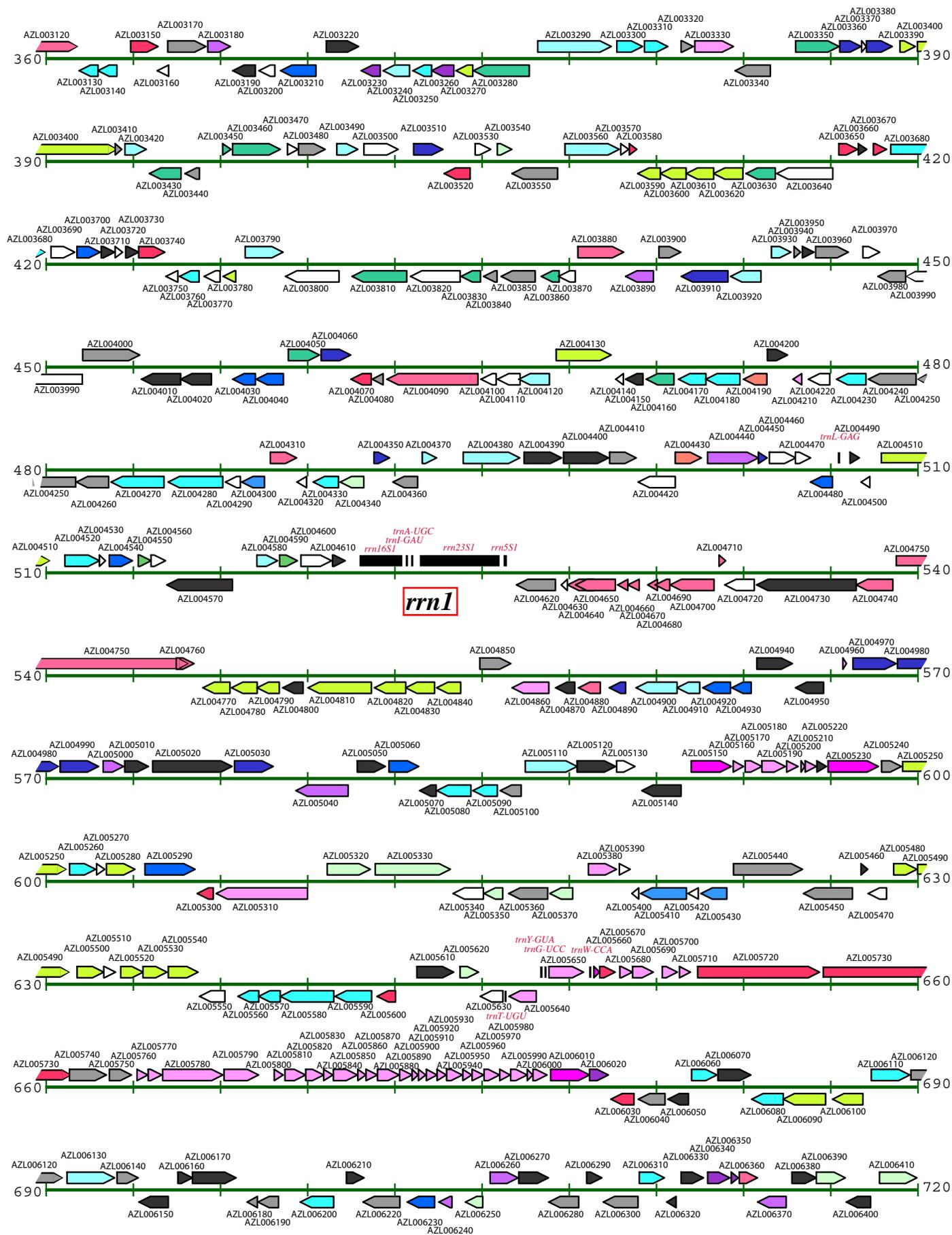
Supplementary Figure 10. Comparison of the pRhico-plasmid and B510-genome sequences. The pRhico plasmid is presented at the top of the figure, with contigs 1 to 5 arranged in numerical order. The B510 genome is shown at the bottom of the figures, with genomic elements arranged in the following order: 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.

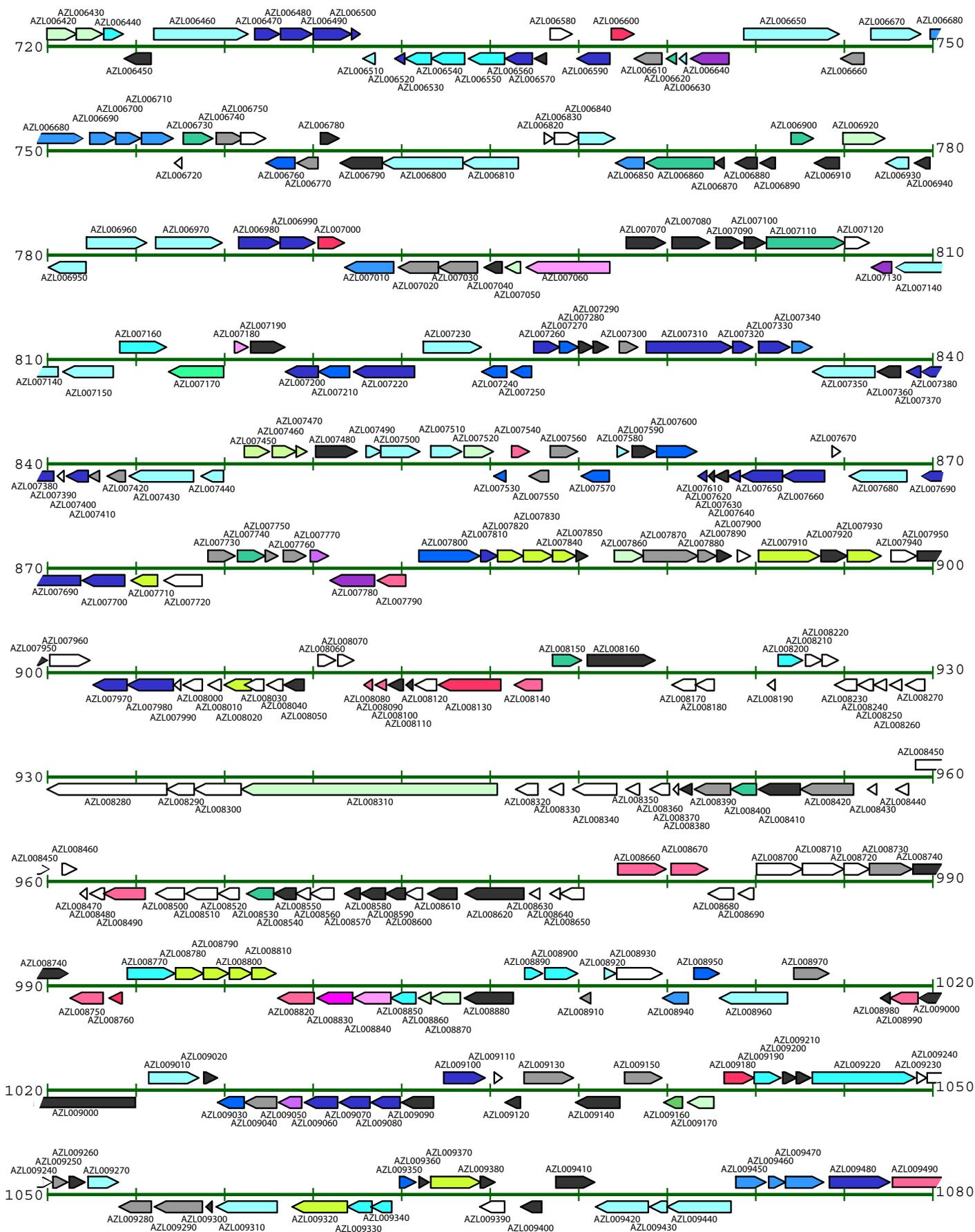


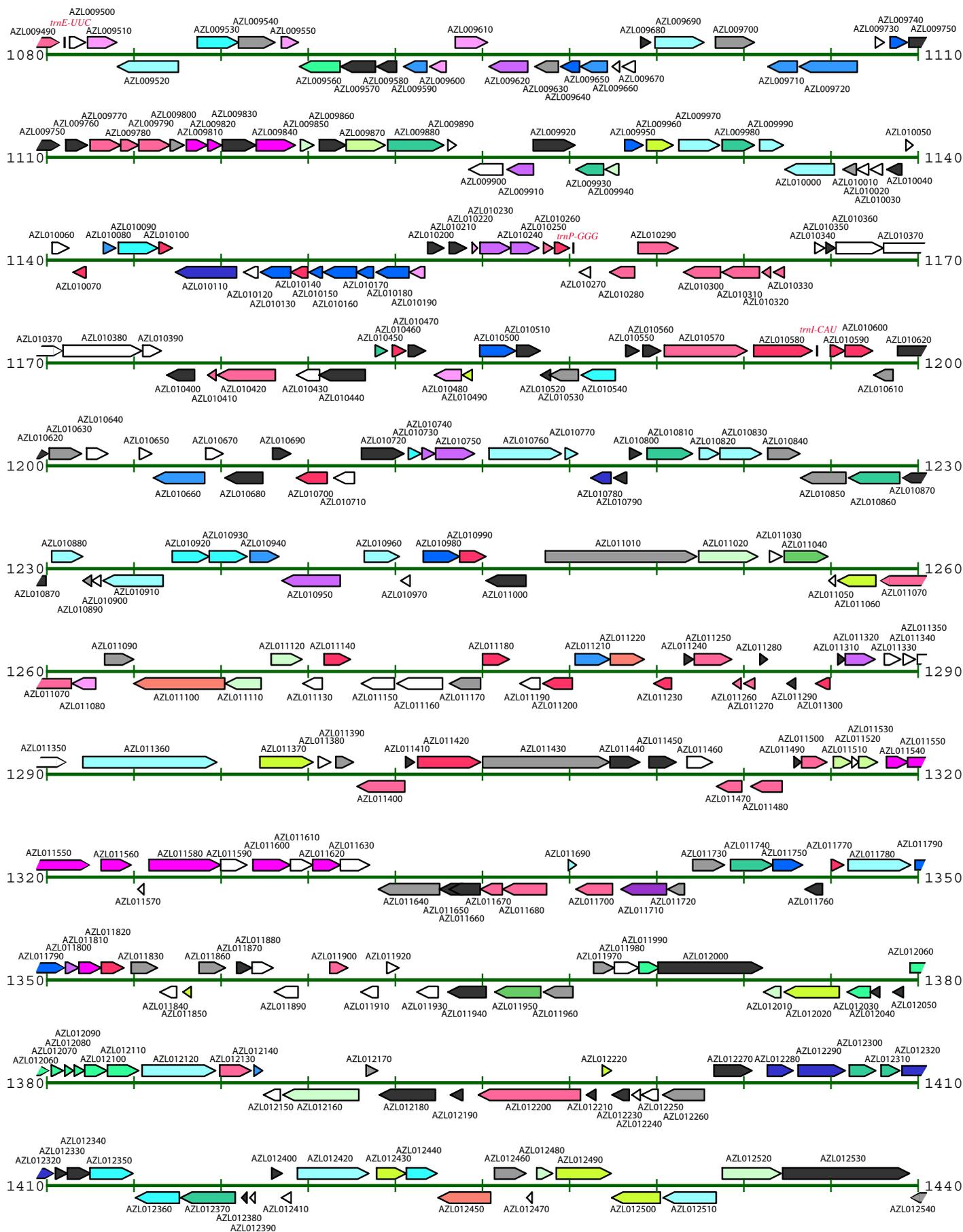
Supplementary Figure 1

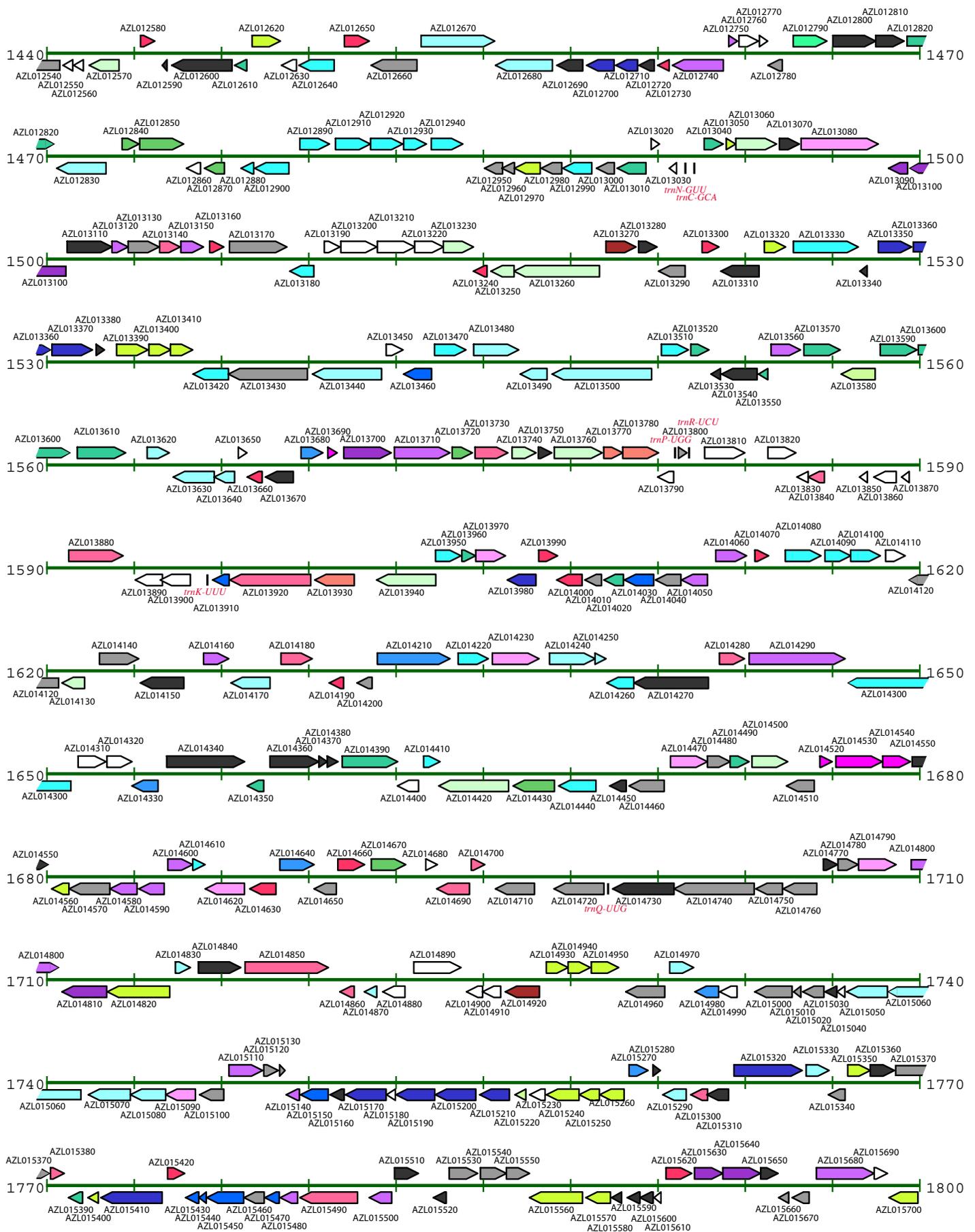
chromosome

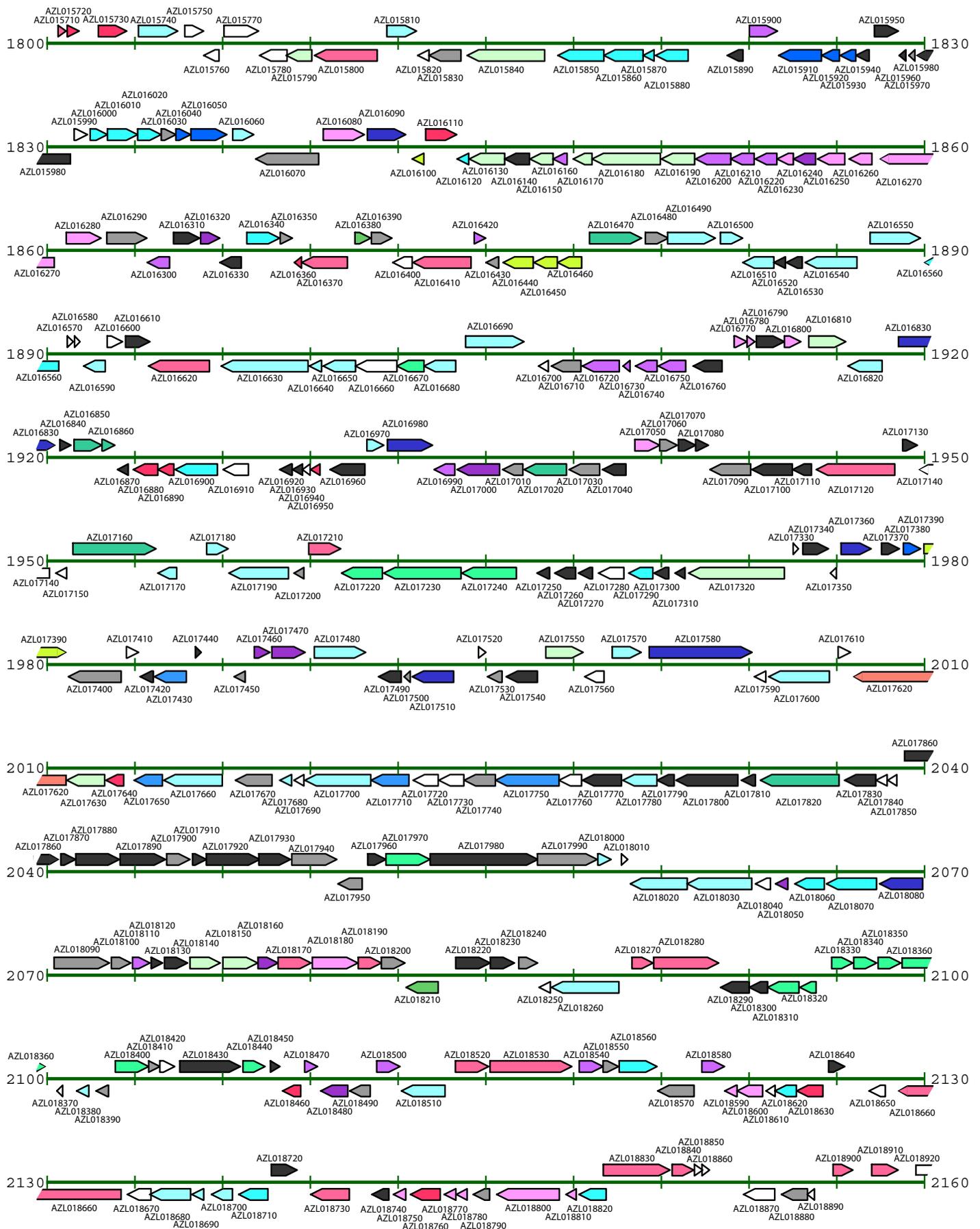


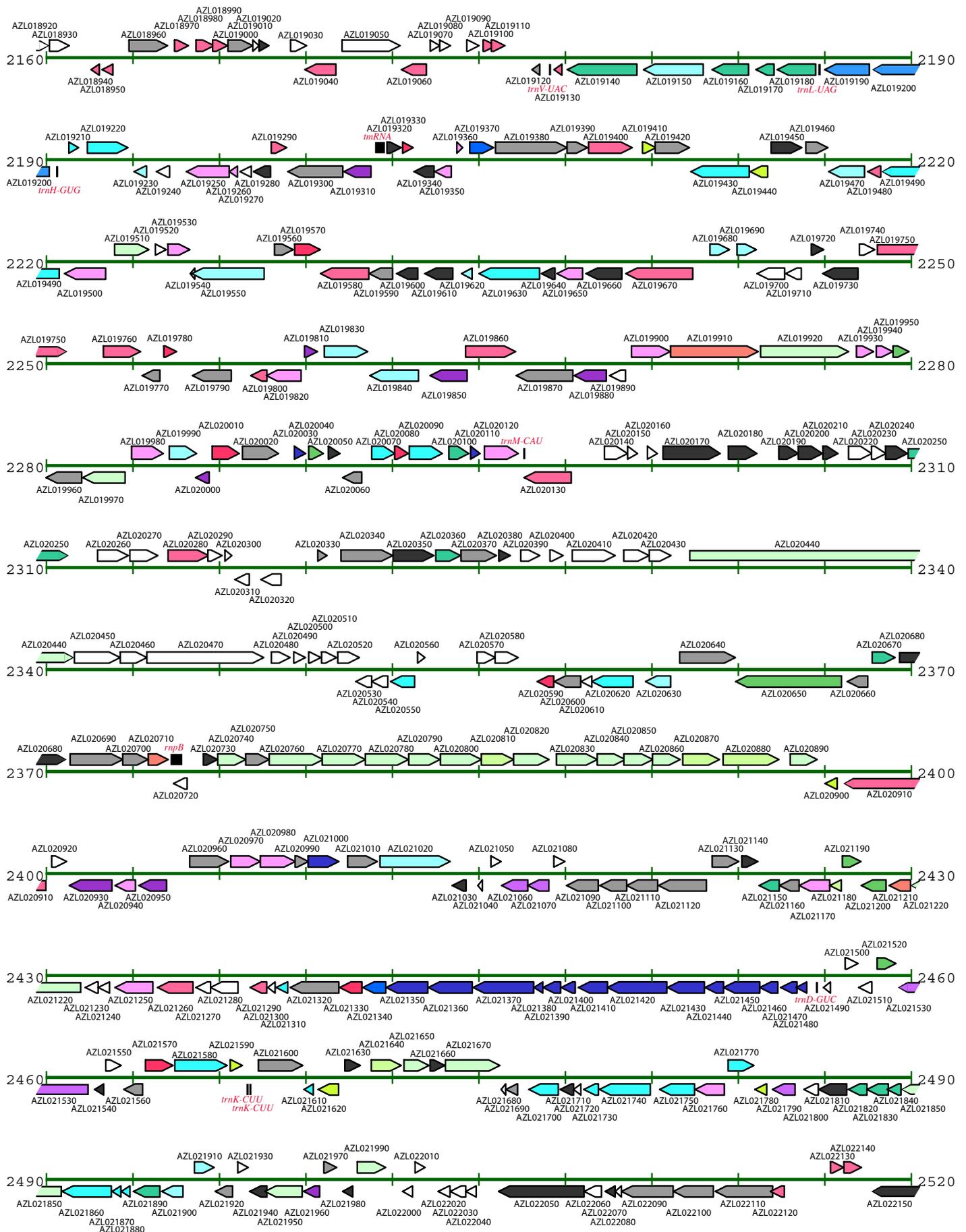


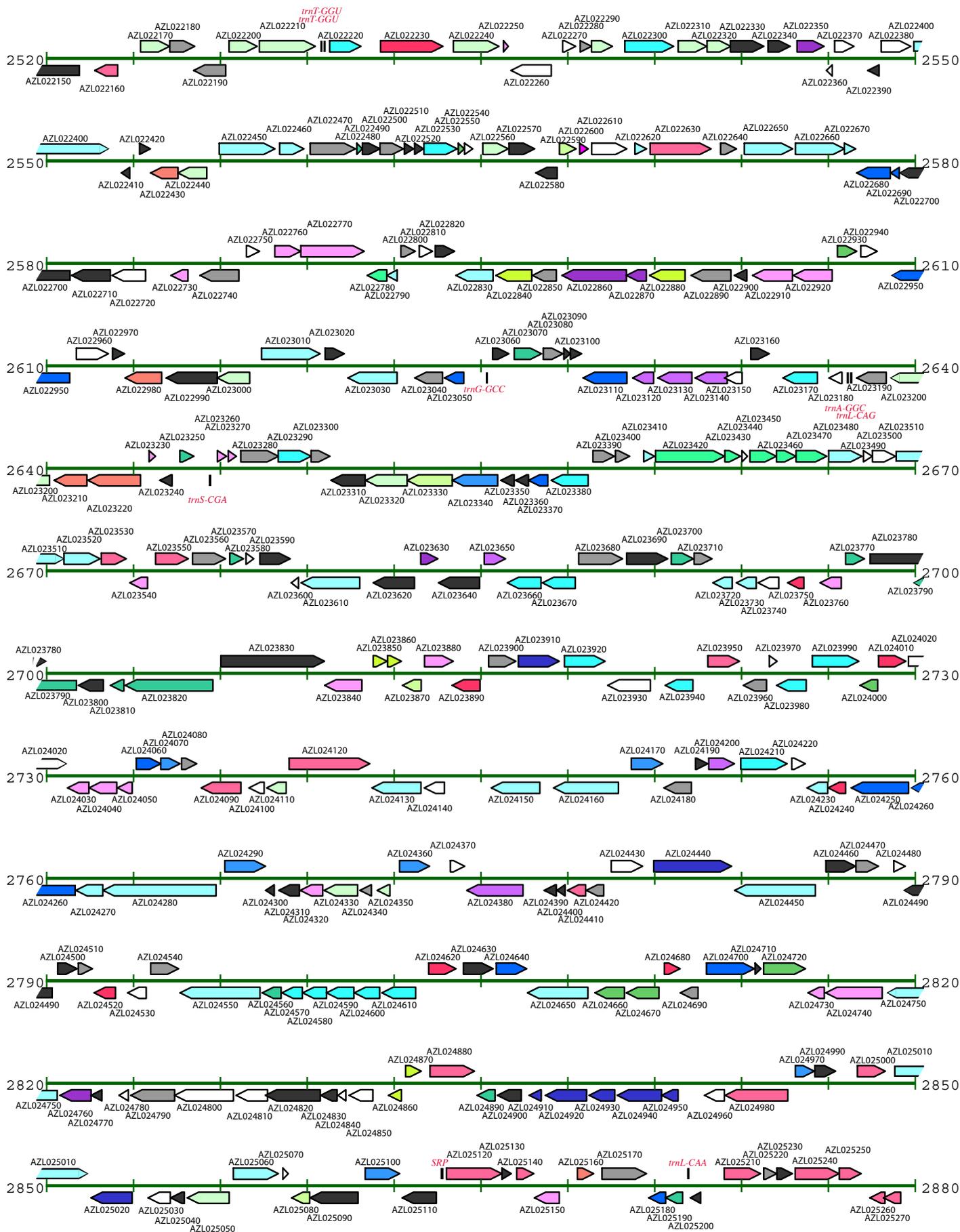


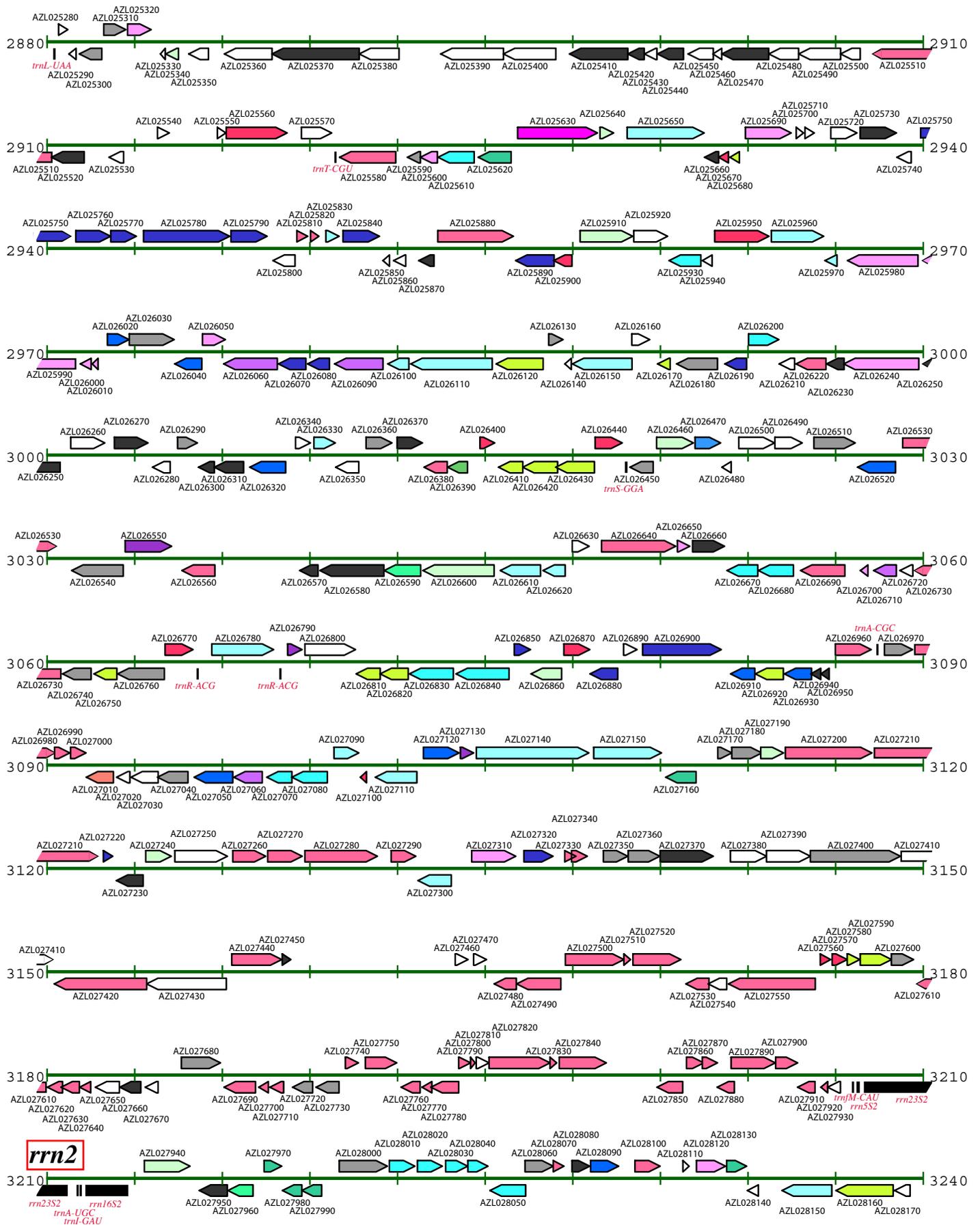


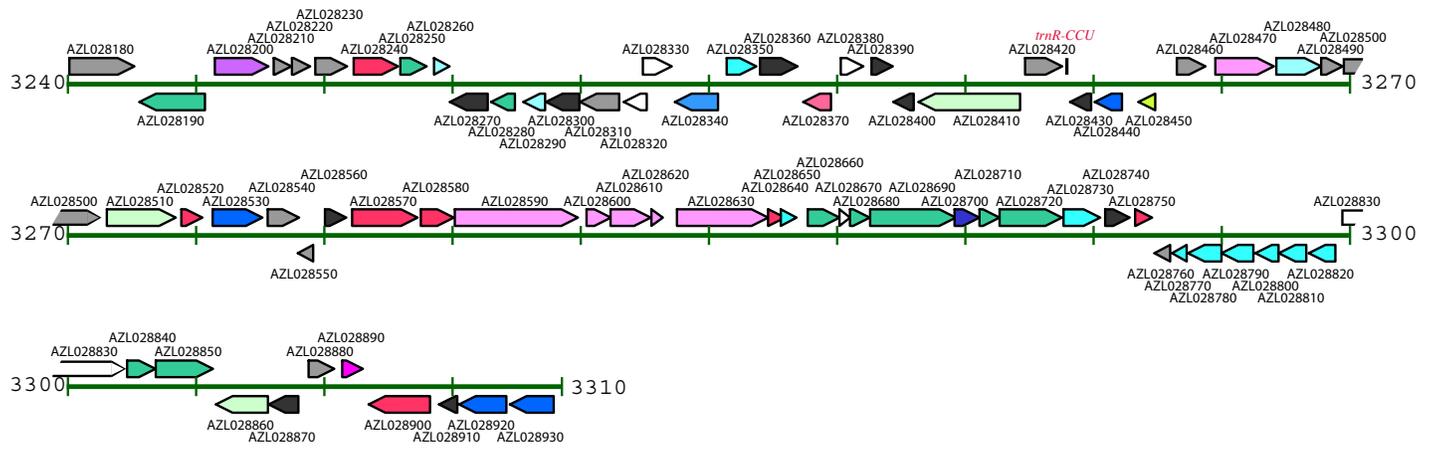




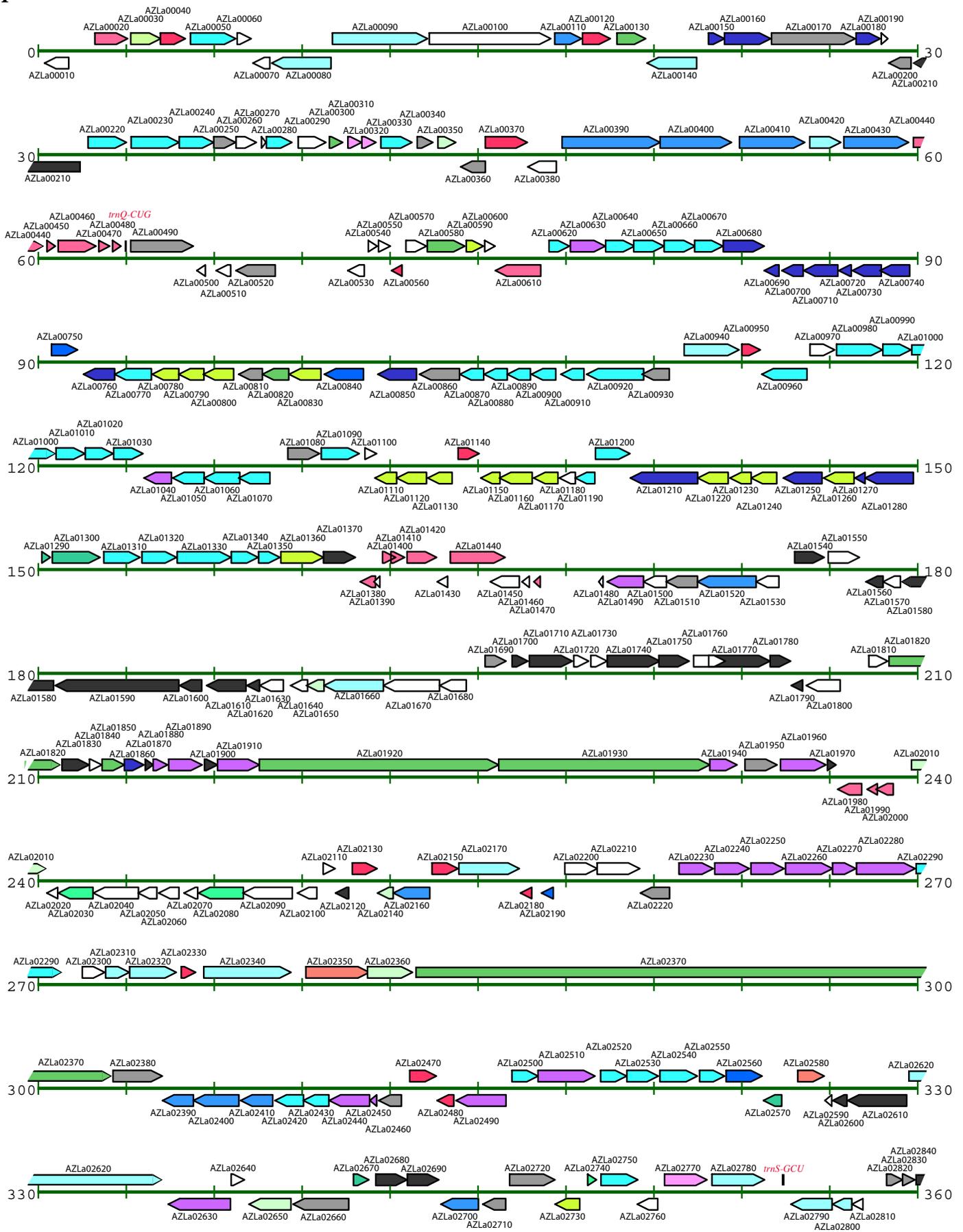


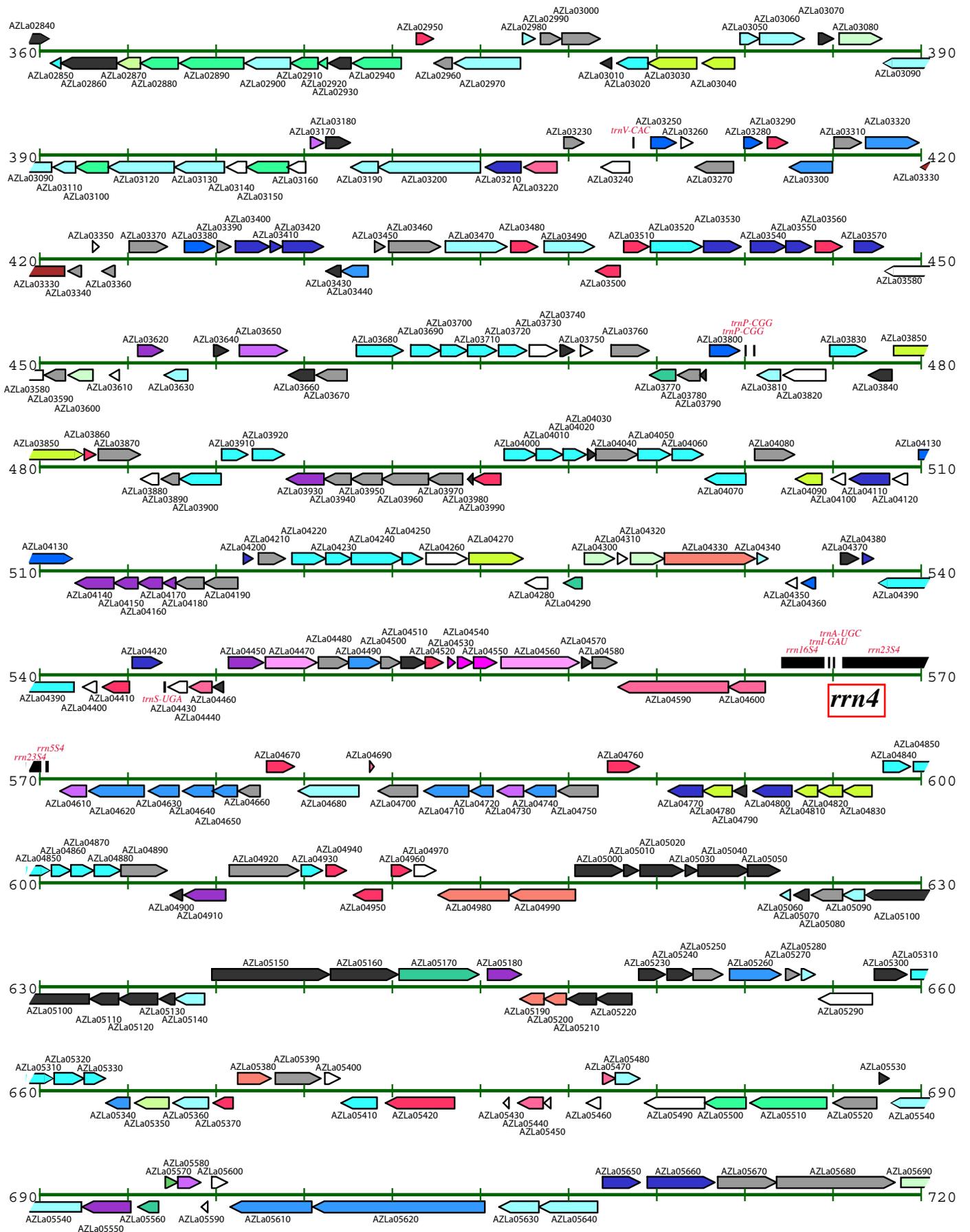


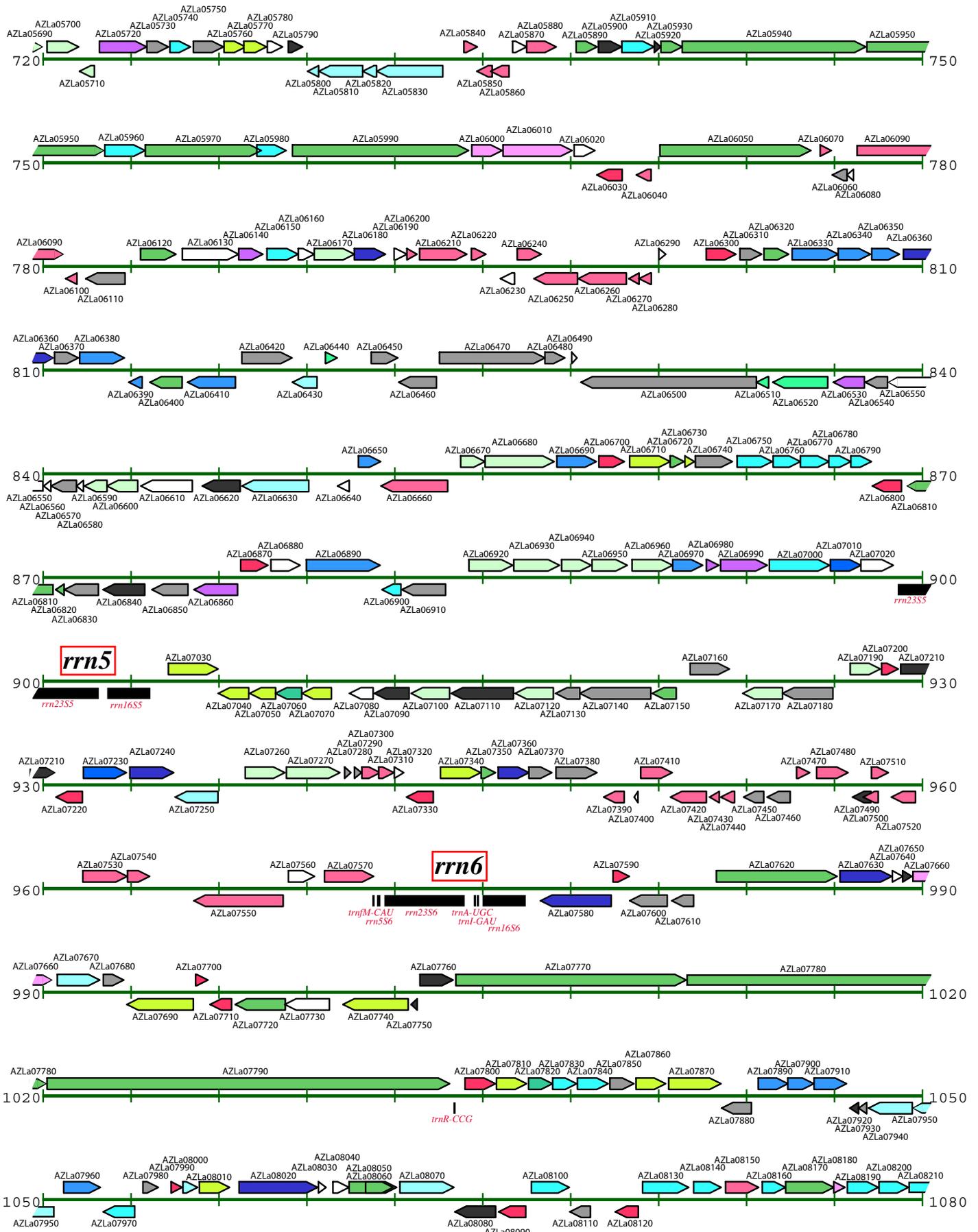


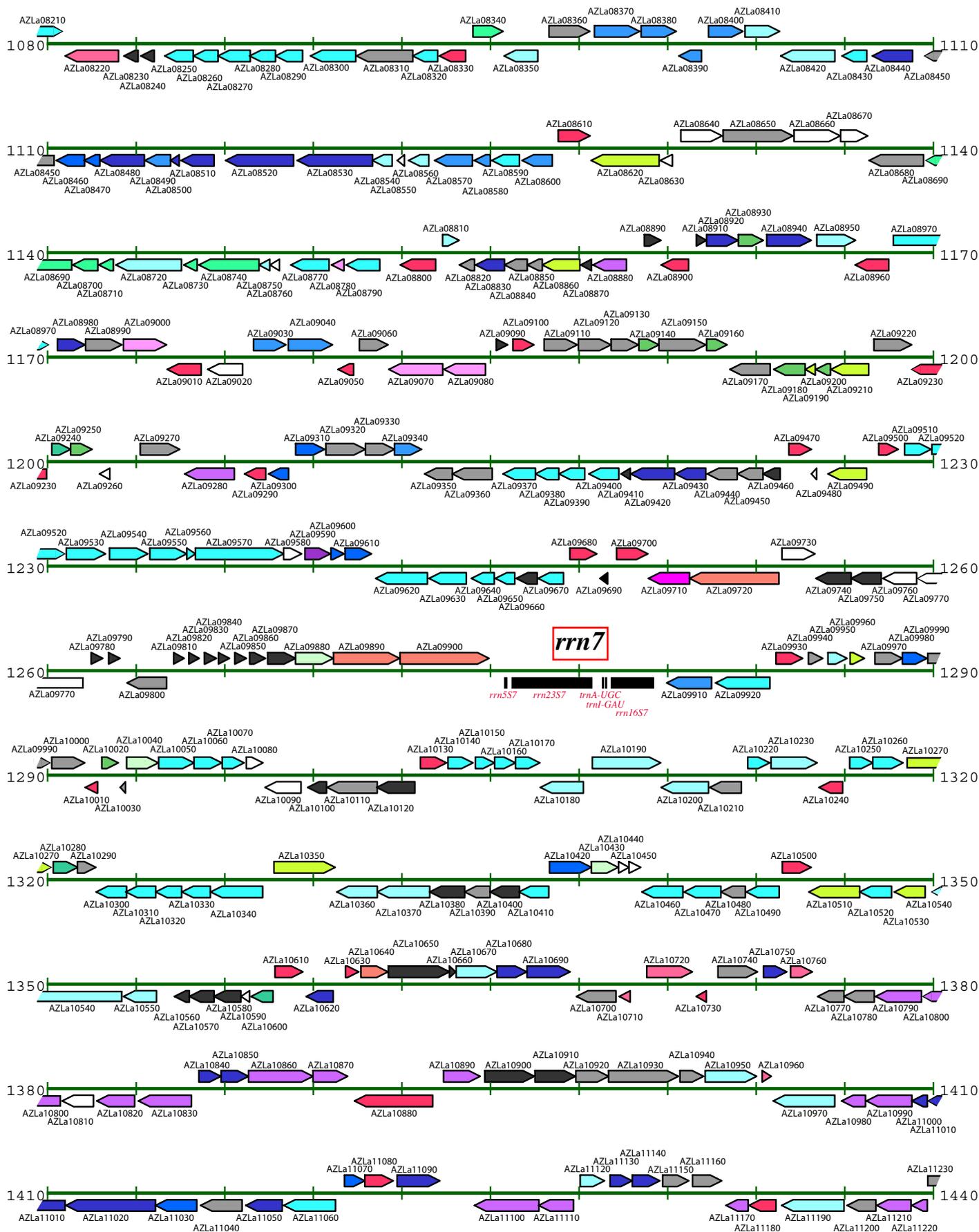


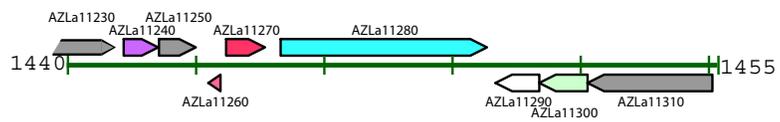
pAB510a



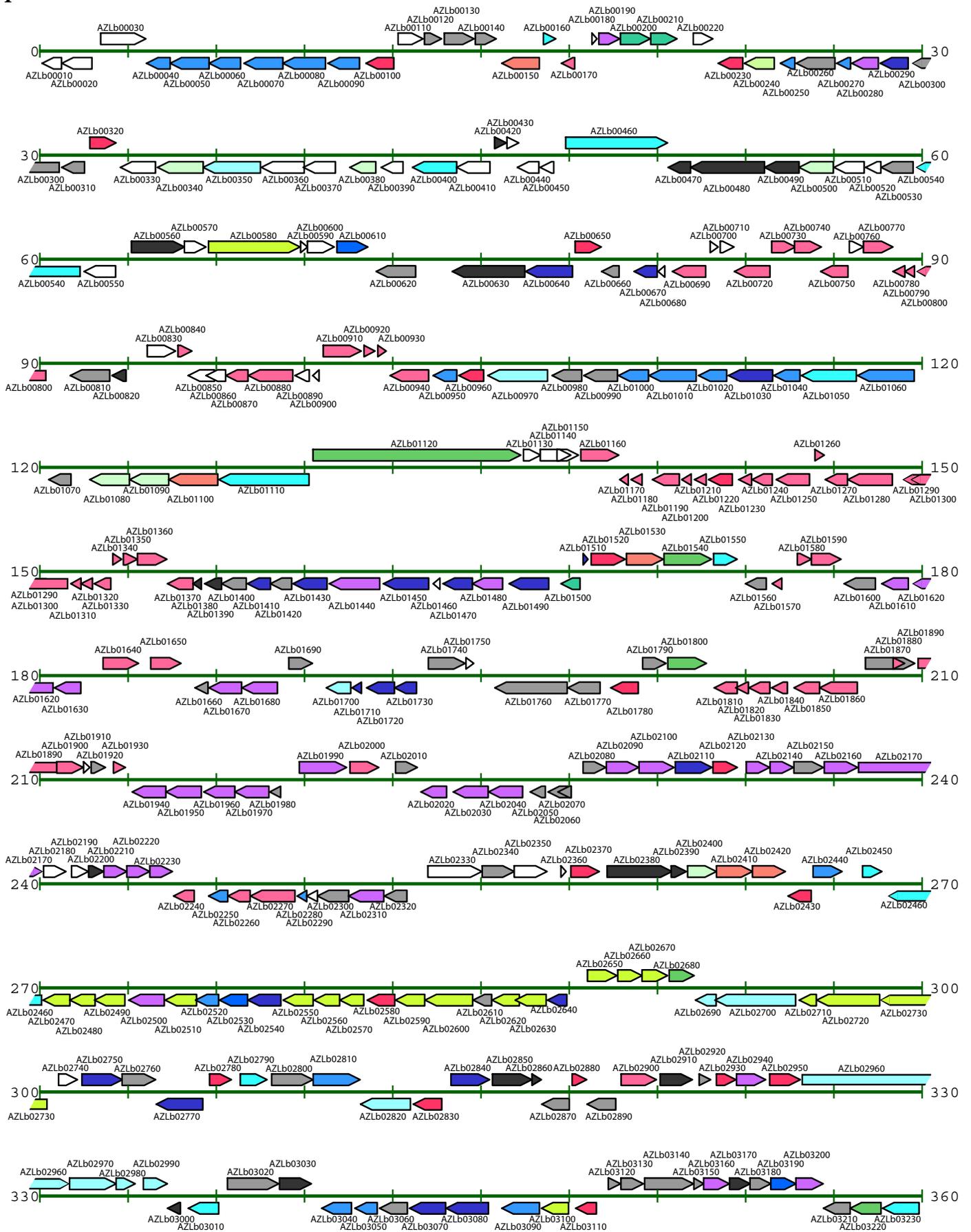


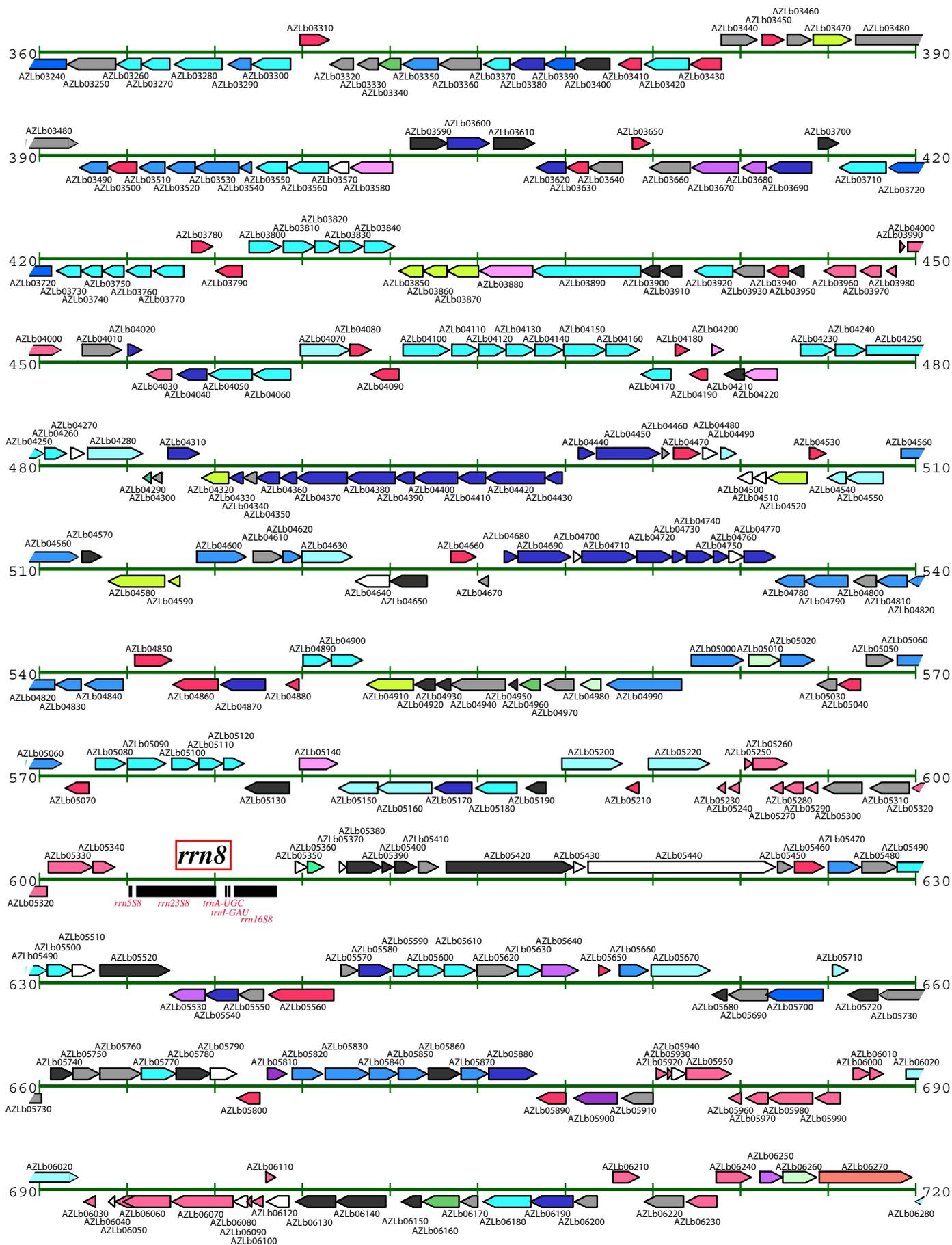


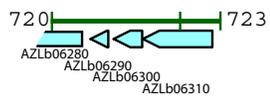




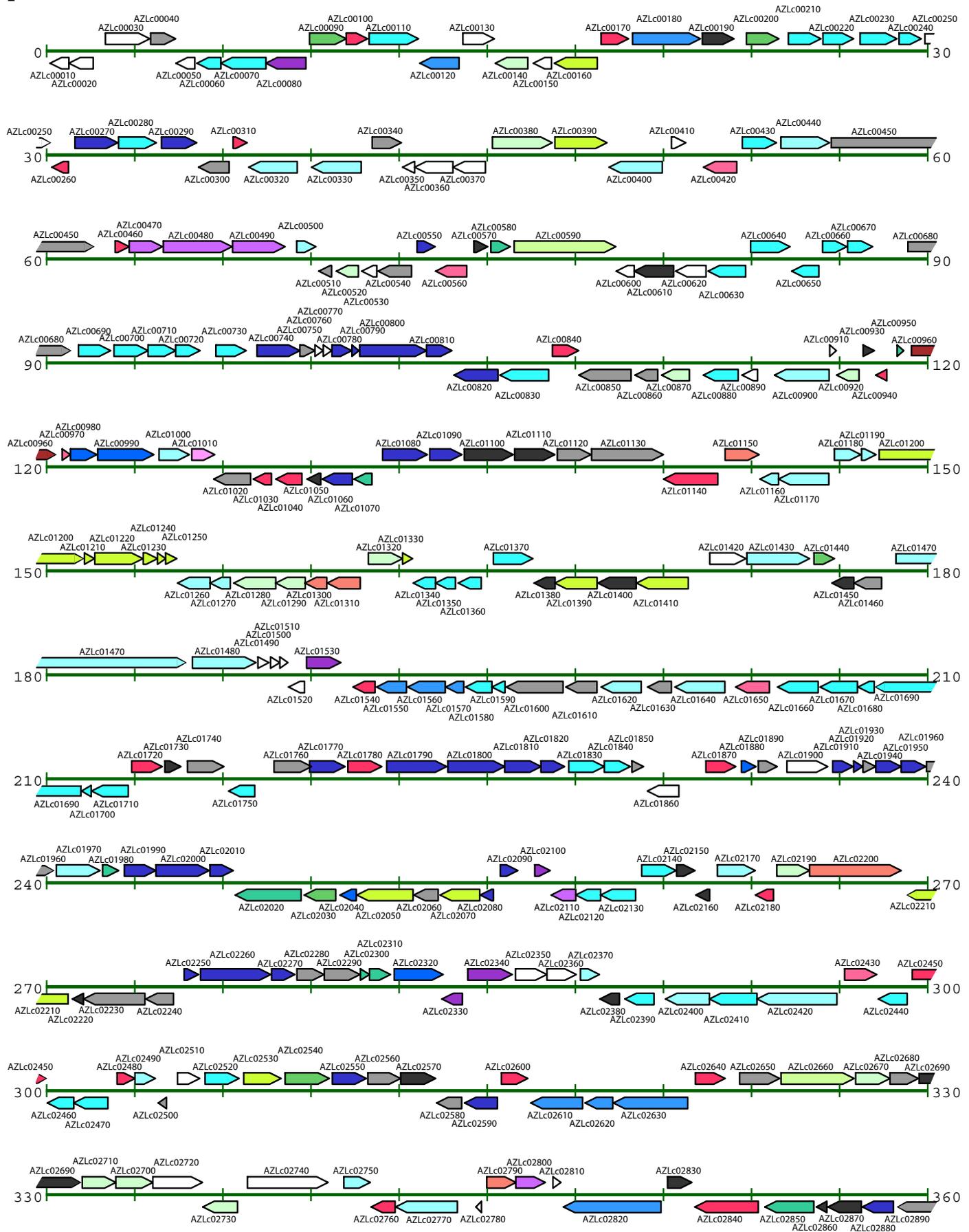
pAB510b

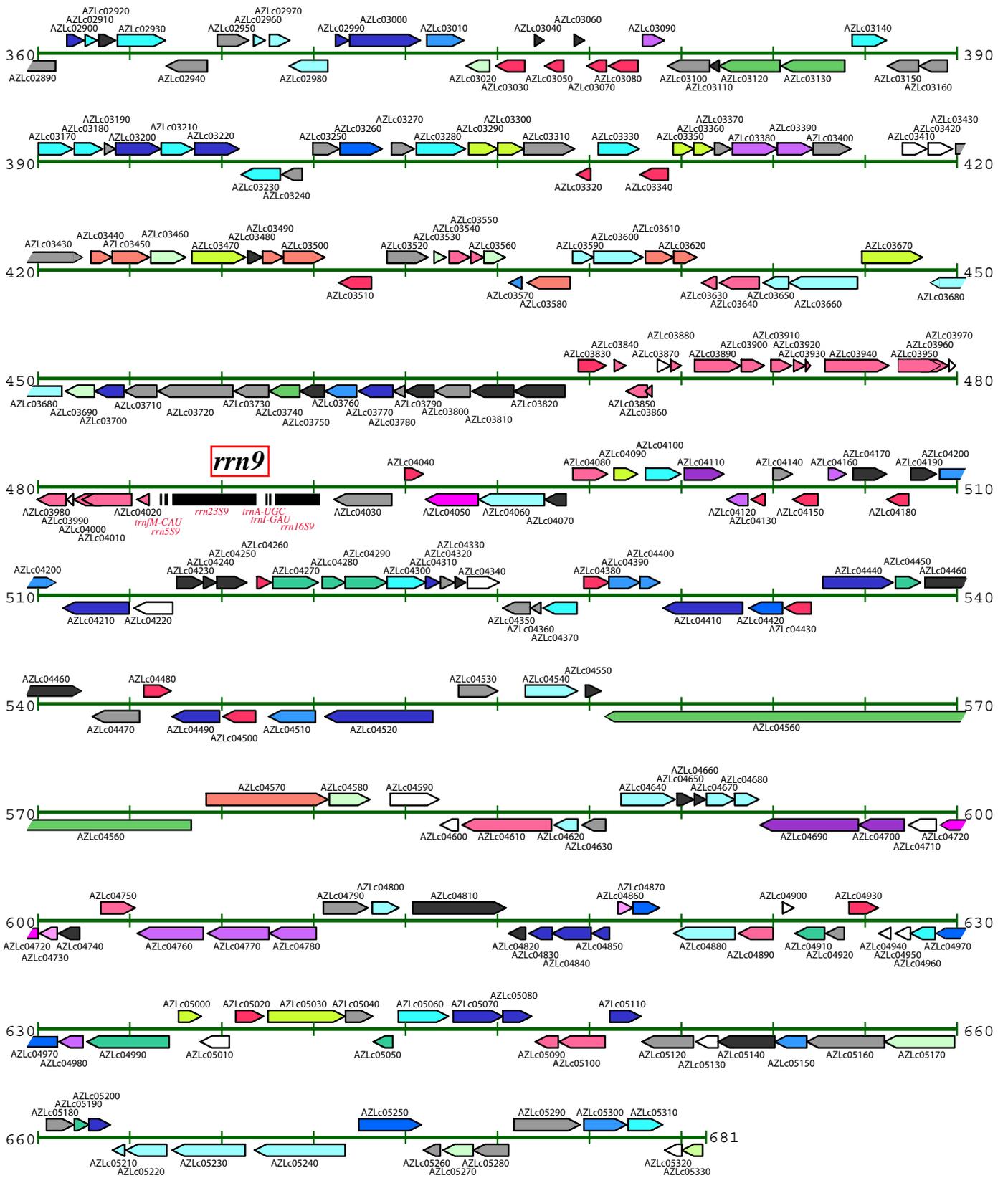




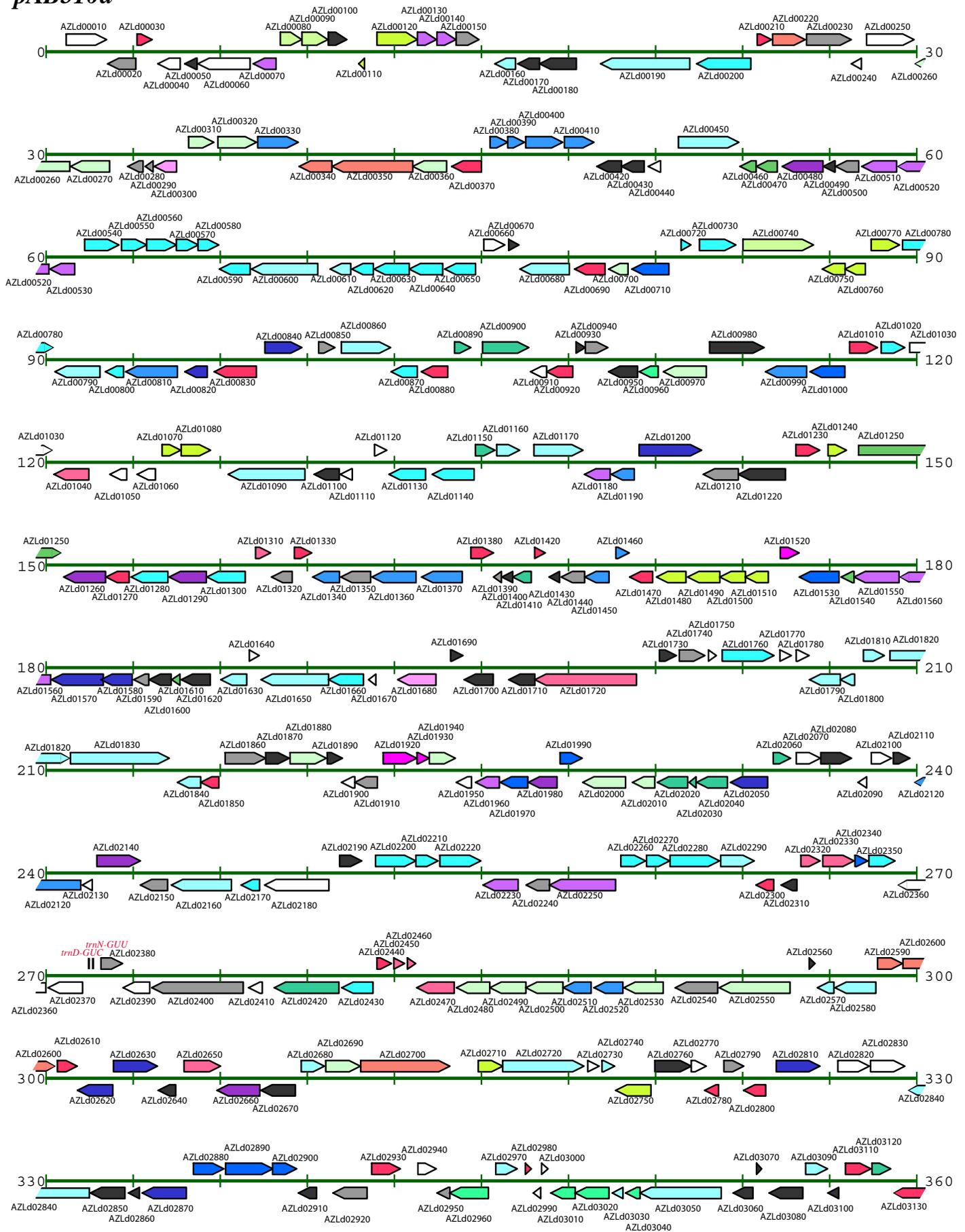


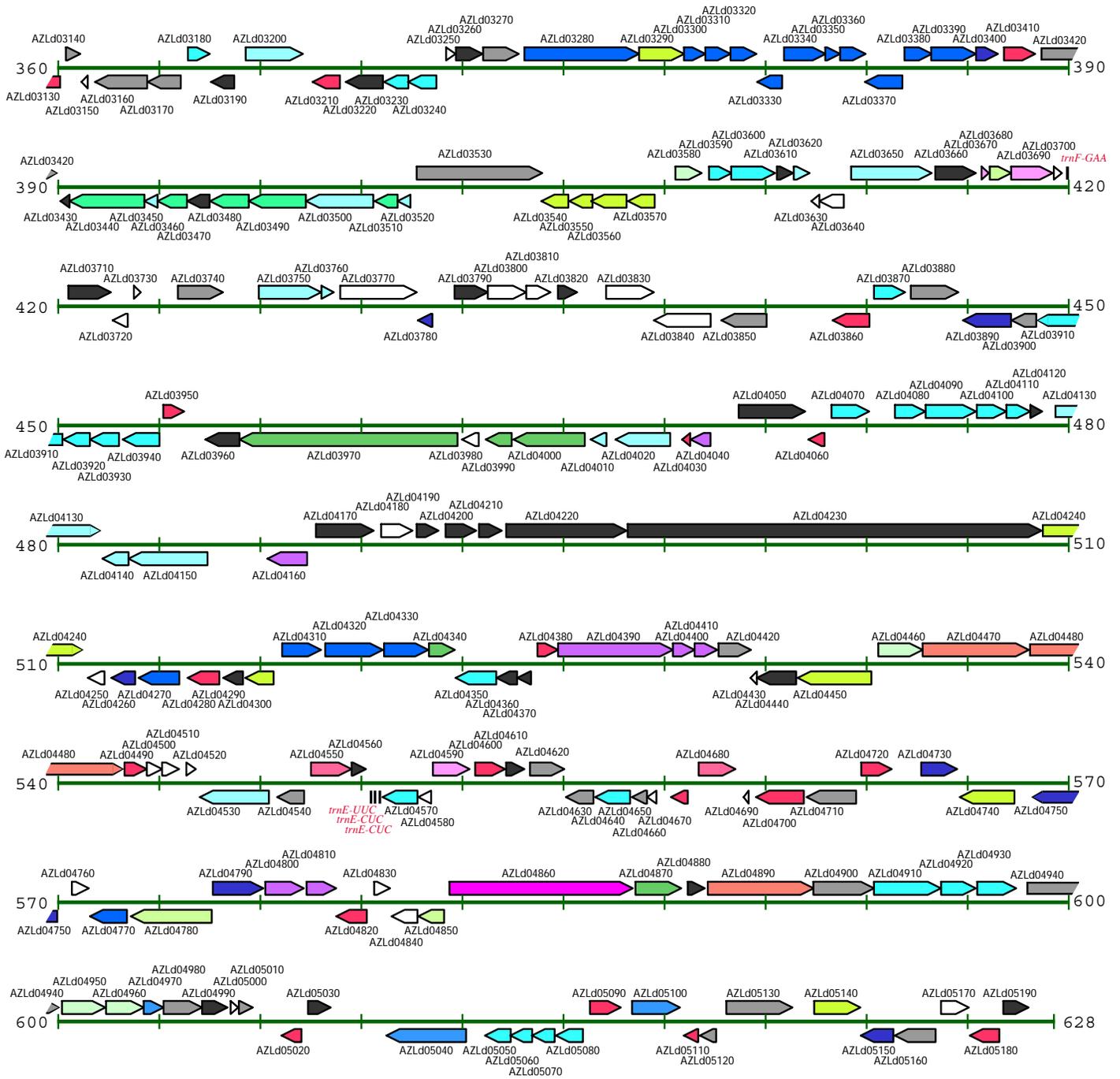
pAB510c



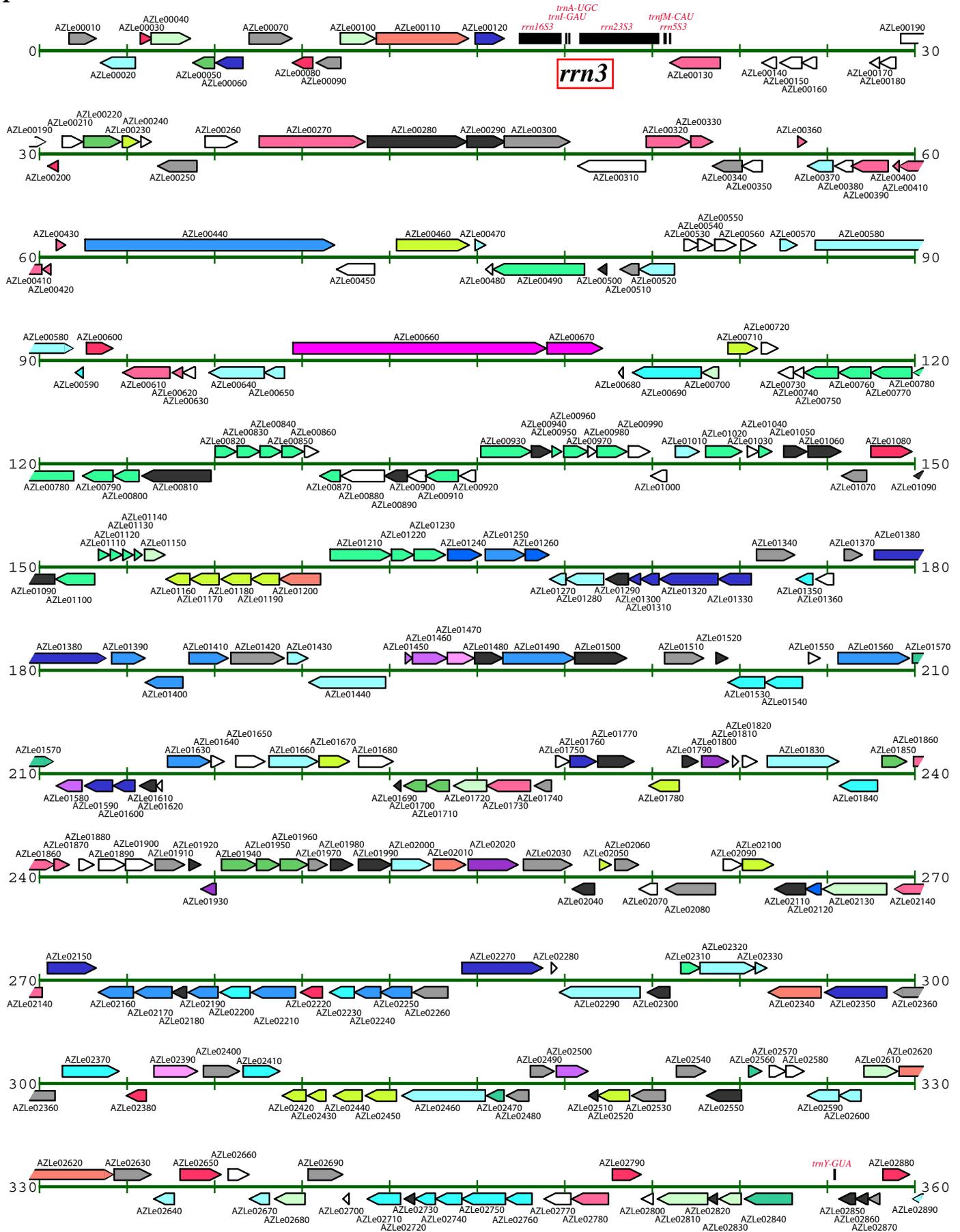


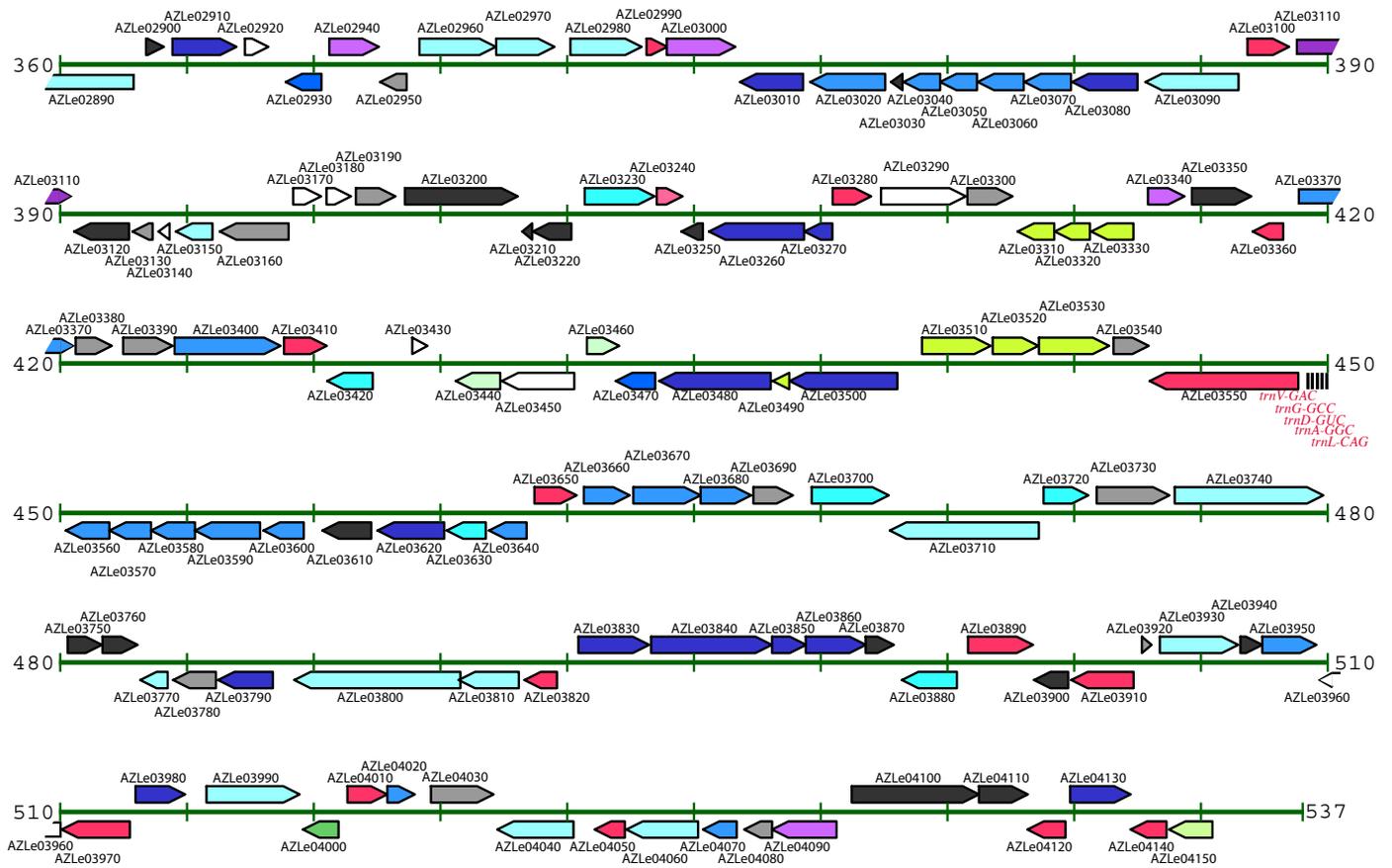
pAB510d



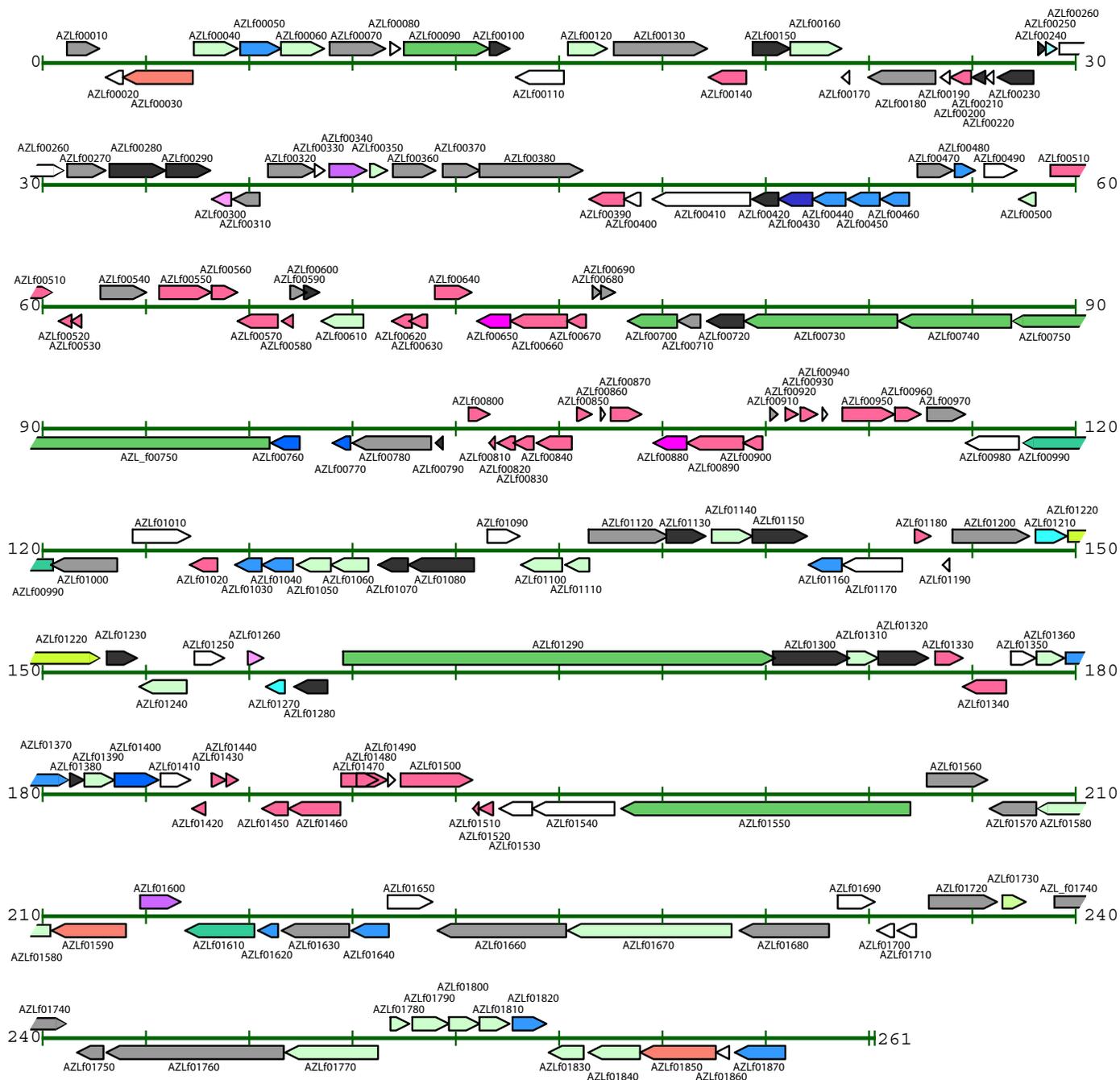


pAB510e

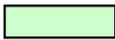


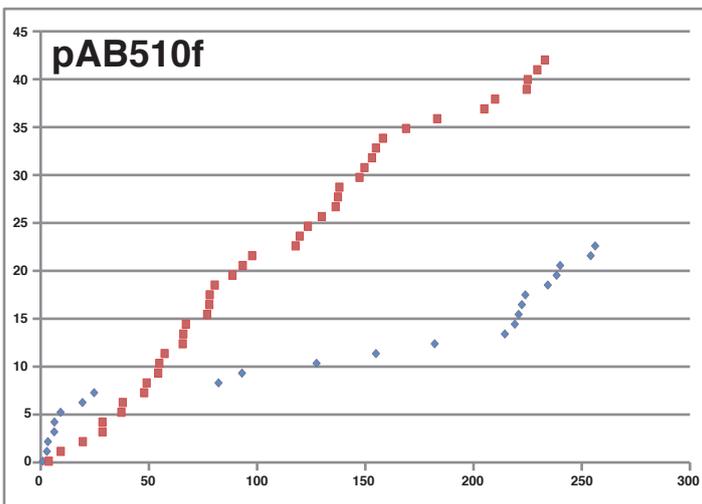
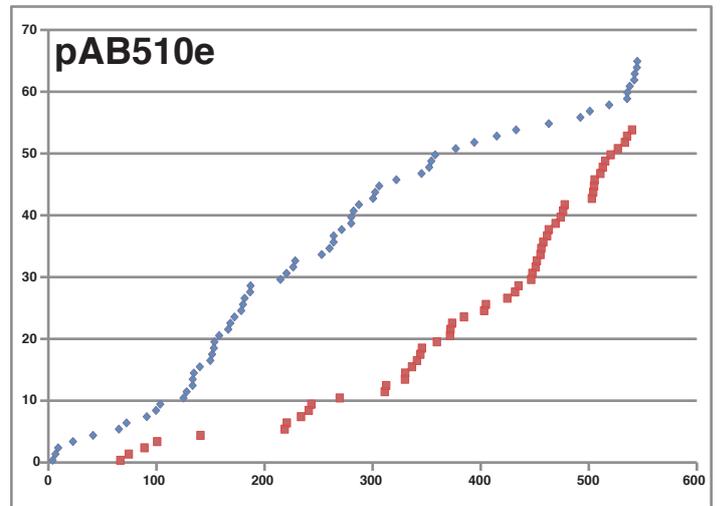
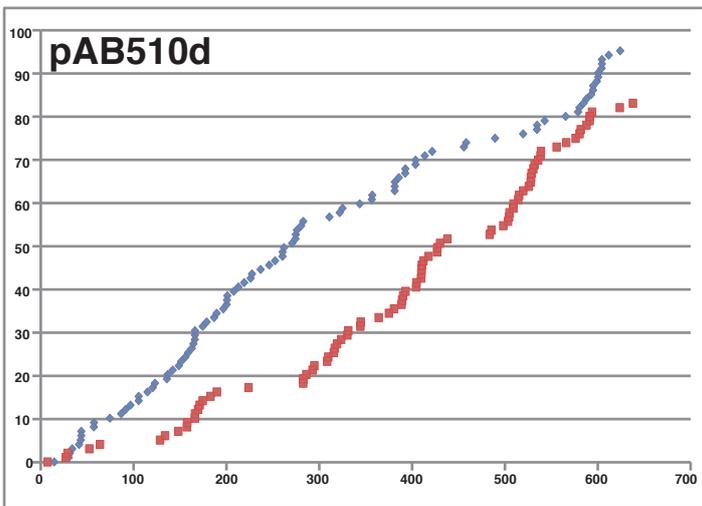
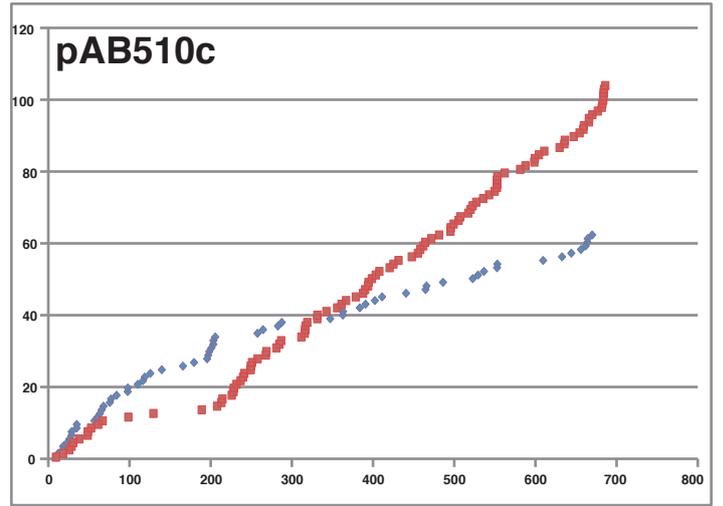
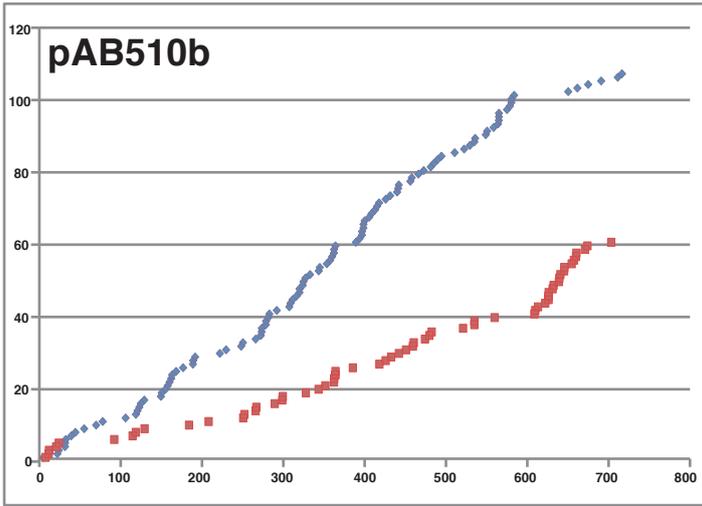
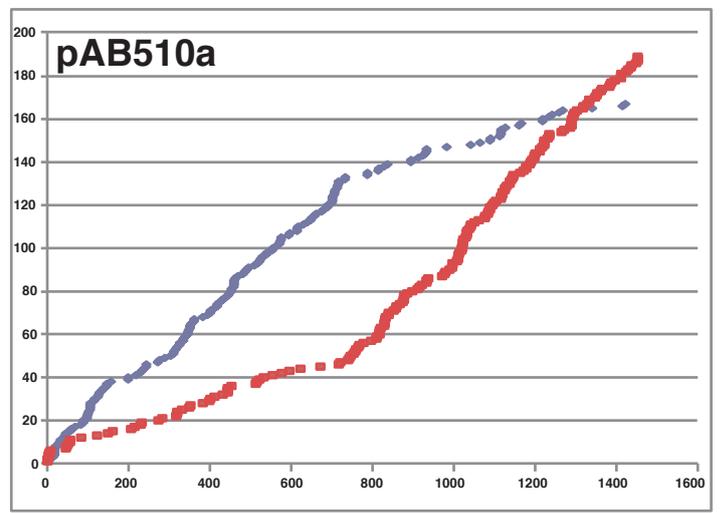
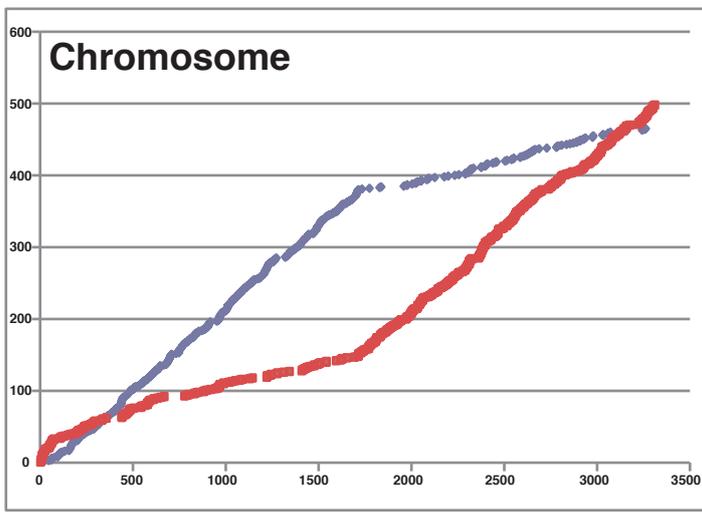


pAB510f



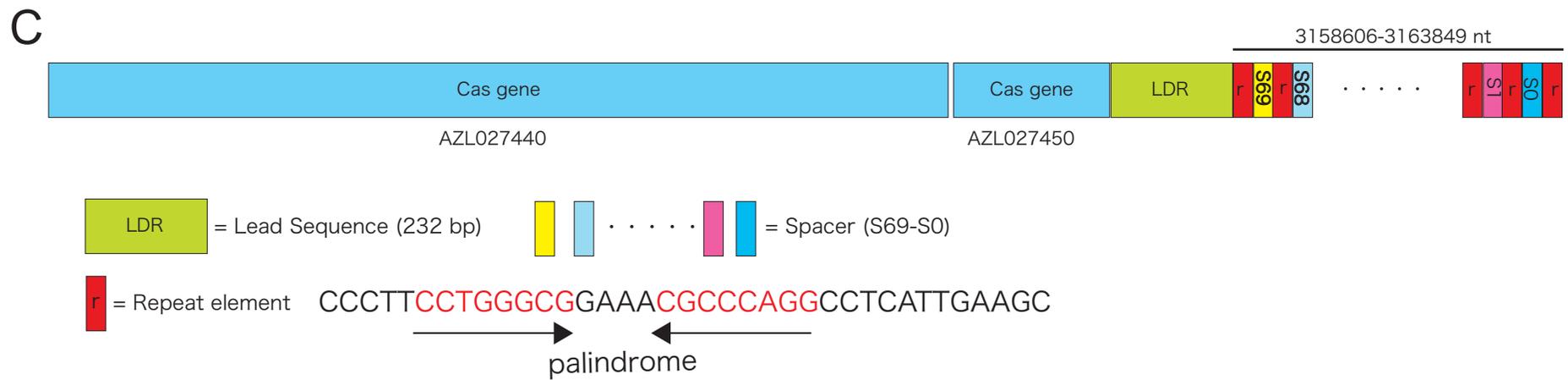
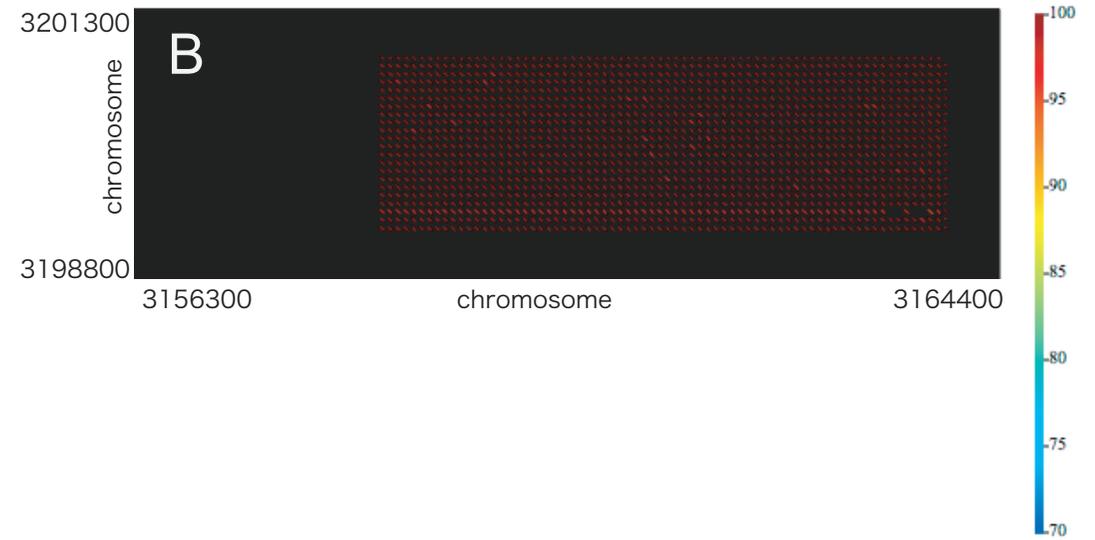
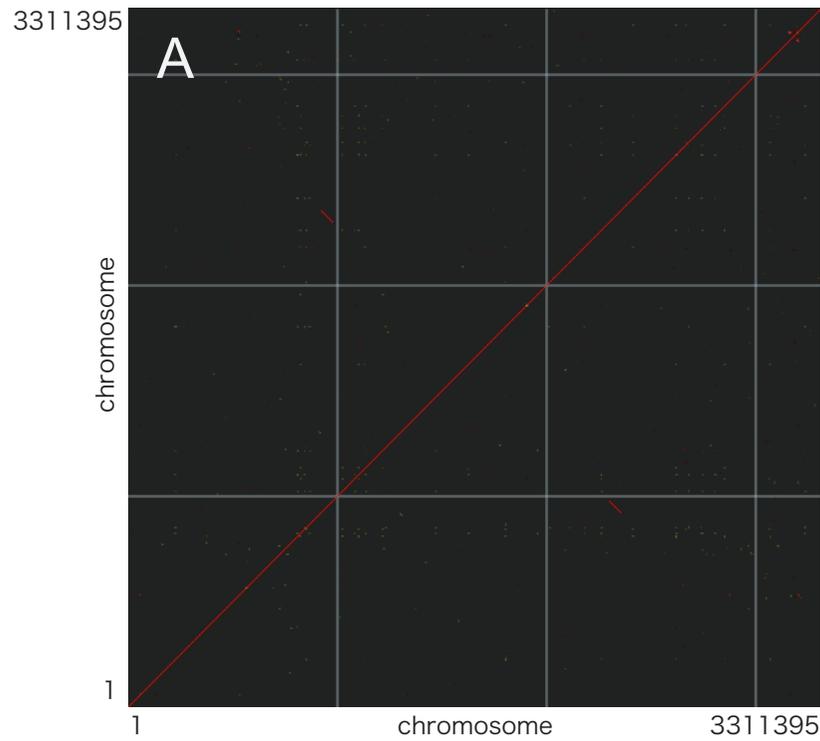
Color codes for functional categories

	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	TACATCTGCGACATCGTTTCGCCTTCGCGGTGGCCCGT	
3158680	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGCAAGGATCGAGTTTCGTAAGTCCGCGAAGGAGAGC	3158679
3158754	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGCAAGGATCGAGTTTCGTAAGTCCGCGAAGGAGAGC	3158753
3158830	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGCAAGGATCGAGTTTCGTAAGTCCGCGAAGGAGAGC	3158829
3158901	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TCAGCGTCCTTAGGATCGTCCGTGAGAAGGTCGGCCAT	3158900
3158976	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGCAGGATGGACCCAACCTCGAACTGGAAGAGGCCG	3158975
3159049	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGTCTGCCAAGTCCATCGGGGCGAAAGAGACGGGAT	3159048
3159123	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GTTCGGAAATATTGATCCGACCCCATTCGCGATCAT	3159122
3159196	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGGCCCGCCGCGTGC TGGGCGCCTCCAATGTCGGACC	3159195
3159270	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GGCGTCTGCCGTCTCCAGCCACTTCGCCAGCACTCC	3159269
3159343	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GGGTGAAGCGGGGTTCGTGGTAGGGGATGGGTGGTAG	3159342
3159418	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CCCGCGGCACCGGATCGAGGGGCCGCGGCGCAACGGC	3159417
3159492	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	ATCACAATGAACGGCGAACCACCTCGCCCACCCTTG	3159491
3159565	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGGGCGCGGGGAGGGGGAGCGCCCCCATTTCCCTCAGCC	3159564
3159641	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GCCATGTCGCCGTCTGCTGTGGTGTTCATCGTGTGCCACTT	3159640
3159720	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGCAAGGATCGAGTTTCGTAAGTCCGCGAAGGAGAGC	3159719
3159794	CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TCAGTCGGTGC GGATACCGCCTTGGCGATCACAGAGACGG	3159793
3159871			3159870

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

CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	3161426
3161427	
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGGAGGGAAACCACCAGCCGAGCAACGTCGCTCGGCTTG
3161503	3161502
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GACGCATGAAGTTCGTACCGGTGGCGTCCCGCGCATATC
3161579	3161578
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CCCTGCCACCCTGATCTGCAGCTTGTGCATGTCGAGCCG
3161655	3161654
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGCCGCTCTATGAGTTCGTGCCACGCGGTCCATCACGC
3161731	3161730
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGGCCATGGTCTATGACAGGAACGCCCTGGCCCTTGA
3161807	3161806
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	ATAAGAGGCTCTGGGCGGAAACGCCCTGGCCCTTGA
3161885	3161884
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GCCATTCCCTTGTGCTCTGGTCAAATGGGCAA
3161957	3161956
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGCCATACAGGCCAGGACTTCAGACCTTCTGGGAAATC
3162034	3162033
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	GACTCCCAACGGACGCTGACGAAATGGGGCGCACGGC
3162110	3162109
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TTGAACACTGCGGCGGAACACCCCATCGTGGTAACG
3162183	3162182
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AATGATATGGCTCTCTCGGGCTTCGTCAATGTAGACCC
3162258	3162257
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	AGCAAGAACTATGTGGATCGAAACATCCTCAATCTGAA
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	SGTNTTGGCTCAGTTCCGACCGCTGCGGATAAATC
3162848	3162847

CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TCTACCGGCAGCAATGCCGGCAGATCCTGATCGGGG	3162920
3162921		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	TGAATGCTGGGGCATCATGTCCGGCCGTATCCCGA	3162993
3162994		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CCGTGGAAAGTGACCGATCTGCGTTTGCTATGAACTGAG	3163069
3163070		
CCCTTCCTGGGCGGAAACGCCAGGCCTCATTGAAGC	CGTCAACCCAATCAGCCAGACCCTGGGTGCGGGCT	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 stem loop	stem	Anticodon domain stem loop	stem	Variable region	TψC domain stem loop	stem	Aminoacyl stem
AZL_r001	Type: Gly	GCTC AATGGTA	Anticodon: CCC (51753-51826)	GAGC AG AAGC TTCCCAA GCTT	ACGAC	GAGGG	TTCGATT CCCTT	CACCCGC TCCA
AZL_r002	Type: Leu	GCGG AATTGGTAG	Anticodon: CAG (115296-115382)	ACGC G CAGGT TTCAGGT ACCTGT	GGCAGAAATGTCGT	GGAAG	TTCGAGT CTCTT	CCTGGGC ACCA
AZL_r003	Type: Gly	GCTC AGGGGTA	Anticodon: GCC (265257-265331)	GAGC A CAACC TTGCCAA GGTTG	GGGTC	GAGGG	TTCGAAT CCCTT	CGCCCCG TCCA
AZL_r004	Type: Val	GCTC AGCGGGA	Anticodon: GAC (322374-322448)	GAGC A CTACG TTGACAT CGTAG	GGGTC	ACTGG	TTCAACT CCAGT	ACCGCC ACCA
AZL_r005	Type: Ala	GCTC AGCTGGGA	Anticodon: GGC (324313-324238)	GAGC G CTACA ATGGCAT TGTAG	AGGTC	AGGAG	TTCGATC CTCTT	TGGCTCC ACCA
AZL_r006	Type: Leu	GCGG AATTGGTAG	Anticodon: GAG (507267-507351)	ACGC G CAGCG TTGAGGT CGCTG	TGGGGCAACCCGT	GGAAG	TTCGAGT CTCTT	CGACCCG ACCA
AZL_r007	Type: Ile	GCTC AGTTGGTTA	Anticodon: GAT (522402-522478)	GAGC G CGCGC TTGATAA GCGTG	AGGTC	GGAGG	TTCAAAT CCTCC	CTGGCCC ACCA
AZL_r009	Type: Ala	GCTC AGTTGGGA	Anticodon: TGC (522508-522583)	GAGC G CCGTC TTGCAA GCAGG	AGGTC	GTCGG	TTCGATC CCGTC	TGCTCC ACCA
AZL_r012	Type: Thr	GCTC AGCTGGTA	Anticodon: TGT (645871-645796)	GAGC AG CTGA TTTGTA TCAG	AAGGTC	ACGGG	TTCGATT CCGT	ATCCGGC ACCA
AZL_r013	Type: Tyr	GCC GACTGGTTAAA	Anticodon: GTA (647024-647109)	GGC A GCAGA CTGTAAA TCTG	CCCGCTACCGCTAC	GCTGG	TTCGAAT CCAGC	CCCTTCC ACCA
AZL_r014	Type: Gly	GCTC AATGGTA	Anticodon: TCC (647175-647248)	GAGC AG AAGC TTCCCAA GCTT	ACGAC	GGGGG	TTCGATT CCCTT	CACCCGC TCCA
AZL_r015	Type: Trp	GCTC AATTGGTA	Anticodon: CCA (648712-648787)	GAGC A CCGGT CTCCAAA ACCGG	GGGTT	GGGGG	TTCGAGC CCCTC	CACTCCT GCCA
AZL_r016	Type: Glu	GTCT AGAGCCCTA	Anticodon: TTC (1080602-1080677)	GGAC A CCGCC CTTTCCG GCGCG	CGAC	ACGGG	TTCGACT CCGCT	CGGGGAC GCCA
AZL_r017	Type: Pro	GCGC AGCCTGGTA	Anticodon: GGG (1158120-1158196)	GCGC A TCAGA CTGGGGG TCTGG	GGGTC	GCAGG	TTCAAAT CCTGT	CGCTCCG ACCA
AZL_r018	Type: Ile (wobble)	GTTT AGCGGTTA	Anticodon: CAT (1196512-1196587)	GAAC G CGCCG CTCATAA CGGTG	TTGTC	GTCGG	TTCGAAT CCGGC	CGGGCCC ACCA
AZL_r019	Type: Asn	GCTC AGTTGGTA	Anticodon: GTT (1492007-1491932)	GAGC A GCGGA CTGTAAA TCCCG	TTGTC	GCTGG	TTCGAGT CCAGC	ACCGGGA GCCA
AZL_r020	Type: Cys	GCA GAGTGGTGA	Anticodon: GCA (1492312-1492239)	TGC A GCGGA CTGCAA TCCGT	GTAC	GTGGG	TTCGATT CCGCC	CCGTGCC TCCA
AZL_r021	Type: Pro	GCGC AGTCTGGTCA	Anticodon: TGG (1581577-1581654)	GCGC G CCAGT TTTGGGT ACTGG	AGGCC	CCCGG	TTCGAAT CCGGG	CGCCCCG ACCA
AZL_r022	Type: Arg	GCTC AGTTGGATA	Anticodon: TCT (1582059-1582135)	GAGC A CCGCC CTTCTAA GCGCC	AGGTC	GCAGG	TTCGAAT CCTGC	TGGGGTC GCCA
AZL_r023	Type: Lys	GCTC AGTCCGTA	Anticodon: TTT (1595581-1595506)	GAGC A CGTGA CTTTAAA TCATG	TGGCC	GTGGG	TTCGAAT CCCAC	CCGGATC ACCA
AZL_r024	Type: Gln	GCC AAGTGGTAA	Anticodon: TTG (1699351-1699277)	GGC A CCGGT TTTTGGT ACCGC	CATTC	CCAGG	TTCGAAT CCTGG	GCCCCCA GCCA
AZL_r025	Type: Val	GCTC AGCGGTA	Anticodon: TAC (2177531-2177457)	GAGC G TCTCG TTTACAC CGAGA	ATGTC	GGGGG	TTCGATC CCCTC	ATCGCCC ACCA
AZL_r026	Type: Leu	GCGG GAATTGGTAGA	Anticodon: TAG (2186884-2186800)	CGC G CTGGA TTTAGGT TCCAG	TGACTTCGCTCGT	ATGGG	TTCAAAT CCCTT	CGCCCCG ACCA
AZL_r027	Type: His	GCTC AGTTGGTA	Anticodon: GTG (2190435-2190360)	GAGC T CTCGG TTGTGGT CCAGG	CGGTC	GTGGG	TTCGAGT CCGAT	CCATCGC CCGA
AZL_r029	Type: Met	GCTC AGTGGTA	Anticodon: CAT (2296562-2296636)	GAGC A GGGGA ATCATAA TCCCT	TGTC	GGGGG	TTCAAAT CCCTC	CGCTGCT ACCA
AZL_r031	Type: Asp	GCTC AGTTGGTAA	Anticodon: GTC (2456790-2456714)	GAGC G CCGCC CTGTAC GCGCG	AGGCC	CGGGG	TTCAAAT CCGCT	CACTCGC GCCA
AZL_r032	Type: Lys	GCTC AGTTGGTA	Anticodon: CTT (2467038-2466963)	GAGC A GCTGA CTCTTAA TCAGC	GGGTC	CAAGG	TTCGAGT CCTTG	TGCGCCC ACCA
AZL_r033	Type: Lys	GCTC AGTTGGTA	Anticodon: CTT (2467137-2467062)	GAGC A GCTGA CTCTTAA TCAGC	GGGTC	CAAGG	TTCGAGT CCTTG	TGCGCCC ACCA
AZL_r034	Type: Thr	GCTC AGTGGTA	Anticodon: GGT (2529479-2529553)	GAGC A CTCCC TTGGTAA GGGAG	AGGTC	GAGAG	TTCAACT CTCTC	TGGCAGC ACCA
AZL_r035	Type: Thr	GCTC AGTGGTA	Anticodon: GGT (2529590-2529664)	GAGC A CTCCC TTGGTAA GGGAG	AGGTC	GAGAG	TTCAACT CTCTC	TGGCAGC ACCA
AZL_r036	Type: Gly	GCTC AGGGGTA	Anticodon: GCC (2625267-2625193)	GAGC A CAACC TTGCCAA GGTTG	GGGTC	GAGGG	TTCGAAT CCCTT	CGCCCCG TCCA
AZL_r037	Type: Ala	GCTC AGCTGGGA	Anticodon: GGC (2637734-2637659)	GAGC G CTACA ATGGCAT TGTAG	AGGTC	AGGAG	TTCGATC CTCTT	TGGCTCC ACCA
AZL_r038	Type: Leu	GCGG AATTGGTAGA	Anticodon: CAG (2637859-2637773)	CGC G CTAGG TTCAGGT CCTAG	TGGCTGAAAGCCGT	GGGGG	TTCGAGT CCCTC	CAAGCGC ACCA
AZL_r039	Type: Ser	GCA GAGCGGTTGAA	Anticodon: CGA (2645714-2645625)	TGC A CCGCA CTCGAAA TGCGG	CATACCCCTCACGGTATC	GGGAG	TTCGAAT CTCCC	CCCATCC GCCA
AZL_r041	Type: Leu	GCTC AATTTGGTAGA	Anticodon: CAA (2872146-2872230)	CGC G CCGGA CTCAAAA TCCGG	TTCCGCAAGGAGT	GTCGG	TTCGAGT CCGAC	CGCCCCG ACCA
AZL_r042	Type: Leu	GCGG AATTTGGTAGA	Anticodon: TAA (2880319-2880235)	CGC A AGGGA CTTAAAA TCCCT	CACCTCTTGGTAT	GTGGG	TTCGAGT CCCAC	CGCCCCG ACCA
AZL_r043	Type: Thr	GCTC AGGGGTA	Anticodon: CGT (2919939-2919865)	GAGC A ACTGA TTCGTAA TCAGT	AGGTC	CGGGG	TTCAAAT CCCTG	TTGGCGC ACCA
AZL_r044	Type: Ser	GCC GAGTGGTTGAA	Anticodon: GGA (3019901-3019811)	GGC G CAGCG CTGGAAA GTGTG	TATACGTCAAAAGCGTATC	GAGGG	TTCGAAT CCCTC	TCCTTCC GCCA
AZL_r045	Type: Arg	GCTC AGCTGGATA	Anticodon: ACG (3065221-3065145)	GAGC A TCAGA CTACGAA TCTGA	GGGTC	AGGAG	TTCGAAT CTCTT	CGGGCGC GCCA
AZL_r046	Type: Arg	GCTC AGCTGGATA	Anticodon: ACG (3068054-3067978)	GAGC A TCAGA CTACGAA TCTGA	GGGTC	AGGAG	TTCGAAT CTCTT	CGGGCGC GCCA
AZL_r047	Type: Ala	GCTC AGCTGGGA	Anticodon: CGC (3088422-3088497)	GAGC G ACAGC TTCCGAA TCGGT	AGGTC	GGGAG	TTCGATC CTCTT	CGGCTCC ACCA
AZL_r048	Type: fMet	GAGC AGCCCCGTA	Anticodon: CAT (3207660-3207584)	GCTC G TCAGG CTCATAA CCGTA	AGGCC	GCAGG	TTCAAAT CCTGC	CCCCGCA ACCA
AZL_r051	Type: Ala	GCTC AGTTGGGA	Anticodon: TGC (3211110-3211035)	GAGC G CCGTC TTGCAA GCAGG	AGGTC	GTCGG	TTCGATC CCGTC	TGCTTCC ACCA
AZL_r052	Type: Ile	GCTC AGTTGGTTA	Anticodon: GAT (3211216-3211140)	GAGC G CGCGC TTGATAA GCGTG	AGGTC	GGAGG	TTCAAAT CCTCC	CTGGCCC ACCA
AZL_r054	Type: Arg	GCTC ATCAGGATA	Anticodon: CCT (3263365-3263441)	GAGC G CGAGT TTCTTAA ACTTG	AGGCA	GCAGG	TTCGAGT CCTGC	CGGGGTC ACCA

Aminoacyl Stem	D domain		Anticodon domain		Variable region	T _W C domain		Aminoacyl stem
	stem	loop	stem	loop		stem	loop	
AZL_ra01	Type: Gln		Anticodon: CTG (62973-63046)					
TGGGGGA TC	GTC T AGCGGTA		GGAC A GCGGA CTCGAC 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	CCCCTC GCCA
AZL_ra03	Type: Val		Anticodon: CAC (410196-410270)					
GGCGCGG TA	GCTC AGCGGGA		GAGC A TTCGC TTCACAC GCGAA		GGTC	ACAGG	TTCAACT CCTGT	CGCGCC ACCA
AZL_ra04	Type: Pro		Anticodon: CGG (474014-474090)					
CGGAGAG TA	GCTC AGCCTGGTA		GAGC A CTGCT TTCGGGA GGCAG		GGCC	GGAGG	TTCGAAT CCTCT	CTCTCC ACCA
AZL_ra05	Type: Pro		Anticodon: CGG (474308-474384)					
CGGAGAG TA	GCTC AGCCTGGTA		GAGC A CTGCT TTCGGGA GGCAG		GGCC	GGAGG	TTCGAAT CCTCT	CTCTCC ACCA
AZL_ra06	Type: Ser		Anticodon: TGA (544324-544235)					
GGATGGG TG	GCCG AGTGGTTTA		AGGC A GCGGT CTTGAAA ACCGC		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	CTGGCC ACCA
AZL_ra09	Type: Ala		Anticodon: TGC (566961-567036)					
GGGGGCG TA	GCTC AGTTGGGA		GAGC G CCTGC TTTGCAA GCAGG		AGGTC	GTCGG	TTCGATC CCGTC	TGCTCC ACCA
AZL_ra14	Type: fMet		Anticodon: CAT (971336-971260)					
CGCGGGG TG	GAGC AGCCCGGTA		GCTC G TCAGG CTCATAA CCTGA		AGGCC	GCAGG	TTCAAAT CCTGC	CCCCGCA ACCA
AZL_ra17	Type: Ala		Anticodon: TGC (974794-974719)					
GGGGGCA TA	GCTC AGTTGGGA		GAGC G CCTGC TTTGCAA GCAGG		AGGTC	GTCGG	TTCGATC CCGTC	TGCTCC 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 CTCGAA GGCAG		AGGTC	ATGAG	TTCGAAT CTCGT	CGGGTC GCCA
AZL_ra23	Type: Ala		Anticodon: TGC (1278873-1278798)					
GGGGGCA TA	GCTC AGTTGGGA		GAGC G CCTGC TTTGCAA GCAGG		AGGTC	GTCGG	TTCGATC CCGTC	TGCTCC ACCA
AZL_ra24	Type: Ile		Anticodon: GAT (1278979-1278903)					
GGGCTAG TA	GCTC AGTTGGTTA		GAGC G CGCGC TTGATAA GCGTG		AGGTC	GGAGG	TTCAAAT CCTCC	CTGGCC ACCA
AZL_rb03	Type: Ala		Anticodon: TGC (606446-606371)					
GGGGGCA TA	GCTC AGTTGGGA		GAGC G CCTGC TTTGCAA GCAGG		AGGTC	GTCGG	TTCGATC CCGTC	TGCTCC ACCA
AZL_rb04	Type: Ile		Anticodon: GAT (606563-606487)					
GGGCTAG TA	GCTC AGTTGGTTA		GAGC G CGCGC TTGATAA GCGTG		AGGTC	GGAGG	TTCAAAT CCTCC	CTGGCC ACCA
AZL_rc01	Type: fMet		Anticodon: CAT (484080-484004)					
CGCGGGG TG	GAGC AGCCCGGTA		GCTC G TCAGG CTCATAA CCTGA		AGGCC	GCAGG	TTCAAAT CCTGC	CCCCGCA ACCA
AZL_rc04	Type: Ala		Anticodon: TGC (487527-487452)					
GGGGGCA TA	GCTC AGTTGGGA		GAGC G CCTGC TTTGCAA GCAGG		AGGTC	GTCGG	TTCGATC CCGTC	TGCTCC ACCA
AZL_rc05	Type: Ile		Anticodon: GAT (487644-487568)					
GGGCTAG TA	GCTC AGTTGGTTA		GAGC G CGCGC TTGATAA GCGTG		AGGTC	GGAGG	TTCAAAT CCTCC	CTGGCC ACCA
AZL_rd01	Type: Asn		Anticodon: GTT (271587-271662)					
TCCCGGT TA	GTC T AGAGCCCTA		GGAC A GCGGA CTGTAA TCCGC		TTGTC	GCTGG	TTCGAT CACAG	ACCGGA GCCA
AZL_rd02	Type: Asp		Anticodon: CTC (271667-271743)					
GCGGGTG TA	GCTC AGTTGGTTA		GAGC G CCGGC CTGTAC GCCGG		AGGCC	GCGGG	TTCAGT CCCGT	CACTGC GCCA
AZL_rd03	Type: Phe		Anticodon: GAA (419945-420020)					
GCCAGG TA	GCTC AGTTGGTA		GAGC A GGGGA CTGAAA TCCCC		GTGTC	GGCGG	TTCAAAT CCGTC	CCTGGCC ACCA
AZL_rd04	Type: Glu		Anticodon: TTC (549351-549276)					
GTCCCGC TC	GTCT AGAGCCCTA		GGAC A CCGCC CTTTAC GCGGG		CGAC	ACGGG	TTCGACT CCCGT	CGGGAC GCCA
AZL_rd05	Type: Glu		Anticodon: CTC (549471-549396)					
GTCCCA TC	GTCT AGAGCCCTA		GGAC A CCGCC CTCTAC GCGGG		TAAC	AGGGG	TTCGACT CCCCT	TGGGAC GCCA
AZL_rd06	Type: Glu		Anticodon: CTC (549614-549539)					
GTCCCA TC	GTCT AGAGCCCTA		GGAC A CCGCC CTCTAC GCGGG		TAAC	AGGGG	TTCGACT CCCCT	TGGGAC GCCA
AZL_re02	Type: Ile		Anticodon: GAT (18041-18117)					
GGGCTAG TA	GCTC AGTTGGTTA		GAGC G CGCGC TTGATAA GCGTG		AGGTC	GGAGG	TTCAAAT CCTCC	CTGGCC ACCA
AZL_re03	Type: Ala		Anticodon: TGC (18147-18222)					
GGGGGCG TA	GCTC AGTTGGGA		GAGC G CCTGC TTTGCAA GCAGG		AGGTC	GTCGG	TTCGAT CCGTC	TGCTCC ACCA
AZL_re06	Type: fMet		Anticodon: CAT (21600-21676)					
CGCGGGG TG	GAGC AGCCCGGTA		GCTC G TCAGG CTCATAA CCTGA		AGGCC	GCAGG	TTCAAAT CCTGC	CCCCGCA 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		GGTTC	ACTGG	TTCAACT CCAGT	CGGGAC ACCA
AZL_re09	Type: Gly		Anticodon: CCC (449699-449625)					
GCGGGCG TA	GCTC AGGGGTA		GAGC A CAACC TTGCCAA GGTTC		GGTTC	GAGGG	TTCGAAT CCCTT	CGCCGC TCCA
AZL_re10	Type: Asp		Anticodon: GTC (449780-449704)					
GCGGGTG TA	GCTC AGTTGGTTA		GAGC G CCGGC CTGTAC GCCGG		AGGCC	GCGGG	TTCAGT CCCGT	CACTGC GCCA
AZL_re11	Type: Ala		Anticodon: GGC (449929-449854)					
GGGGGCA TA	GCTC AGTTGGGA		GAGC G CTACA ATGGCAT TGTAG		AGGTC	AGGAG	TTCGAT CCTCT	TGCTCC ACCA
AZL_re12	Type: Leu		Anticodon: CAG (450054-449968)					
CCCGTTG TG	GCGG AATTGGTAG		ACGC G CTAGG TTCAGGT CTAG		TGGCTGAAAGCCGT	GGGGG	TTCGAT CCTCT	CAAGGC ACCA

Supplementary Figure 6.

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



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



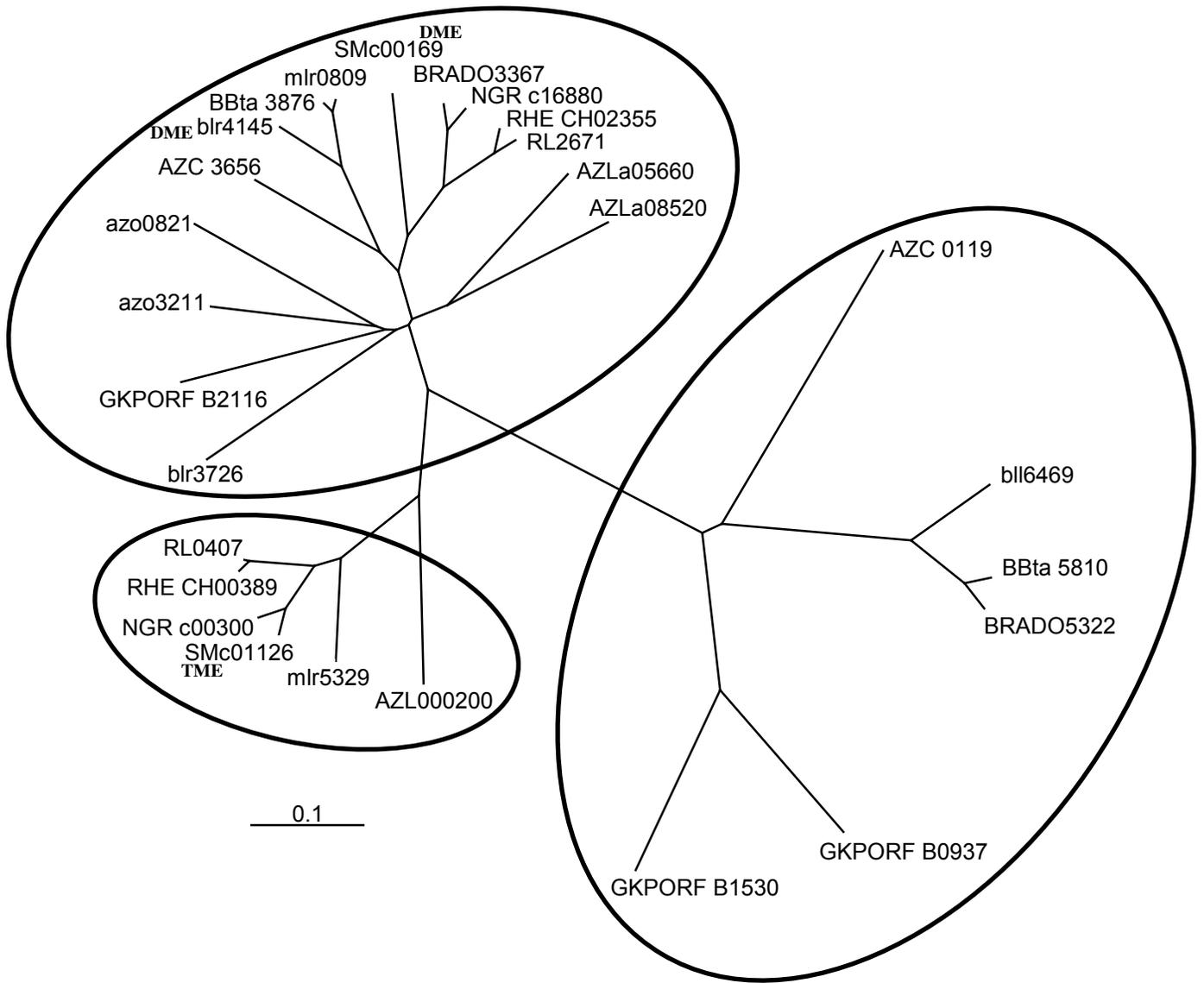
variation3 (*rrn5*)



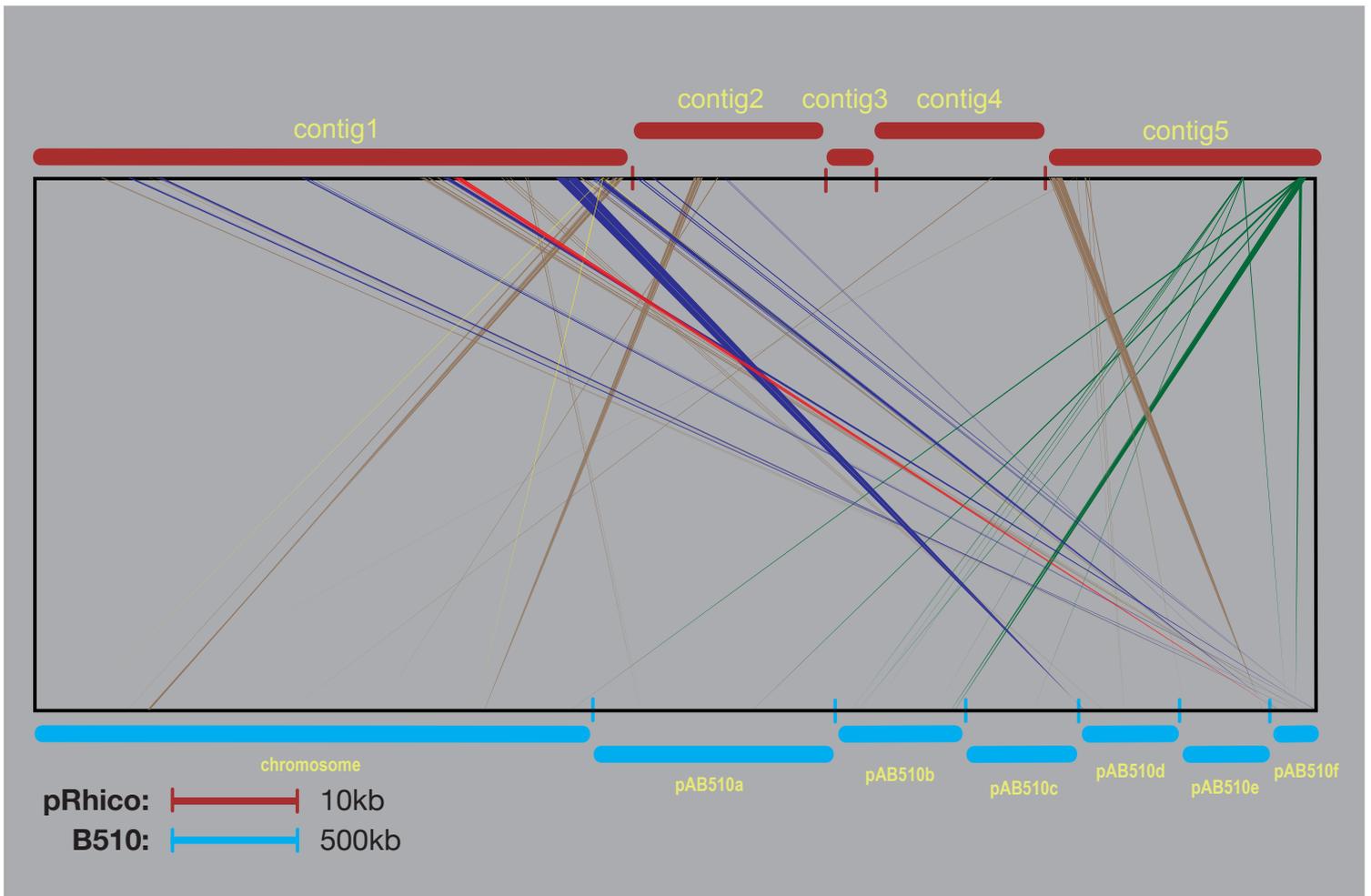
5' (471791)

TCAGAAATCCAGCATCGCGTTGTCGATCACCTCCTTCATGACGAAGAAGGTGCGGTCTGCCGCACGCCGGAAGCGCGATCAGTTGCTGCCCGTGCAG
ACGGTTGAAATCCGCGATGTCGCTGACCCGGATCTTCAGGAAGAAATCGAAGTCCCCGGCGACGAGATGACAGTCCAGCACGAAGGGCAGCTTGGCGAT
CGCCTGCTCGAACCGCGCGAAGCTTCCGGCGTCGAGCGGTCCAACCTCGACCCCGACCATACCAGCGTGCCCGATCCACCTTCGCCGGCTCGATCTC
GGCCCGGACGGAACGGATGTAGCCCTCGGCGAAAAGCCGTTGGGTGTGCCGGTGGCAGGTGGCCGGGCTGAGATTACCCGCCGCCGAGATCGGCGTT
CGTCATCCGGCCATCCCTTTGAAGAAGCTGGAGAATCTTGGGTGCGGCCGGGCTGGGCCCTCATCGGCGGTTGGTTTCTCTCATTTTCATCCATGAATT
TGGGATGGATAGCGCACAAAATGGCATTTATCAGCCTGAAACAGCAATCCGATAGCGAAGTTTGAGAGCACCTTTCATTTCTACGGTGATACCGTCCG
TTCACTTGAGACAGTGGAGCCCCGCCATGCTGCGCCTCGATCGCTTCGAACGTTATCCGCTCACCTTCGGCCCGACTCCCATCGAACATCTGCCCCGCC
TGACCGAGGCGCTCGGCGGCAAGGTCGAGATCTACGCCAAGCGCGACGACTGCAACTCCGGCCTCGCCATGGGCGGCAACAAGCTGCGCAAGCTGGAAT
ACATCGTTCCCGACGCCATCGCGGCCGGCGCCGACACGCTGGTGTCCATCGGCGGCGTGCAGTCGAACCACACCCGCATGGTGGCGGCCACCGCCGCA
AGCTCGGCATGAAGTGCCTGGTTCGTCAGGAGAGCTGGGTTCCGCACGAGGACGCCGCTCTATGACCGCGTCGGCAACATCCTGCTGACCCGCTGATGG
GCGCCGACAGCCGCATCGTTTCCGACGGCTTCGACATCGGCATCCGCAAGAGCTGGGAGGACGCCATCCAGTCGGTGAAGGATGTCGGCGGCAAGCCCT
ACGGCATTCCGGCCGGCGGCTCCGTCCACAAATACGGCGGTCTCGGCTATGTCCGCTTCGCCGAGGAGGTCCGCAAGCAGGAGGCCGAGCTCGGCTTCA
AGTTCGACTACATCGTCTGCTGCTGACCGGCTCGACCCAGGGCGGCATGATCGTCCGCTTCGCCGCCGACGACCGCGCCGACCGCGTGATCGGCA
TCGACGCCTCCGGCACGCCGGAACAGACCCGCGCCAGGTCCGCCAGATCGTCGACAACACGGCCGAGCTGGTTCGAACTCGGCCGCAAGGTCCGCGACG
ACGAGATCGTCATCCTGGAAGACTATGCCTATCCGGCCTATGGCGTGCCAGCGCCGAGACCAACGAGGCGATCCGCCTGGCCGCCCGCACCAGGCGCA
TGATCACCGACCCGGTCTATGAGGGCAAGTCGATGCAGGGCATGATCGACCTCGTGAAGAAGGGCTGGTTCCCCGAAGGCTCGAAGGTCCTCTACGCC
ATCTCGGCGGCGCCCCGGCGCTGAACGGCTACAGCTACACCTACCGGAACGGCTGA-3' (470631)

Supplementary Figure 8



Supplementary Figure 9



Supplementary Figure 10