

Modular design of the *nic*-gene cluster within the *Arthrobacter* genus



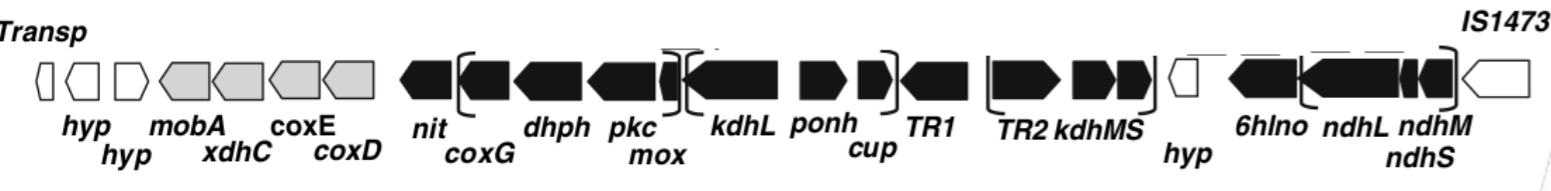
UNIVERSITATEA "ALEXANDRU IOAN CUZA" din IAȘI

www.uaic.ro

Marius I. Mihășan*

Laboratory of Biochemistry, Faculty of Biology, University "A. I. Cuza", Iași

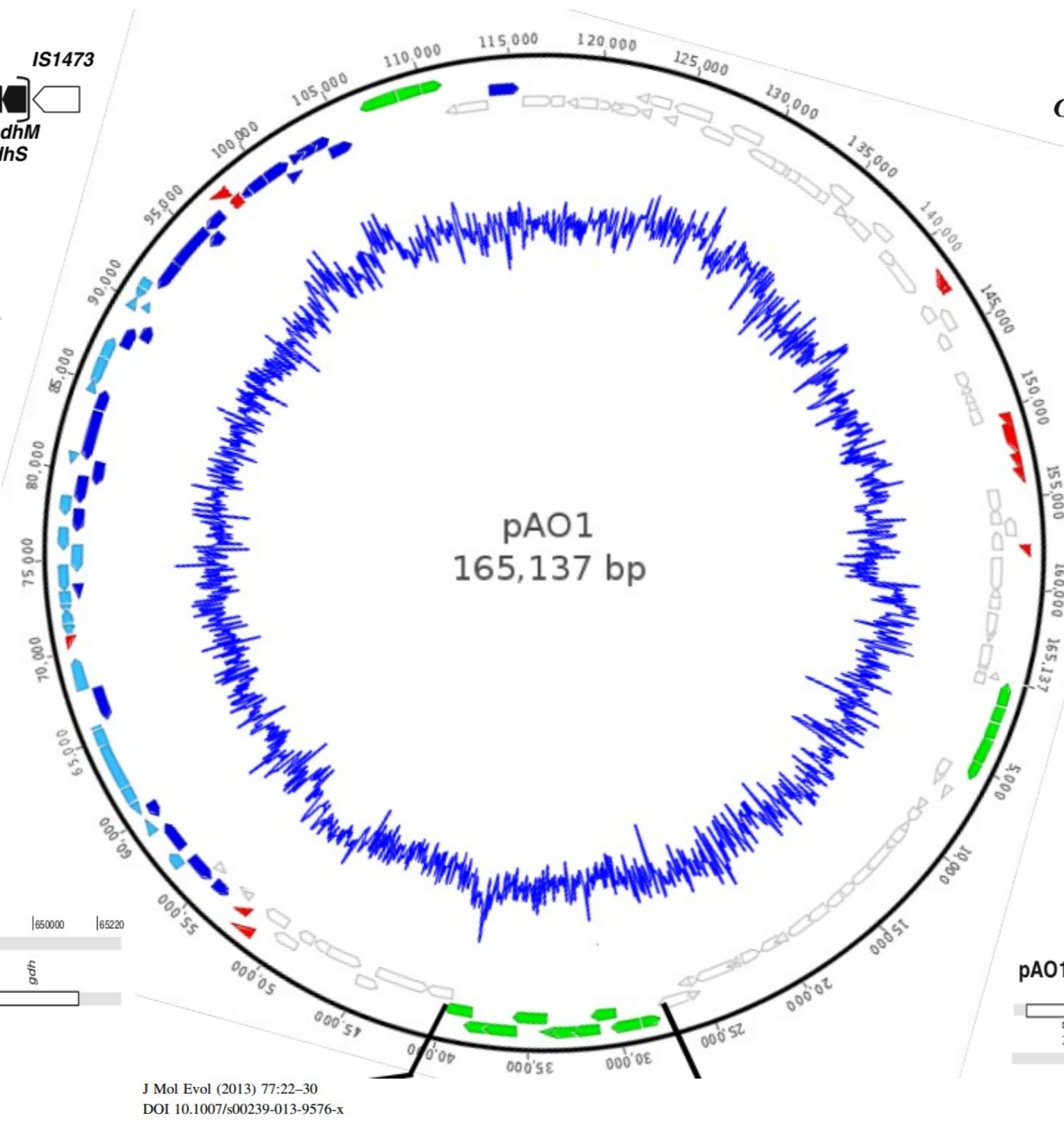
Nicotine-metabolism



Appl Microbiol Biotechnol (2006) 69: 493–498
DOI 10.1007/s00253-005-0226-0

MINI-REVIEW

Roderich Brandsch
Microbiology and biochemistry of nicotine degradation

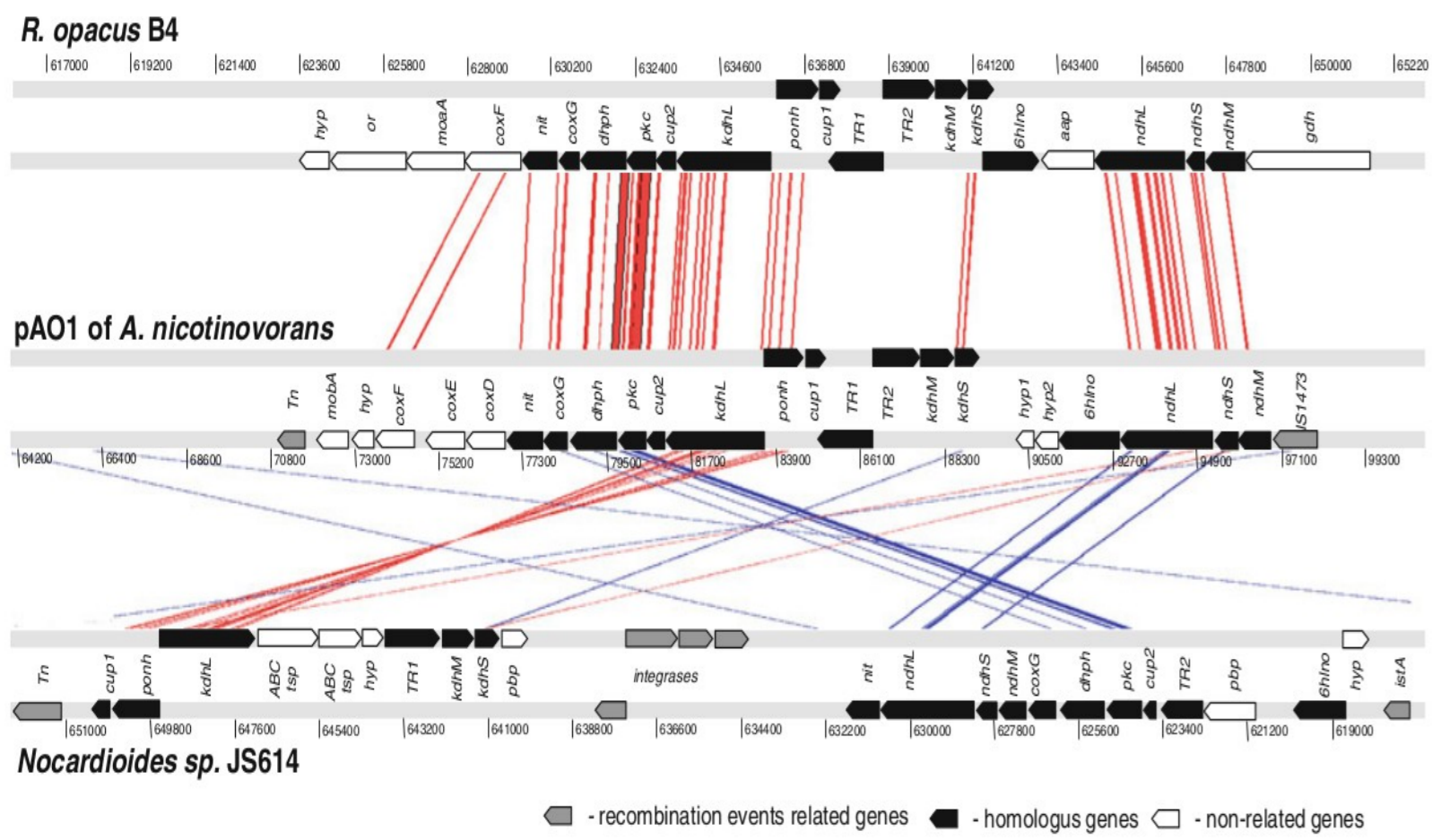


J Mol Evol (2013) 77:22–30
DOI 10.1007/s00239-013-9576-x

ORIGINAL ARTICLE

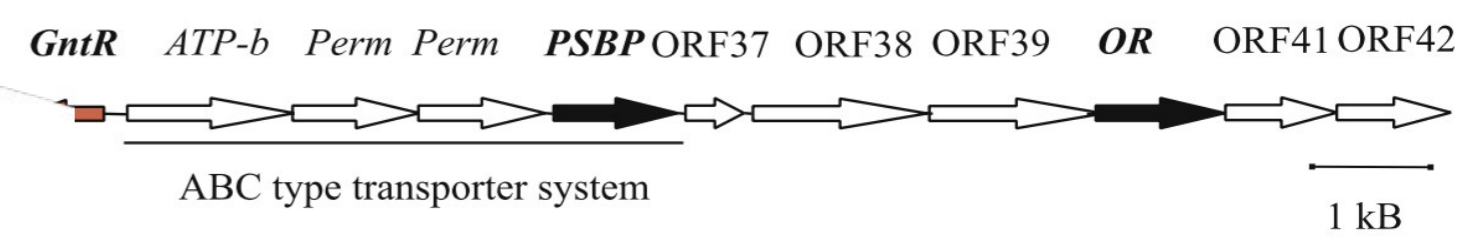
pAO1 of *Arthrobacter nicotinovorans* and the Spread of Catabolic Traits by Horizontal Gene Transfer in Gram-Positive Soil Bacteria

Marius Mihasan · Roderich Brandsch

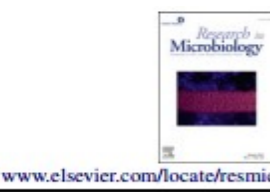


◻ - recombination events related genes ◼ - homologous genes ◻ - non-related genes

Xylose-metabolism

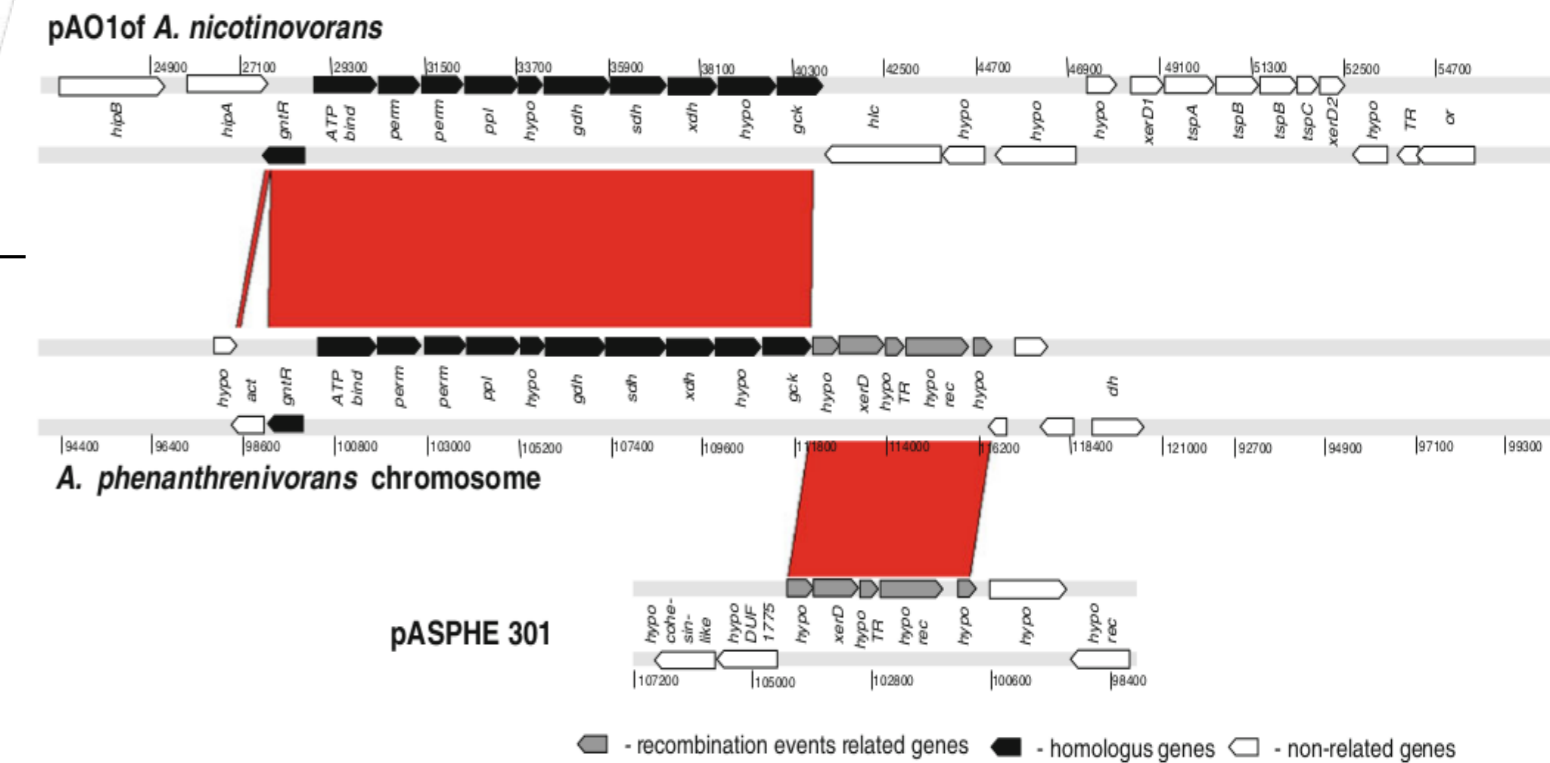


Institut Pasteur
Research in Microbiology 164 (2013) 22–30



Evidence of a plasmid-encoded oxidative xylose-catabolic pathway in *Arthrobacter nicotinovorans* pAO1

Marius Mihasan ^{a,*}, Marius Stefan ^{a,1}, Lucian Hritcu ^{a,1}, Vlad Artenie ^a, Roderich Brandsch ^b



◻ - recombination events related genes ◼ - homologous genes ◻ - non-related genes



UNIVERSITATEA "ALEXANDRU IOAN CUZA" din IAȘI

www.uaic.ro

Arthrobacter nicotinovorans and pAO1

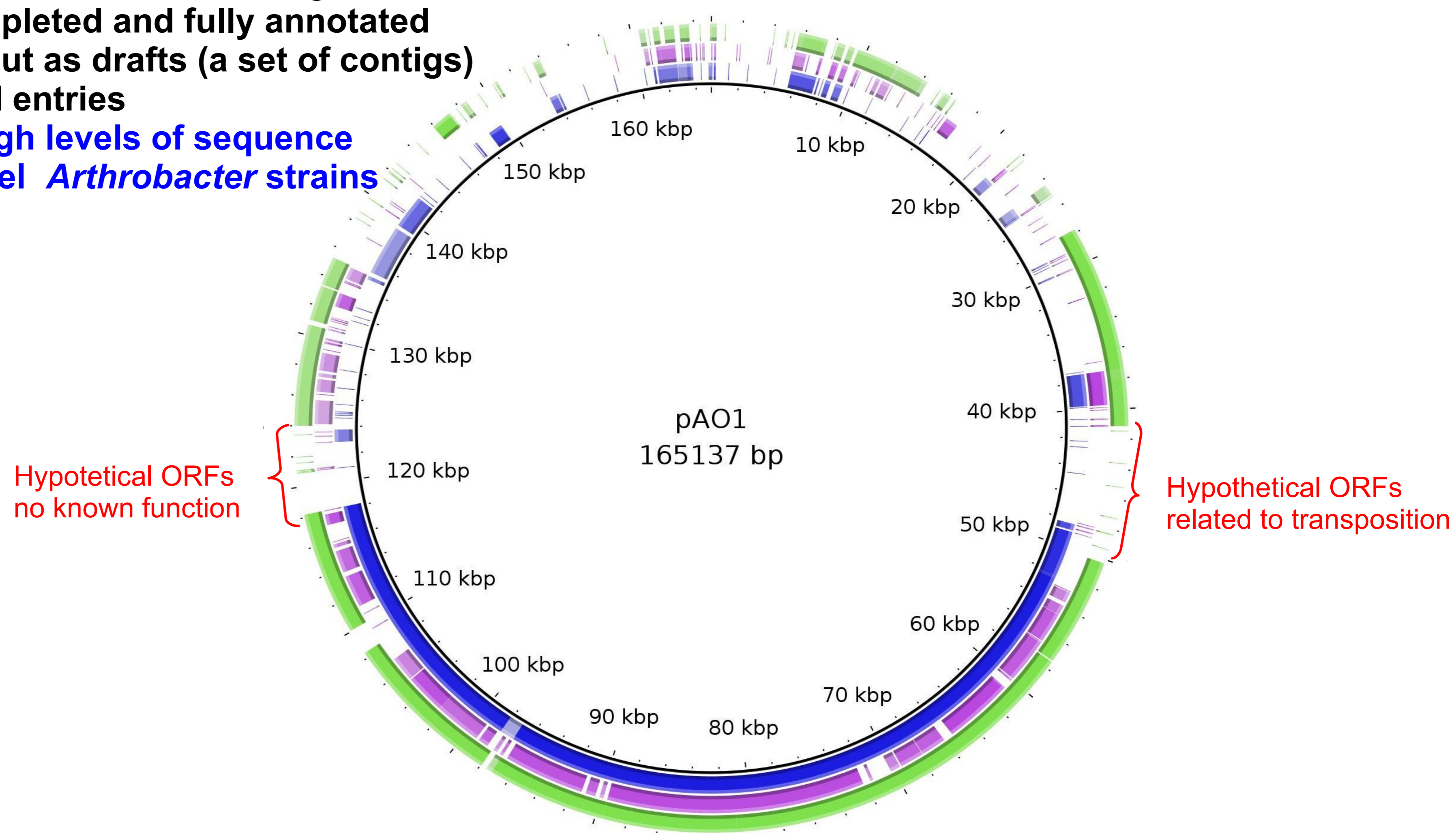
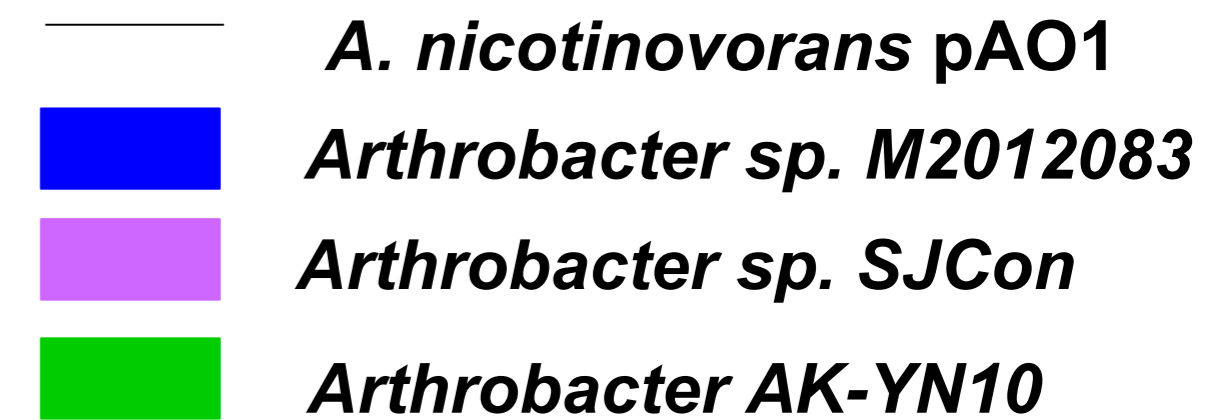
pAO1 shares most of its *nic*-genes with other *Arthrobacter* genomes



25 entries in GenBank for *Arthrobacter* plasmids
similarities between pAO1 and other
Arthrobacter plasmids are rather low

37 entries in GenBank for *Arthrobacter* genomes:
- 5 genomes are completed and fully annotated
- 10 are completed, but as drafts (a set of contigs)
- 22 are uncompleted entries

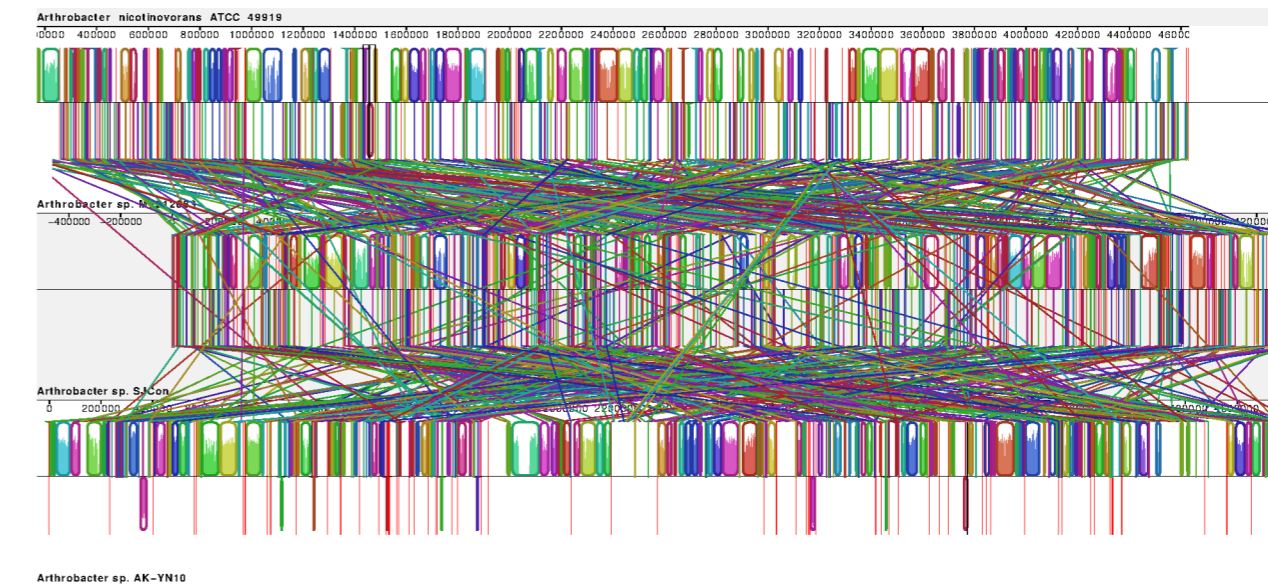
pAO1 shows high levels of sequence
similarity with 3 novel *Arthrobacter* strains





Draft genomes

Arthrobacter sp. M2012083 (GI:NZ_AKKK000000000)
Arthrobacter sp. SJCon (GI: NZ_AOFD000000000)
Arthrobacter sp. AK-YN10 (GI: NZ_AVPD000000000)



Contig assembly using
MAUVE

weight: 200

Min LCB

Scoring Matrix:

HOXD

Reference genomes

Arthrobacter aurescens TC1 (GI:NC_008711),
Arthrobacter chlorophenolicus A6 (GI:NC_011886),
Arthrobacter arilaitensis Re117 (GI:NC_014550)
Arthrobacter phenanthrenivorans Sphe3 (GI:NC_015145)
Arthrobacter nitroguajacolicus Rue61a (GI:NC_018531)

Genome annotation using RAST

v 2.0

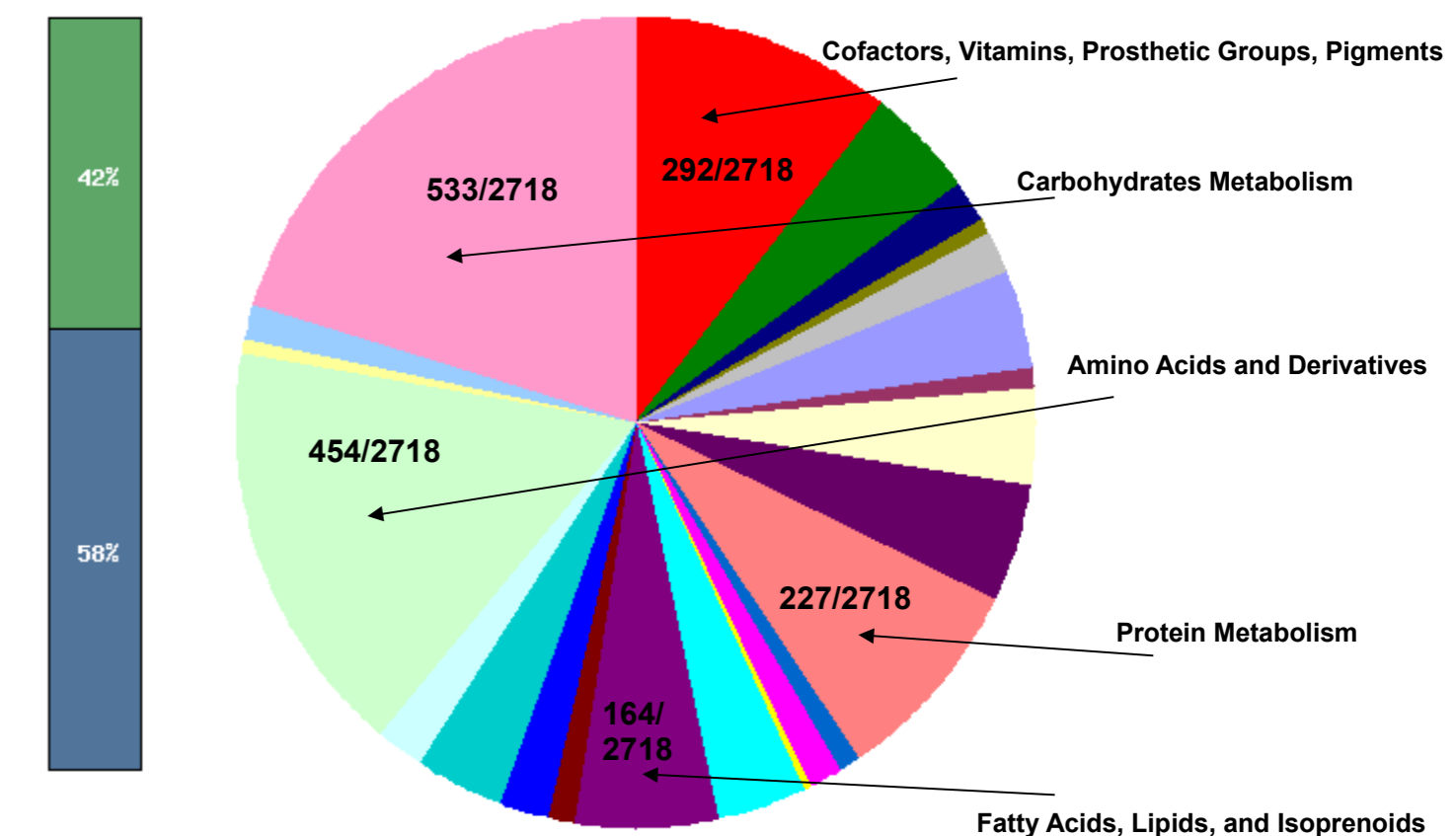
Final genomes

Gene collinearity
MAUVE

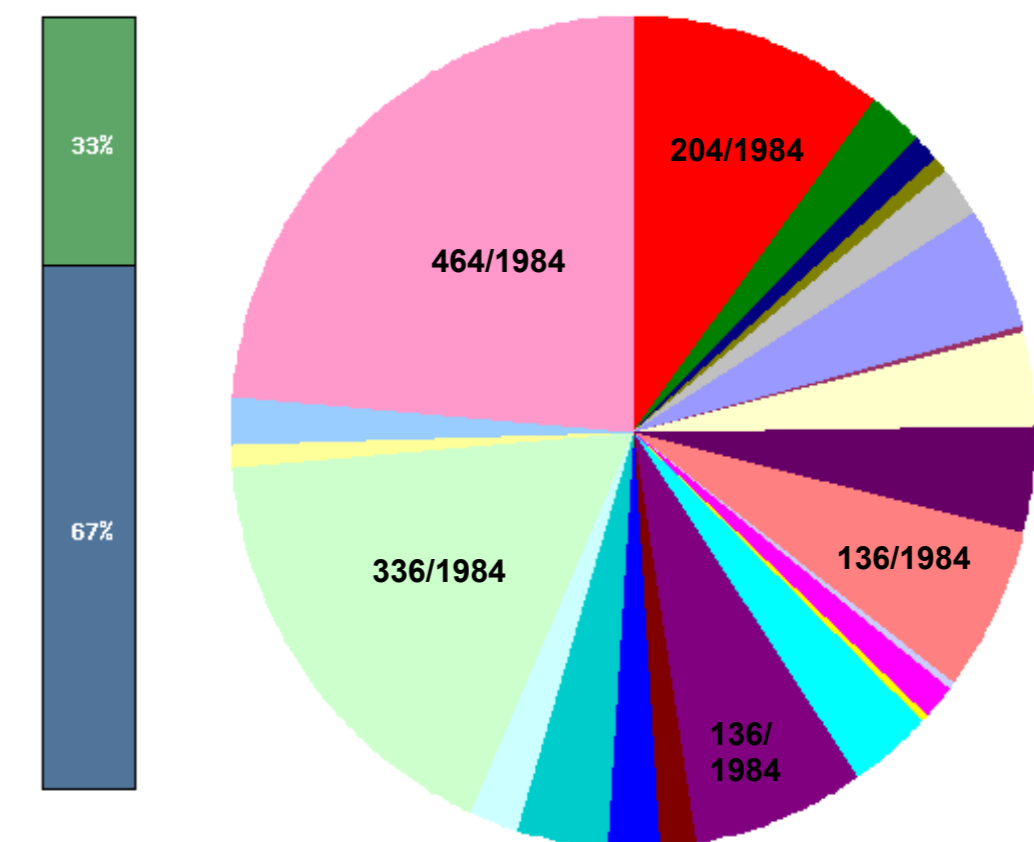
Sequence similarity
BLAST, BRIG

General statistical data and metabolic profiles of the annotated *Arthrobacter* genomes

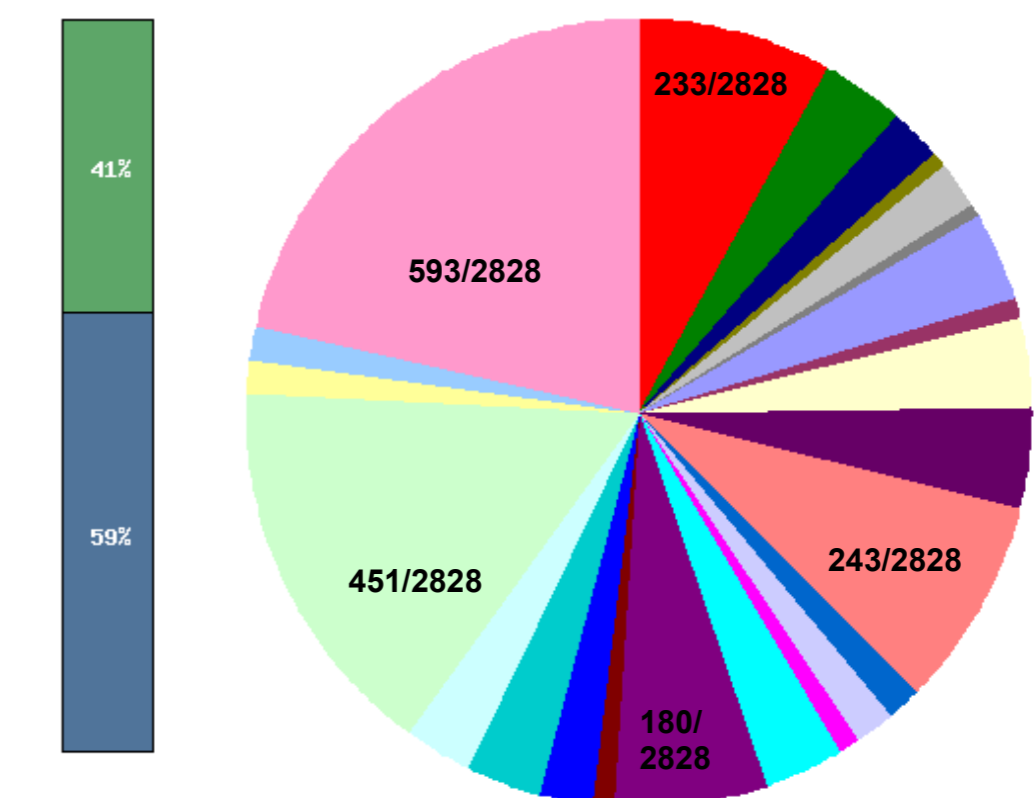
Arthrobacter sp. M2012083



Arthrobacter sp. SJCon



Arthrobacter AK-YN10

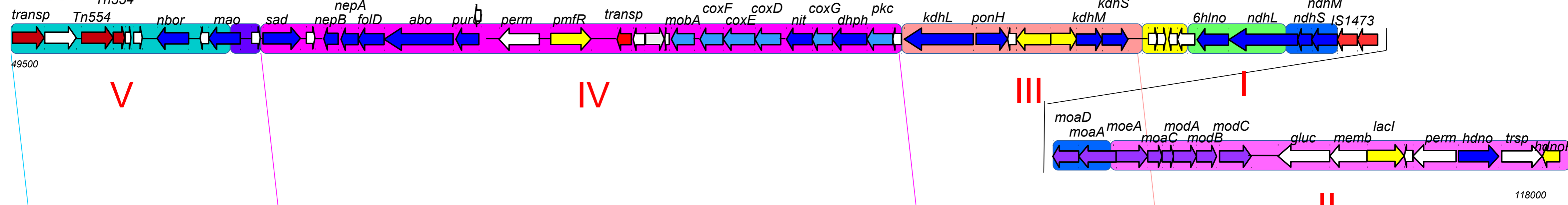


Assembled genomes	<i>Arthrobacter</i> sp. M2012083			<i>Arthrobacter</i> sp. SJConn		<i>Arthrobacter</i> sp. AK-YN10	
Reference genomes	<i>Arthrobacter aureescens</i> TC1	<i>Arthrobacter nitroguajacolicus</i> Rue61a	<i>Arthrobacter chlorophenolicus</i> A6	<i>Arthrobacter phenanthrenivorans</i> Sphe3	<i>Arthrobacter chlorophenolicus</i> A6	<i>Arthrobacter aureescens</i> TC1	<i>Arthrobacter nitroguajacolicus</i> Rue61a
Number of Contigs:	67			142		107	
Number of assembly bases:	4629172			4389620		4839751	
Number of reference bases:	4597686	4736495	4395537	4250414	4395537	4597686	4736495
Number of LCBs:	37	53	257	134	79	98	105
Number of Blocks:	89	100	285	222	173	135	142
Breakpoint Distance:	89	100	285	222	173	135	142
DCJ Distance:	81	91	276	213	164	126	134
SCJ Distance:	178	200	570	444	346	270	284
Number of SNPs:	437118	441536	601223	534528	531846	635011	642135
Number of Gaps in Reference:	3707	3793	8645	6602	6798	6689	6937
Number of Gaps in Assembly:	3583	3643	9053	7788	8342	7207	7349
Total bases missed in reference:	806998	846192	1831672	1132545	1310535	1408168	1523527
Percent bases missed:	17.55%	17.87%	41.67%	26.65%	29.82%	30.63%	32.17%
Total bases extra in assembly:	601474	607774	2003703	1032185	1165005	1230027	1244537
Percent bases extra:	12.99%	13.13%	43.28%	23.51%	26.54%	25.42%	25.71%
Number of extra contigs:	8	8	13	40	37	56	58
Number of Inter-LCB Boundaries:	6	11	25	14	9	13	9
Contig N50:	129641	129641	129641	61782	61782	131717	131717
Contig N90:	33265	33265	33625	18136	18136	28762	28762

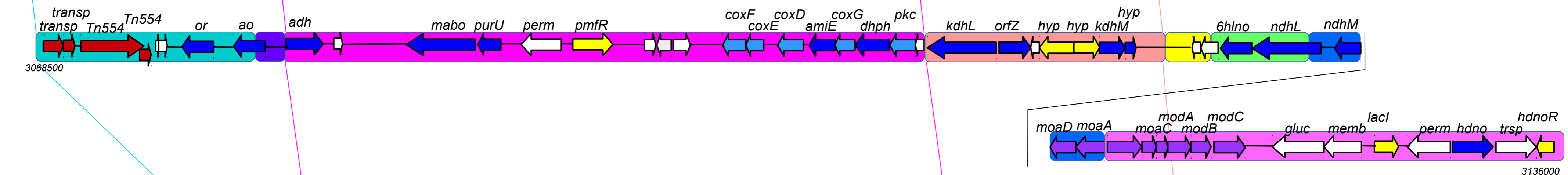
Nic-gene cluster collinearity analysis among 3 *Arthrobacter* genomes



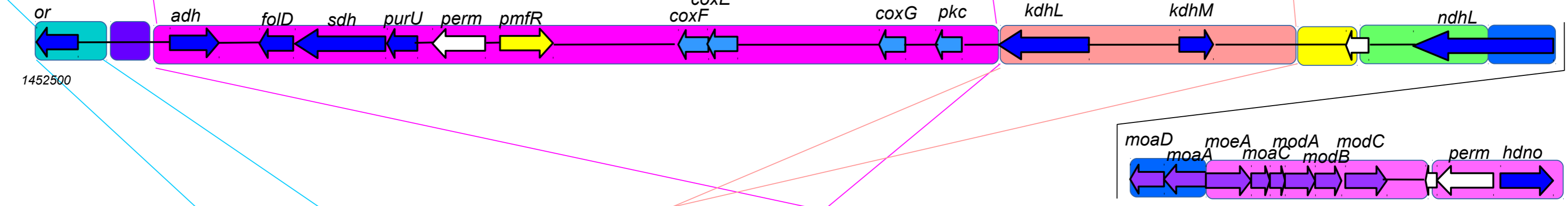
A. *nicotinovorans* pAO1



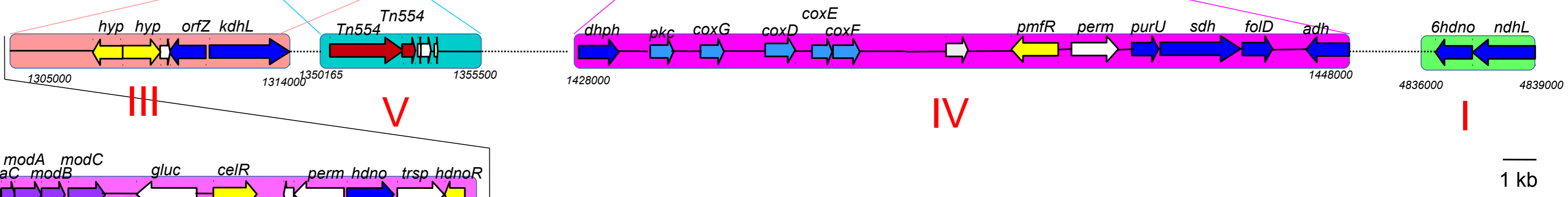
Arthrobacter sp. M2012083



Arthrobacter sp. SJCon

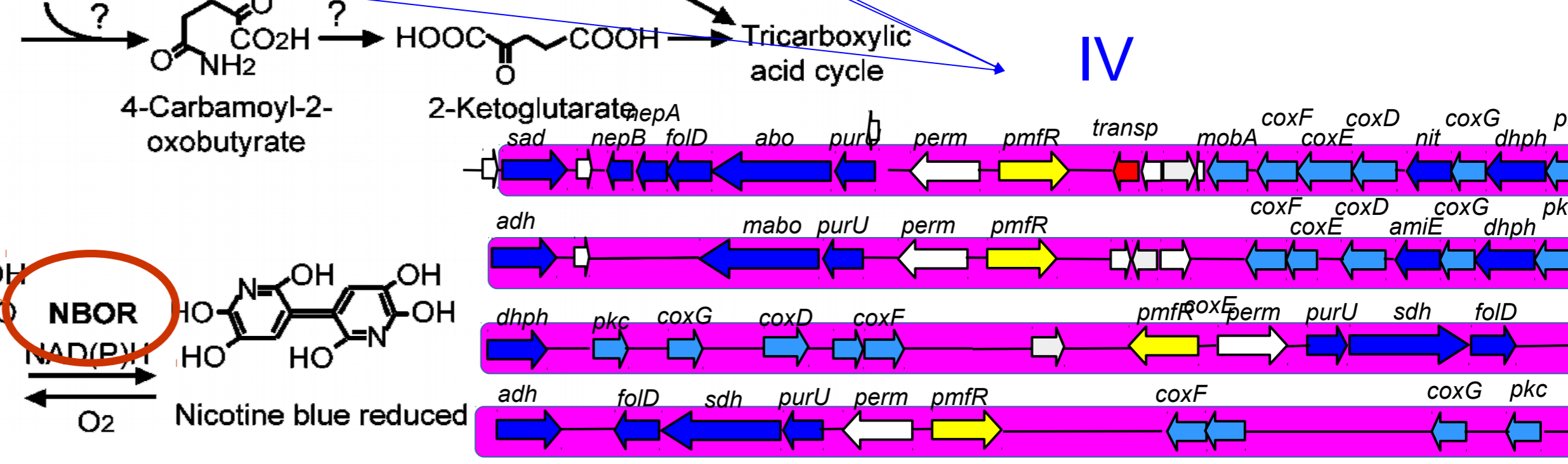
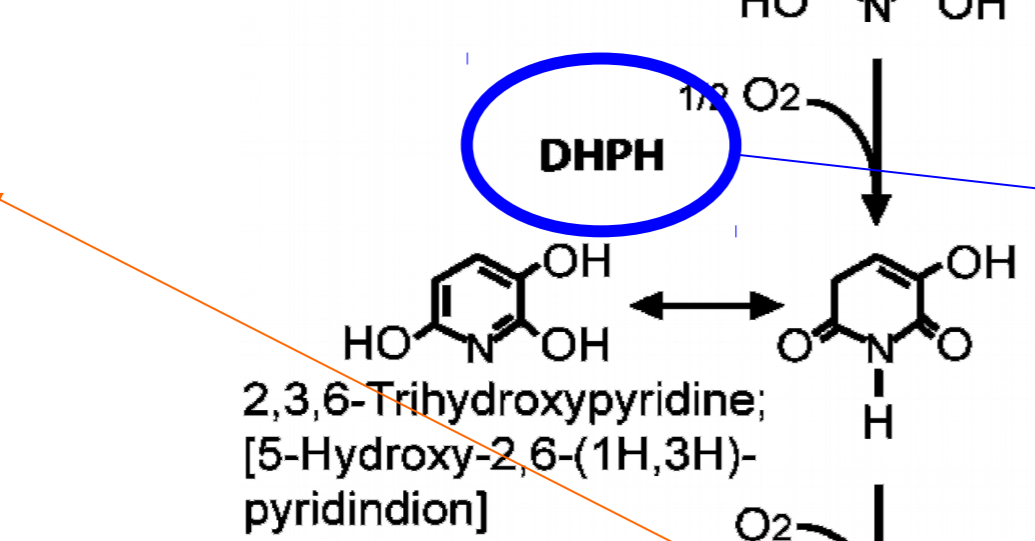
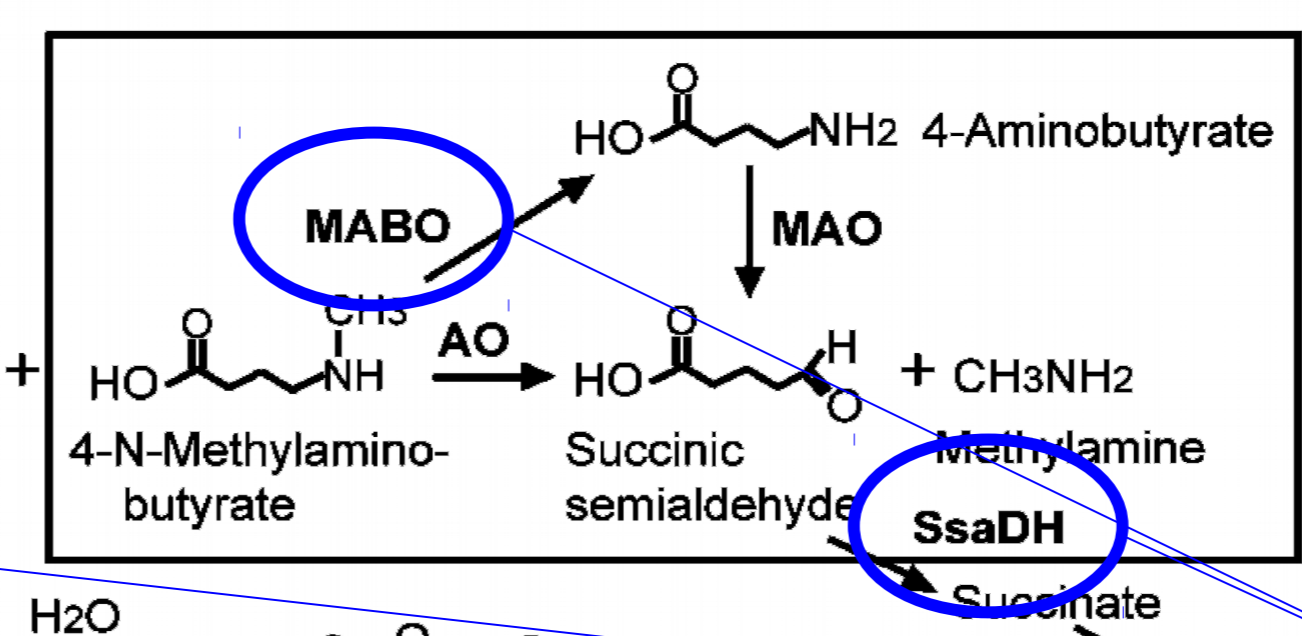
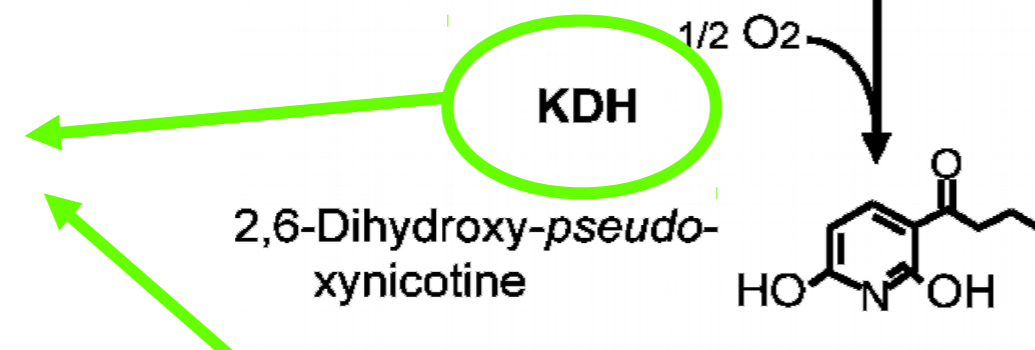
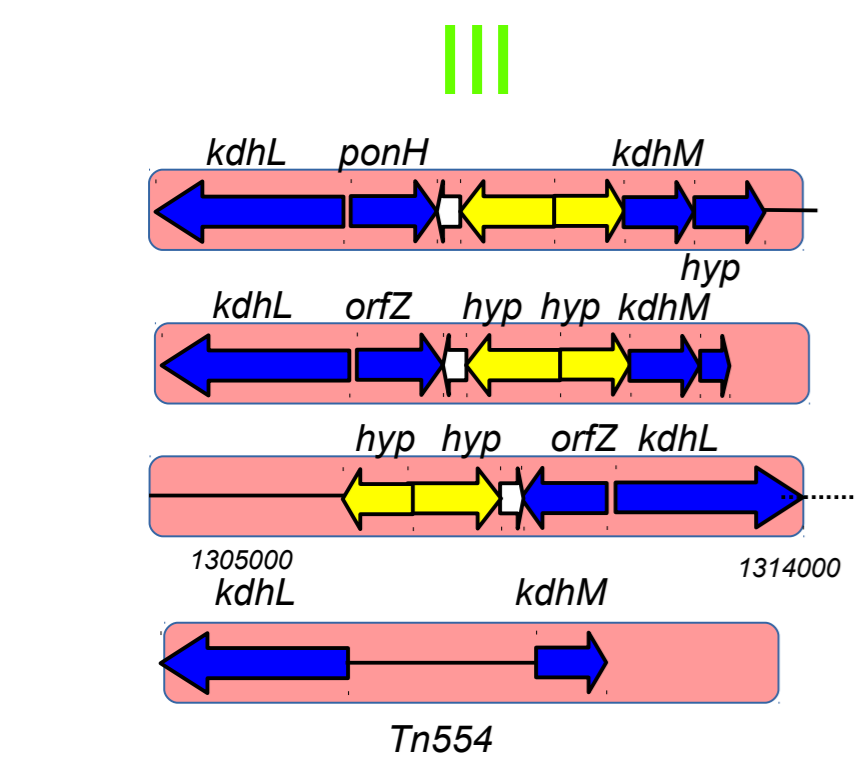
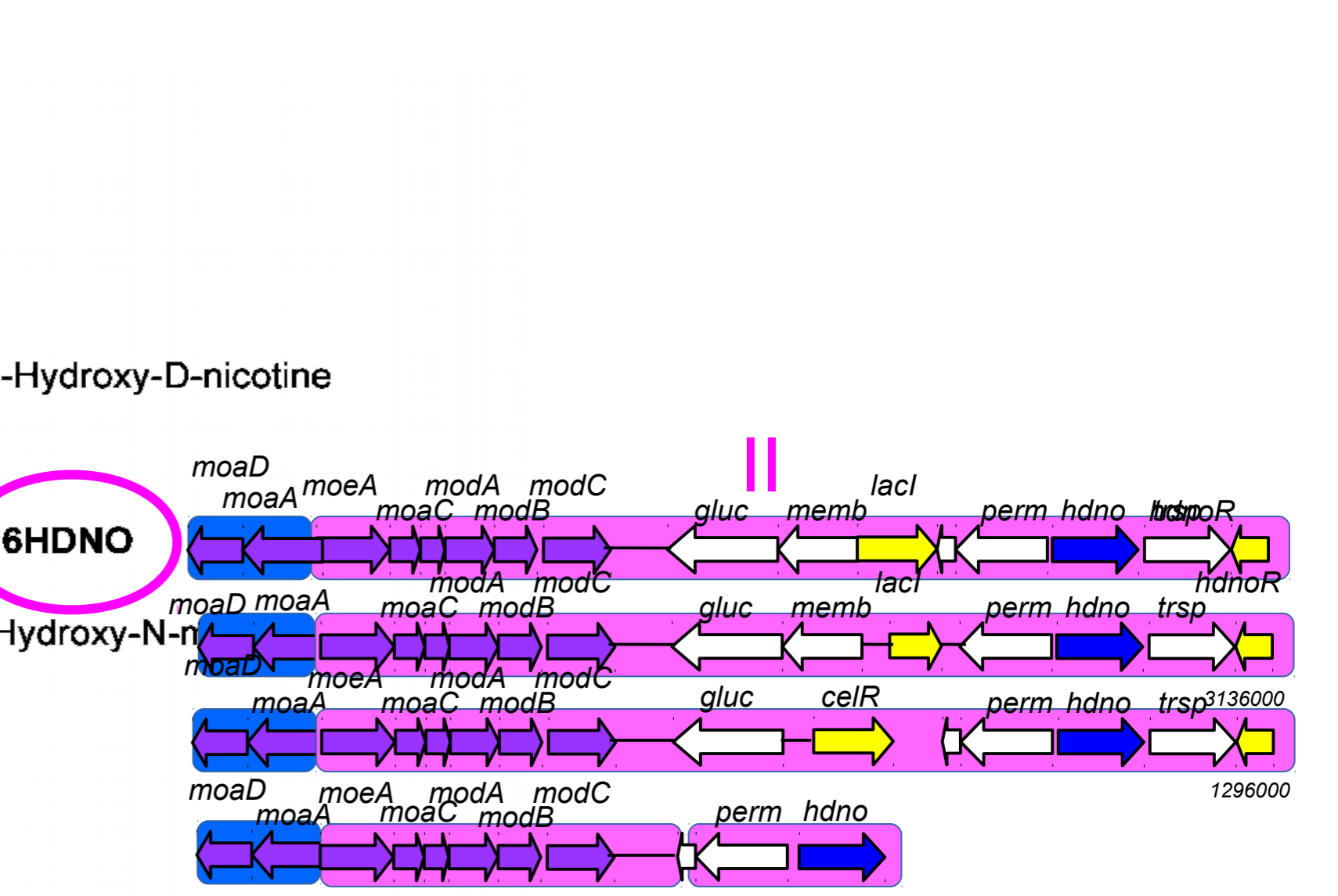
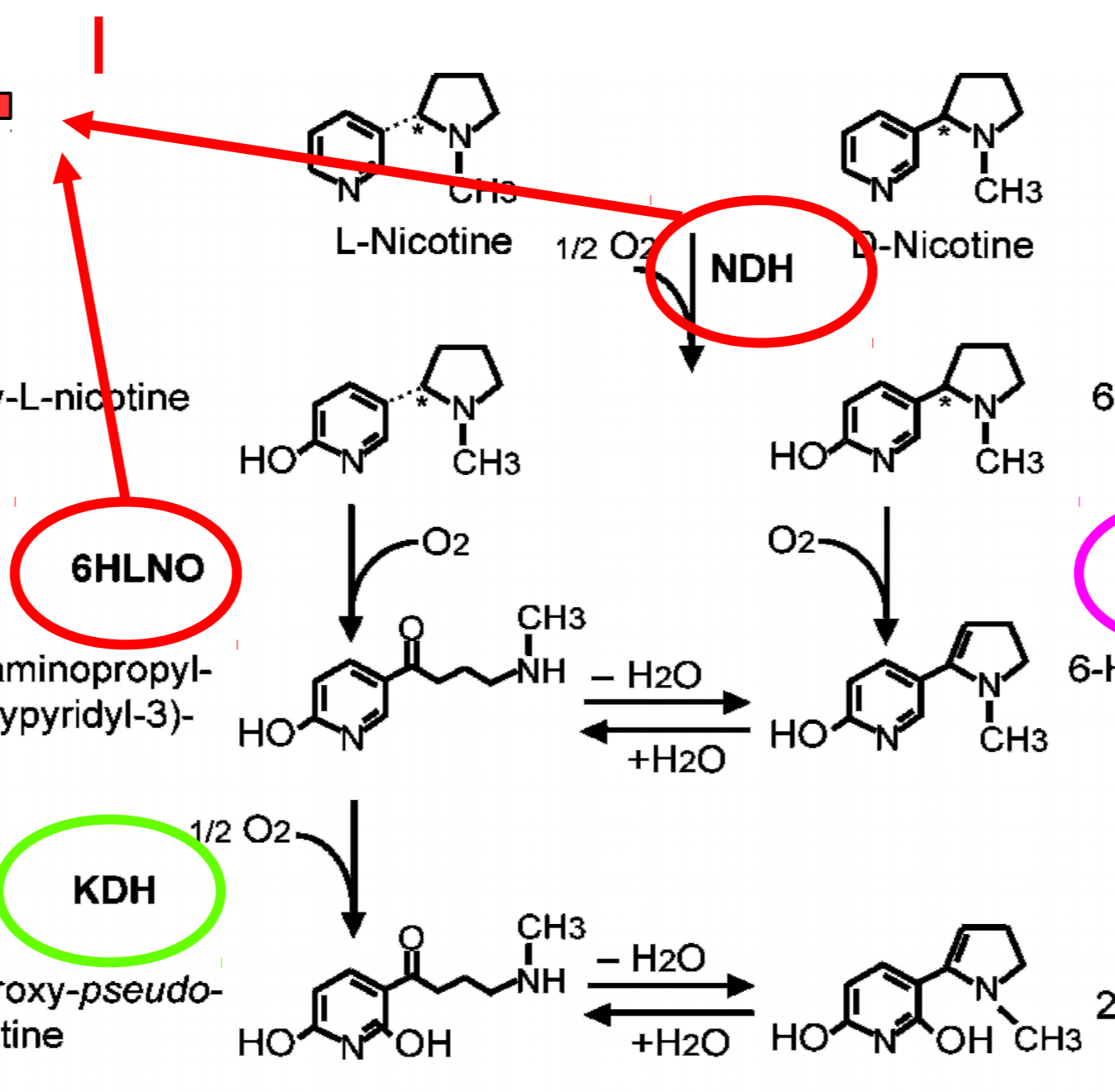
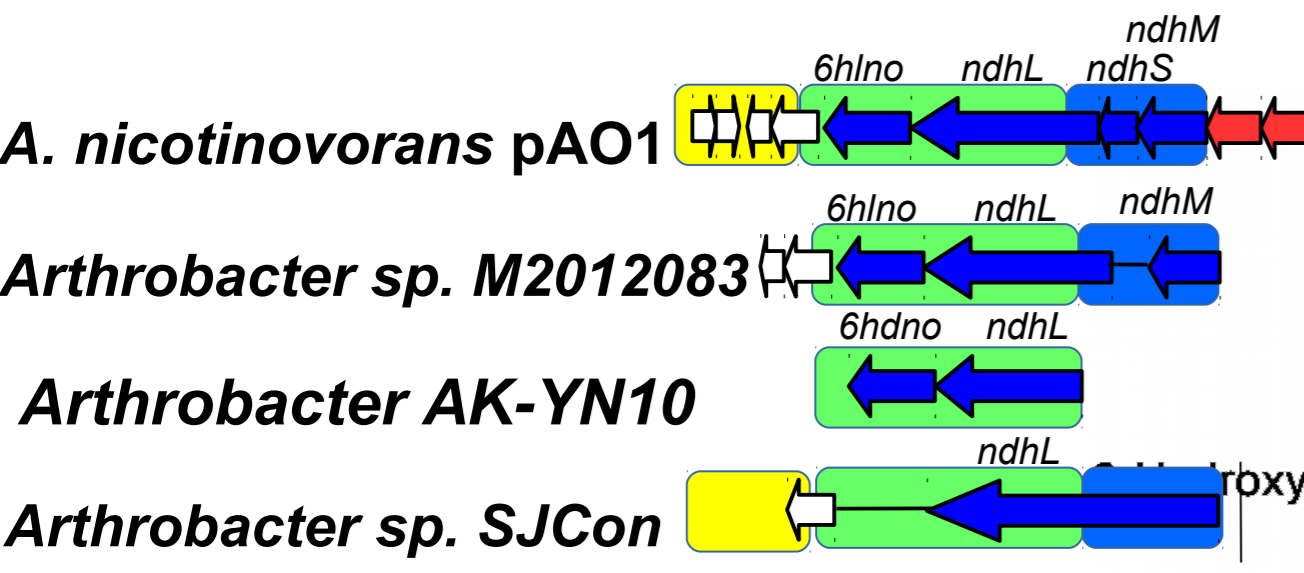


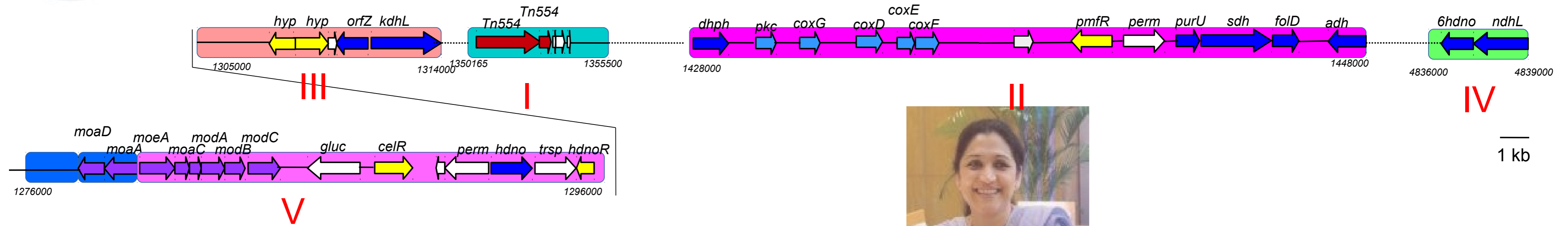
Arthrobacter AK-YN10



- ➔ ORFs involved in transposition
- ➔ ORFs involved in nicotine metabolism
- ➔ ORFs involved in cofactor biosynthesis
- ➔ ORFs with unknown function
- ➔ ORFs involved in expression regulation
- Identical colored blocks indicate LCBs

1 kb

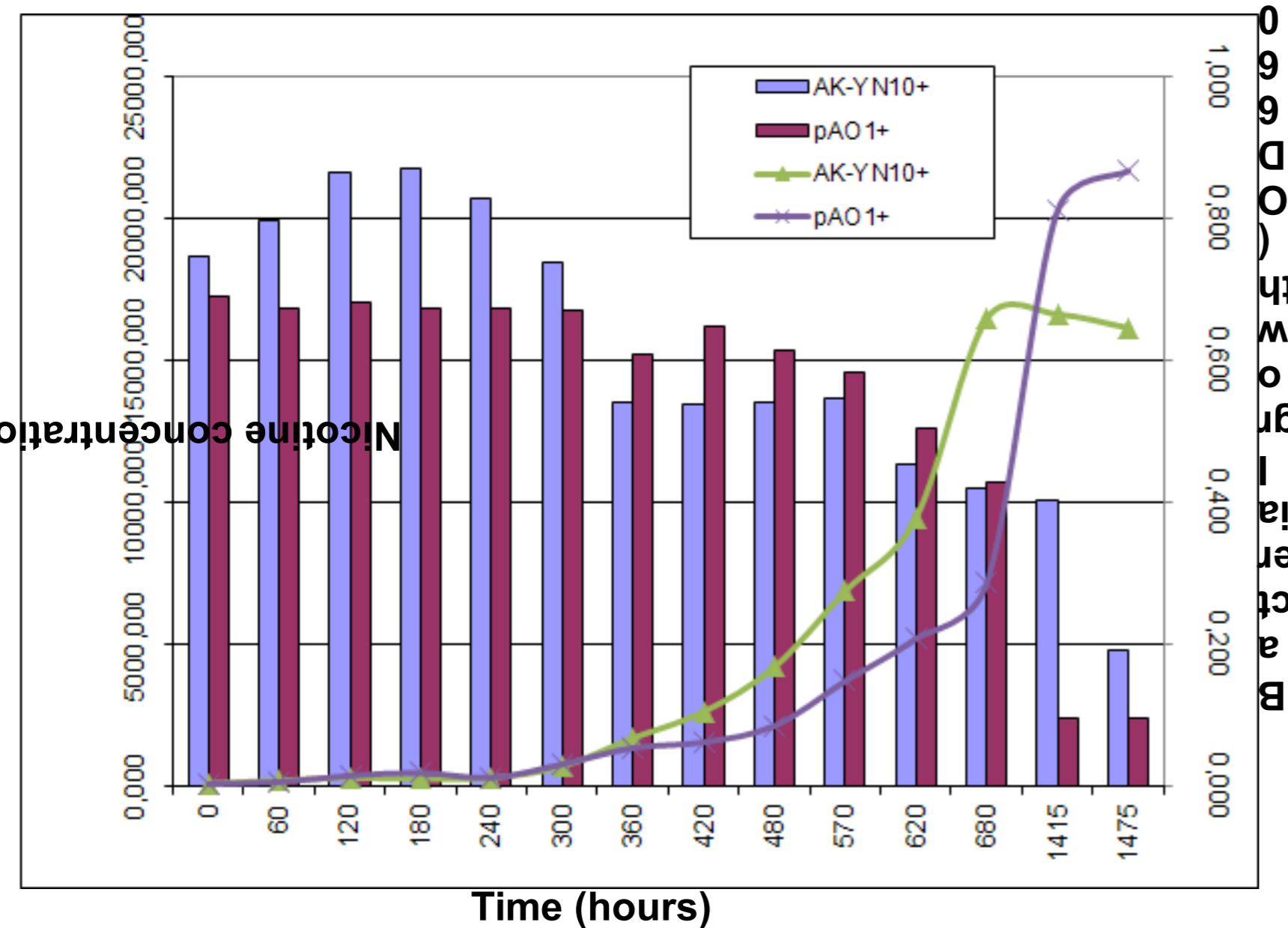




Arthrobacter AK-YN10 strain was a kind gift from Dr. Atya Kapley
National Environmental Engineering Research Institute, CSIR-NEERI, Nagpur, India

Nicotine resistance – 6 g/L for AK-YN10 vs: 1,5 g/L for *Pseudomonas sp. HF-1*

- 3 g/L for *Pseudomonas sp. Y22***
- 4 g/L for *Pseudomonas geniculata***
- 5 g/L for *Shinella sp. H2N7***
- 6 g/L for *Arthrobacter nicotinovorans pAO1***



Arthrobacter AK-YN10 can grow on nicotine containing medium, but does not for the blue pigment.

Conclusions:



- 1. Within the *Arthrobacter* genus, the *nic*-gene cluster is not singular to the pAO1 megaplasmid. Three strains (*Arthrobacter* sp. M2012083, *Arthrobacter* AK-YN10, *Arthrobacter* sp. SJCon) have been identified here as containing the *nic*-gene cluster.**
- 2. A modular design for the *nic*-gene cluster can be described, each module coding for a step in the nicotine catabolic pathway**
- 3. *Arthrobacter* sp. AK-YN10 can degrade nicotine, but the catabolic pathway might be slightly different compared to the pAO1 encoded pathway.**

Open questions:

- What is the evolutionary path that led to the *nic*-gene cluster? Did it formed within the *Arthrobacter* genus or was it acquired by horizontal gene transfer? Are the different modules evolving independently?**
- What is the origin of the pAO1 megaplasmid as the similarity with other plasmids is low?**

Funny facts:

- genomic data used as input (reference and draft genomes): 60 MB**
- number of generated files: approx: 5000**
- total amount of generated data (genomes and alignments): 32 GB**
- total computation time (assembly and alignments): about 52 hours**