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

Marius I. Mihăşan*

Laboratory of Biochemistry, Faculty of Biology, University “A. I. Cuza”, Iaşi

*marius.mihasan@uaic.ro*, Carol I Bvd., No 20 A, 700506, Iasi, Romania
**Nicotine-metabolism**

**Xylose-metabolism**

Roderich Brandeck

**Microbiology and biochemistry of nicotine degradation**

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

Marius Mihasan, Marius Stefan, Lucian Hritcu, Vlad Artemie, Roderich Brandeck.

**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)

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)

Contig assembly using MAUVE
- weight: 200
- HOXD Scoring Matrix:

Genome annotation using RAST v 2.0

Final genomes

Sequence similarity
BLAST, BRIG

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

<table>
<thead>
<tr>
<th>Assembled genomes</th>
<th>Arthrobacter sp. M2012083</th>
<th>Arthrobacter sp. SJConn</th>
<th>Arthrobacter sp. AK-YN10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference genomes</td>
<td>Arthrobacter aurescens TC1, Arthrobacter nitroguajacolicus Rue61a, Arthrobacter chlorophenolicus A6, Arthrobacter phanerothenia Sphe3, Arthrobacter chlorophenolicus A6, Arthrobacter aurescens TC1, Arthrobacter nitroguajacolicus Rue61a</td>
<td>Arthrobacter aurescens TC1, Arthrobacter nitroguajacolicus Rue61a, Arthrobacter chlorophenolicus A6, Arthrobacter phanerothenia Sphe3, Arthrobacter chlorophenolicus A6, Arthrobacter aurescens TC1, Arthrobacter nitroguajacolicus Rue61a</td>
<td>Arthrobacter aurescens TC1, Arthrobacter nitroguajacolicus Rue61a, Arthrobacter chlorophenolicus A6, Arthrobacter phanerothenia Sphe3, Arthrobacter chlorophenolicus A6, Arthrobacter aurescens TC1, Arthrobacter nitroguajacolicus Rue61a</td>
</tr>
</tbody>
</table>

| Number of Contigs: | 67 | 142 | 107 |
| Number of assembly bases: | 4629172 | 4389620 | 4839751 |
| Number of reference bases: | 4597686 | 4736495 | 4395537 |
| Number of LCBs: | 37 | 79 | 98 |
| Number of Blocks: | 89 | 222 | 173 |
| Breakpoint Distance: | 89 | 222 | 173 |
| DCJ Distance: | 81 | 213 | 164 |
| SCJ Distance: | 178 | 346 | 270 |
| Number of SNPs: | 437118 | 441536 | 601223 |
| Number of Gaps in Reference: | 3707 | 3793 | 8645 |
| Number of Gaps in Assembly: | 3583 | 3643 | 9053 |
| Total bases missed in reference: | 806998 | 846192 | 1831672 |
| Percent bases missed: | 17.55% | 17.87% | 41.67% |
| Total bases extra in assembly: | 601474 | 607774 | 2003703 |
| Percent bases extra: | 12.99% | 13.13% | 43.28% |
| Number of Inter-LCB Boundaries: | 6 | 14 | 9 |
| Contig N50: | 129641 | 61782 | 131717 |
| Contig N90: | 33265 | 18136 | 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 nicotine metabolism**  
**ORFs involved in expression regulation**  
**ORFs involved in cofactor biosynthesis**  
**ORFs with unknown function**

Identical colored blocks indicate LCBs

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

UNIVERSITATEA "ALEXANDRU IOAN CUZA" din IAŞI

www.uaic.ro
Nic-gene cluster modular design

A. nicotinovorans pAO1
Arthrobacter sp. M2012083
Arthrobacter AK-YN10
Arthrobacter sp. SJCon
Arthrobacter AK-YN10 can degrade nicotine

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 Pseuomonas 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

This work was supported by the strategic grant POSDRU/159/1.5/S/133652, co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007 – 2013