The GEMO project: Hitchhiking DNA in Magnaporthe oryzae

Ludovic Mallet, Cyprien Guerin Guérin, Joelle J. Amselem, Elisabeth E. Fournier, Helene H. Chiapello

The GEMO project: Hitchhiking DNA in *Magnaporthe oryzae*

MALLET L.[1,2], GUÉRIN C.[2], AMSELEM J.[3,4], FOURNIER E.[5] and CHIAPELLO H.[6]

[1] Evolutionary Bioinformatics University of Münster, DE
[2] INRA, MaIAGE Jouy-en-Josas, FR
[3] INRA, URGI Versailles, FR
[4] INRA, UMR BIOGER Thiverval-Grignon, FR
[5] INRA, UMR BGPI Montpellier, FR
[6] INRA, MIAT Castanet-Tolosan, FR

ANR 2009-GENM-029-01

Abstract

ludovic.mallet@wwu.de

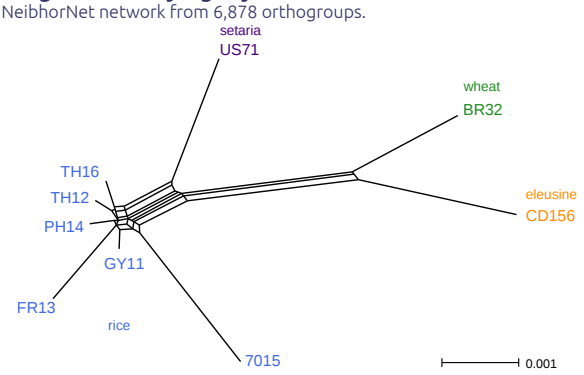
The Evolutionary Genomics of *Magnaporthe oryzae* (**GEMO**) **project** is an attempt to identify the genomic determinants and evolutionary events involved in **pathogenesis, host specificity and adaptation**. Ten closely related genomes of *M.oryzae* strains and one of the sister species *M. grisea* with different main host and host range were sequenced and analysed [1].

We focused here on the detection and analysis of potentially **horizontally acquired DNA regions** that we characterized using compositional methods.

We examined the general content, the **functional profiles** and the **potential donors** of the putative acquried genes. We reviewed the acquired regions and investigated a few target candidates for pathogenicity.

Figure 1: Phylogeny of strains.

NeibhorNet network from 6,878 orthogroups.



Methods

Horizontal Transfer detection [2]

- Tetranucleotide composition in sliding windows, 5kb long, 100bp step
- Kullback-Leibler divergence of composition: windows vs whole genome
- Parametric determination of threshold

Horizontal Transfer filtering

- To avoid false positive detection from known repeats, we filtered out:
- Transposable elements, annotated with REPET [3].
- rRNA, annotated with REPET base and blast.
- Homopolymers >100bp, Nannotator, PC
- Simple repeats, annotated with RepeatMasker [4]

Potential origin identification

- Inference with GOHTAM [1], a database of species composition
- Neighbor species with a distance over 160 arbitrary units were discarded

Functional profile

- Functional annotation was predicted with InterProScan [5]
- Analysis and presentation with DAVID [6], REVIGO [7] and Genes2WordCloud [8]

Potential origin

Class	# of regions	Species
Virus	3	Gorilla gorilla cytomegalovirus 2.1 (x2) ; Simian adenovirus 18
Fungi	6	Humicola grisea ; Colletotrichum higginsianum (x4) ; Chrysosporium lucknowense
Arthro-pods	3	Drosophila parabipectinata ; Lychas mucronatus ; Richardia teevani
Alguae	3	Pyropia yezoensis ; Acetabularia peniculus ; Cruoriaceae sp.
Choanof lagellate	2	Monosiga ovata

Table 2: Most credible donors

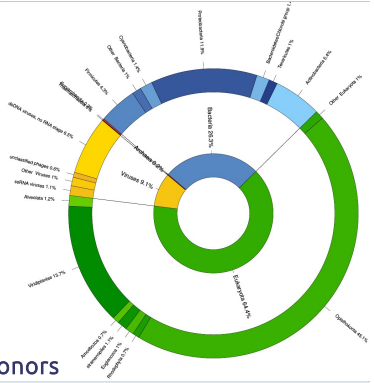
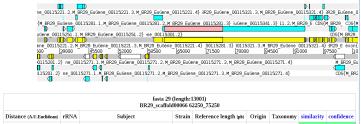


Figure 4: Potential donors

Candidate(s)

The putative transfered region is in red.

The composition of this region is very similar to the whole genome composition of another fungi: *Colletotrichum higginsianum*, *a* **plant pathogen**.



The gene is a repeat assemlby of 3 protein domains.

The condensation domain is found in many enzymes which synthesise peptide antibiotics. It catalyses a reaction to form peptide bonds in **non-ribosomal peptide biosynthesis**. We hypotesise an NRPS activity

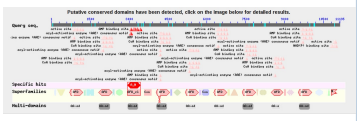


Table 1 : Genomes and statistics

Genomes	*70-15	† BR29	BR32	CD156	FR13	GY11	TH12	PH14	TH16	US71
Main host	rice	rice	wheat	Eleusine	rice	rice	rice	rice	rice	Setaria
Size (Mb)	40.9	40.9	41.9	42.7	43.1	46.3	48.5	49.8	39.1	41.2
Predicted genes	12,827	12,616	14,781	14,415	15,035	20,621	19,811	20,067	13,725	14,013
HT DNA (Mb)	1.22	1.23	2.59	1.55	0.14	0.37	0.67	0.52	0.47	1.26
HT DNA %	2.97	3.0	6.18	3.62	0.32	0.79	1.38	1.05	1.19	3.06
HT Genes	502	374	948	438	41	229	252	175	176	511

‡,* The reference strain *M.o.* 70-15 is sequenced with Sanger technology. † BR29 is a strain from the *M. grisea* species.

Figure 2: Functional profile

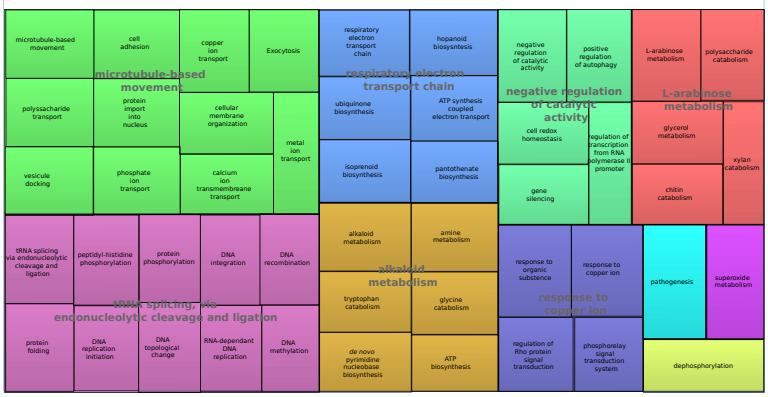
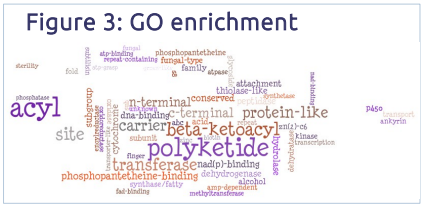


Figure 3: GO enrichment



References

[1] Chiapello *et al.*, 2015,GBE, in revision
[2] Ménigaud *et al.*, 2012, BMC Bioinfo.
[3] Flutre *et al.*, 2011, Plos One
[4] Smit, Hubley & Green. RepeatMasker
[5] Zdobnov *et al.*, 2001, Bioinformatics
[6]Huang *et al.*, 2009, Nature Protoc.
[7] Supek *et al.*, 2011, PloS One
[8] Baroukh *et al.*, 2011, Source Code Biol Med.

Conclusion & perspectives

The genomes of *Magnaporthe oryzae* contain between 0.14 (0.3%) to 3.59 (6.2%) of putative horizontal transfers, likely coming from **various kingdoms.** The functional profiling suggests their implication in various processes, including **pathogenicity, regulation and cell motion**. Early analyses yielded interesting candidates but further work is required for other functional classes. The strain-specific transfers will be analysed in the future to assess their implication in the variable host range of the strains. These investigations are expected to enhance the understanding of the evolutionary origins of the features driving pathogenicity phenotype and host specificity.

Acknowledgments

Funding: **ANR** 2009-GENM-029-01. Travel grant from the **SFBI**, many many thanks! Special thanks to Sophie Schbath and Philippe Leroy.