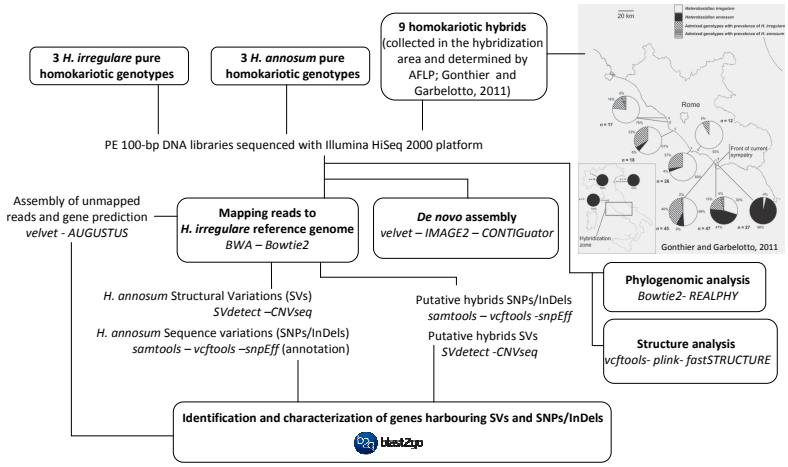
COMPARATIVE GENOMICS BETWEEN TWO ALLOPATRIC SPECIES AND THEIR HYBRIDS PROVIDE INSIGHTS ON THE MECHANISMS OF INVASION OF FUNGAL PLANT PATHOGENS

Fabiano Sillo[1], Matteo Garbelotto[2] and Paolo Gonthier[1]

[1]*University of Torino, Department of Agricultural, Forest and Food Sciences Largo Braccini 2, I-10095 Grugliasco, Italy; e-mail: paolo.gonthier@unito.it*
[2]*University of California at Berkeley, Department of Environmental Science Policy and Management, 151 Hilgard Hall, CA 94720 Berkeley*

Heterobasidion irregulare is a major fungal pathogen of pines in North America and was accidentally introduced into Italy within the natural range of its allopatrically diverged sibling species *H. annosum*, becoming invasive. Phenotypic observations have shown that the invasive species is outcompeting the native one, especially in saprobic and sporulation potential. Moreover, the two species retained high levels of interfertility, leading to massive hybridization between the two and the generation of hybrid swarms in the sympatric area. Here, we report on a comparative genomic study aimed at 1) **elucidating the genomic structure of sibling taxa**, with similar biology and host preference, that underwent speciation in allopatry, 2) **identifying the genomic traits providing the advantage the invasive species *H. irregulare* has over the native one** and 3) **determining the mechanisms underlying the current massive hybridization between them**. In order to reach these aims, a whole-genome sequencing of three pure genotypes of *H. irregulare*, three of *H. annosum* and nine genotypes previously identified as hybrids through Amplified Fragment Length Polymorphism (AFLP) analysis was performed.

Materials and Methods



Results - Comparative genomics between *H. irregulare* and *H. annosum*

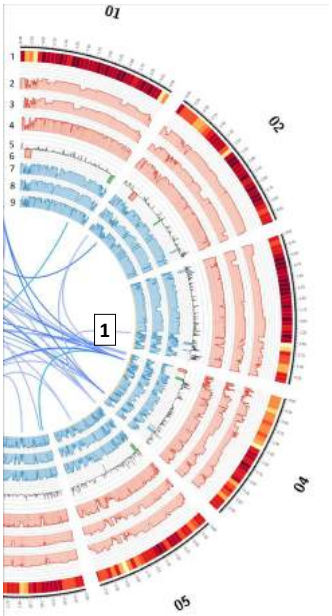


Fig.1. Visualization of reads mapping of the six pure pure genotypes sequenced. The scaffolds corresponding to chromosomes of reference *H. irregulare* TC32-1 are circled and defined. Gene density (1) in a window size of 20K bp is colour coded from yellow to dark red, with deeper red region representing higher gene density. From outer to inner, pink circles represents read depth of *H. irregulare* genotypes, (2,3,4).
Density of consensus SNPs of three *H. annosum* isolates is visualized as line plot in (5). Putative *H. annosum* SVs are represented in (6) and colour coded as follow: red for large deletions, green for large insertions, blue for inversions. Each blue line in the central part represented a putative *H. annosum* translocation event. Read depth of *H. annosum* genotypes was plotted in circles (7,8,9).

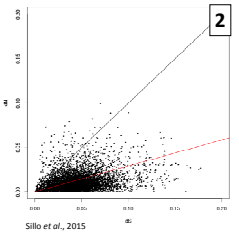
- Comparative genomics analysis among the pure genotypes identified a significant macrosynteny between the genomes of the two species, as well as **inter-chromosomal translocations**, **inversions** and duplications (**Copy Number Variations** – CNVs) differentiating the two pathogens.

- Unmapped reads of *H. annosum* genotypes (25% of the total reads) were assembled in 1,104 contigs. *AUGUSTUS* predicted 249 putative CDSs. After the BLAST2GO analysis, sequences were annotated as **related to retroelements.**

-707 genes were found to harbor more than 40 putative SNPs/InDels per kilobase and were regarded as *H. annosum* **species specific alleles**. Main GO terms assigned to these genes were related to heme binding and oxidative-reduction processes. Wood decay fungi as *Heterobasidion* produce large amounts of oxidoreductases putatively involved in lignin oxidation and conversion, as well as heme-containing proteins, such as manganese and lignin peroxidases.

-In 2012, a clear trade-off between *H. irregulare* gene families expressed during pathogenesis, saprobic wood decay, and fruiting was reported (Olson *et al.,* 2012). In the comparative analysis, **genes involved in pathogenicity appeared to significantly harbour a lower number of non-synonymous mutations** between the two species compared to genes involved in **saprobic growth and sporulation**. This finding provided genomic evidence that differences in *fitness* are more likely to be determined by these two last functions, as previously documented by in vitro experiments (Giordano *et al.*, 2014).

-The panel of detected putative *H. annosum* SNPs was annotated in homologous *H. irregulare* gene models and dN/dS was calculated for each homologous compared pairwise. A large fraction of genes under divergent positive selection (dN/dS>1) was described as involved in **transcriptional** functions and **mitochondrial-related** factors.



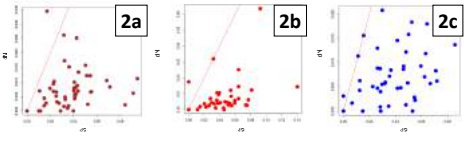
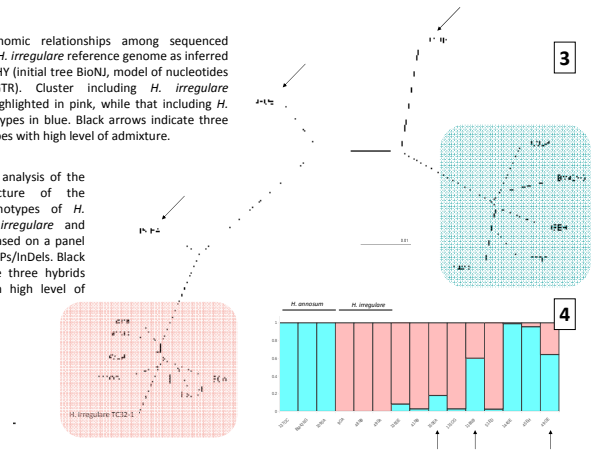


Fig 2. dN/dS distribution among homolog pairs of *H. irregulare* and *H. annosum*. The mean dN/dS value is 0.26 and is indicated by a red line. Black line indicates the neutral expectation where dN = dS. 153 sequences fall above the black line. dN/dS distribution is also measured for genes related to saprobicity (mean dN/dS = 0.22) (2a), to pathogenicity (mean dN/dS = 0.16) (2b), and to sporulation (mean dN/dS = 0.21) (2c) .

Results - Genomics of natural hybrids

Fig.3. Phylogenomic relationships among sequenced genotypes and *H. irregulare* reference genome as inferred through REALPHY (initial tree BioNJ, model of nucleotides substitution GTR). Cluster including *H. irregulare* genotypes is highlighted in pink, while that including *H. annosum* genotypes in blue. Black arrows indicate three hybrids genotypes with high level of admixture.

Fig.4. Bayesian analysis of the genomic structure of the sequenced genotypes of *H. annosum*, *H. irregulare* and their hybrids based on a panel of 1,608,098 SNPs/InDels. Black arrows indicate three hybrids genotypes with high level of admixture.



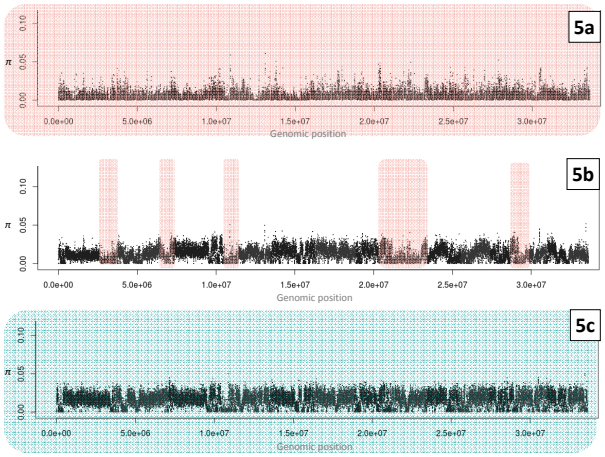


Fig.5. Nucleotide diversity (π) calculated per *site* across the whole genome in 10K bp sliding windows. In (5a), π is calculated by comparing *H. irregulare* genomes *vs* the reference; in (5b) π is calculated by comparing *H. annosum* vs *H. irregulare* genomes, including the reference; in (5c) π is calculated by comparing hybrid *H. annosum* vs *H. irregulare* genomes, including the reference. In hybrids, regions with abnormal level of π putatively belonging to *H. irregulare* (introgressed regions) are highlighted in pink (5b).

Conclusions

Comparative genomics between *H. irregulare* and *H. annosum* (Sillo *et al.*, 2015)
-The genomic structure of *H. irregulare* and *H. annosum* has remained similar

-**Nucleotide interspecific identity reached 98%** in several mapped genomic regions.

-Several SVs including **inversions**, **translocations**, and **CNVs** have played a prominent role in the creation of genomic islands leading to diversification of the two species.

-**Proliferation of transposable elements** have shaped the *H. annosum* genome, probably leading to a size-increasing.

-Adaptive evolution might have remodeled the **transcriptional machinery** and the **mitochondrial-related pathways** during differentiation of the two species.

-Genes encoding **heme binding proteins** and **oxidoreductases**, probably related to saprobic capacity, exclusively found in regions showing high level of sequence variation might also have been influenced by adaptive processes.

-Genes involved in **sporulation** and especially in **saprobic growth** appeared to harbour higher level of non-synonymous mutations, compared with genes involved in **pathogenicity, for which purifying selection** has probably acted.

Genomics of hybrids
-Phylogenomic and STRUCTURE analysis allowed to identify three hybrids with **high level of admixture.**

-Analysis of nucleotide diversity allowed to detect the **genomic regions putatively introgressed from one species to the other.**

-Genes in these regions will be investigated in order to understand which traits were selected by the adaptive processes.

References
-Giordano, L., Gonthier, P., Lione, G., Capretti, P., Garbelotto, M. (2014). The saprobic and fruiting abilities of the exotic forest pathogen *Heterobasidion irregulare* may explain its invasiveness. *Biological invasions*, **16**(4), 803-814.
-Gonthier, P., Garbelotto, M. (2011). Amplified fragment length polymorphism and sequence analyses reveal massive gene introgression from the European fungal pathogen *Heterobasidion annosum* into its introduced congener *H. irregulare*. *Molecular Ecology*, **20**(13), 2756-2770.
-Olson, Å., Aerts, A., Asiegbu, F., Belbahri, L., Bouzid, O., Broberg, A., *et al.* (2012). Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. *New Phytologist*, **194**(4), 1001-1013.
-Sillo, F., Garbelotto, M., Friedman, M., Gonthier, P. (2015). Comparative genomics of sibling fungal pathogenic taxa identifies adaptive evolution without divergence in pathogenicity genes or genomic structure. *Genome biology and evolution*, **7**(12), 3190-3206.

This work was supported by the Italian Ministry of Education, University and Research, within the FIRB program (grant number RBFR1280NN).