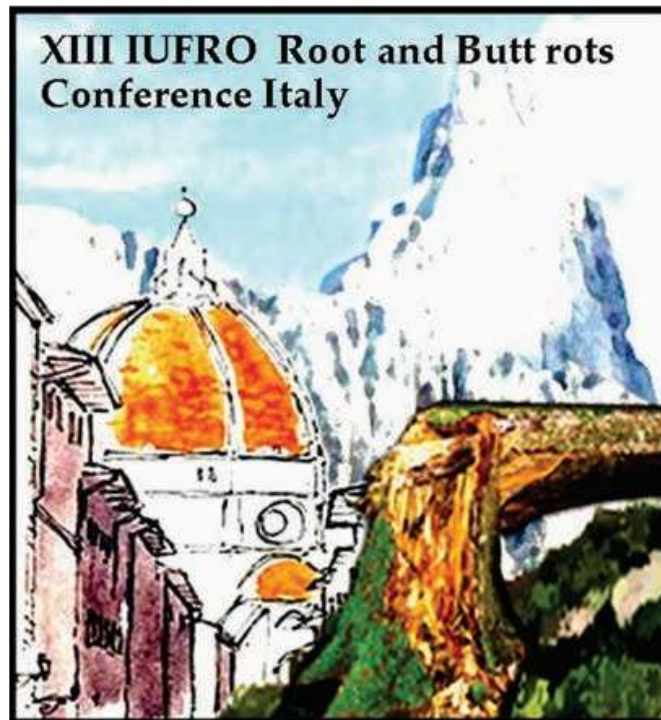


Proceedings e report

93



**XIII Conference
"Root and Butt Rot of Forest Trees"
IUFRO Working Party 7.02.01**

**September 4th – 10th 2011
Firenze – Auditorium di S. Apollonia
S. Martino di Castrozza (TN) – Palazzo Sass Maor, Italy**

XIII Conference Root and Butt Rot of Forest Trees
IUFRO Working Party 7.02.01 / edited by P. Capretti, C.
Comparini, M. Garbelotto, N. La Porta, A. Santini. –
Firenze : Firenze University Press, 2013.
(Proceedings e report; 93)

<http://digital.casalini.it/9788866553533>

ISBN 978-88-6655-352-6 (print)
ISBN 978-88-6655-353-3 (online)

Progetto grafico di Alberto Pizarro Fernández, Pagina Maestra snc

This book was printed with contribution of Fondazione Edmund Mach and Italian Ministry of Education,
University and Research, PRIN program number 2008SBCC9S
Cover photos and photos on pages 1, 65, 77, 115, 141, 173, 233 by Kari Korhonen

Peer Review Process

All publications are submitted to an external refereeing process under the responsibility of the FUP Editorial Board and the Scientific Committees of the individual series. The works published in the FUP catalogue are evaluated and approved by the Editorial Board of the publishing house. For a more detailed description of the refereeing process we refer to the official documents published in the online catalogue of the FUP (<http://www.fupress.com>).

Firenze University Press Editorial Board

G. Nigro (Co-ordinator), M.T. Bartoli, M. Boddi, R. Casalbuoni, C. Ciappei, R. Del Punta, A. Dolfi, V. Fargion, S. Ferrone, M. Garzaniti, P. Guarnieri, A. Mariani, M. Marini, A. Novelli, M. Verga, A. Zorzi.

© 2013 Firenze University Press
Università degli Studi di Firenze
Firenze University Press
Borgo Albizi, 28, 50122 Firenze, Italy
<http://www.fupress.com/>
Printed in Italy

Edited by

Paolo Capretti, Cecilia Comparini, Matteo Garbelotto, Nicola La Porta, Alberto Santini.

Paolo Capretti

Università di Firenze

Dipartimento di Scienze delle Produzioni Agroalimentari e dell'Ambiente

Piazzale delle Cascine 28, 50144 Firenze, Italy

paolo.capretti@unifi.it

Cecilia Comparini

Università di Firenze

Dipartimento di Scienze delle Produzioni Agroalimentari e dell'Ambiente

Piazzale delle Cascine 28, 50144 Firenze, Italy

cecilia.comparini@unifi.it

Matteo Garbelotto

University of California, Berkeley

Department of Environmental Science, Policy and Management – Ecosystem Science

137 Mulford Hall, 3114, Berkeley, CA, USA, 94720

matteog@berkeley.edu

Nicola La Porta

IASMA Centro per la Ricerca e l'Innovazione, Fondazione Edmund Mach

Dipartimento Agrosistemi Sostenibili e Biorisorse

Via E. Mach 1, 38010, S. Michele all'Adige (TN), Italy

nicola.laporta@iasma.it

Alberto Santini

Consiglio Nazionale delle Ricerche – CNR – Istituto per la Protezione delle Piante

Via Madonna del Piano 10, 50019, Sesto Fiorentino (FI), Italy

a.santini@ipp.cnr.it

To cite this book:

Capretti P., Comparini C., Garbelotto M., La Porta N., Santini A., (eds.). Proceeding of the XIII International Conference on Root and Butt Root of Forest Trees. Firenze (FI) – S. Martino di Castrozza (TN), Italy, 4th – 10th September 2012. University Press, Firenze, 2012. 280 pp.

Multivariate analysis revealed translational selection and mutational bias in *Heterobasidion irregulare*-destructive fungal pathogen of conifers in the Boreal hemisphere

G. Sablok¹, K.C. Nayak², E. Potenza³, G. Emiliani⁴, N. La Porta¹

¹*Sustainable Agro-ecosystems and Bioresources Department, IASMA Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, (TN), Italy.*

²*Bioinformatics Centre, Institute of Life Sciences, Department of Biotechnology, Nalco Square, Bhubaneswar, India.*

³*Genomics and Biology of Fruit Crops Department, IASMA Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, (TN), Italy.*

⁴*Trees and Timber Institute, National Research Council, IVaLSA-CNR S. Michele all'Adige (Trento), Italy.*

Corresponding author e-mail address: nicola.laporta@fmac.it

Abstract. Codon usage has a profound effect at the intra and inter- proteome level and even at the organismal level. *Heterobasidion annosum s.l.* is one of the most destructive fungal pathogen of conifers in the Boreal hemisphere. Recently the whole genome sequence of the *H. annosum* was released at DOE Joint Genome Institute (DOE-JGI) with 8.23 X coverage (Olson *et al.*, 2009) covering nuclear genome assemblies in 39 scaffolds of total 33.7 Mbp estimated to cover 98.1% of the complete genome. We have carried out a genome wide codon usage analysis of this pathogenic fungus using multivariate analysis and popular indices of codon usage. The results show that the G+C contents at three positions of codon are different GC1 (mean value of 58.37% ± 0.07); GC2 (mean value of 47.5% ± 0.09); GC3 (mean value of 60.9% ± 0.13) explain heterogeneity in the composition of the genes. For the identification of gene expression values we have systematically searched for the ribosomal proteins and have used them as a reference for calculating codon adaptation index (CAI). The heterogeneity in the genome was also revealed by plotting coordinates of *H. annosum* genes on axis 1 (COA/RSCU) which showed high negative correlation with GC_{3s} and GC content ($r=-0.965$, $P<0.01$ and $r=-0.858$, $P<0.01$), and significant high positive correlation with Nc ($r=0.689$, $P<0.01$) Furthermore, high significant negative correlation was observed between Nc and GC_{3s} ($r=-0.674$, $P<0.01$) which also complements the observation that C-ending codons are preferred over G-ending codons in highly expressed genes. It was also observed that CAI value and GC_{3s} also had a significant correlation ($r=0.789$, $P<0.01$) which suggest that genes with higher expression level tend to use C or G at synonymous positions compared to genes with lower expression level. We further analyzed the correlation between the nucleotide bias and amino acid composition using the two other phylogenetically closed fungal genomes.

Codon usage signature implies the variation in the usage of the codon at the wobble position. This variation in the usage has a profound effect at the intra- and inter-proteome level. It has widely demonstrated that except methionine (Met) and tryptophan (Trp), most of the amino acids are biased towards codon degeneracy (Hershberg and Petrov, 2009). *Heterobasidion irregulare s.l.* is one of the most destructive necrotrophic fungal pathogen of conifers in the Boreal hemisphere that produces a range of extracellular enzymes and a multitude of toxins. Most conifer

trees are susceptible to infection by this basidiomycete, causative agent of a root and butt rot disease and widely regarded as the most economically important forest pathogen in temperate forests devastating conifer plantations and natural forests.

H. irregulare complex has a wide geographical distribution particularly in many parts of Europe, North America, China and Japan (Dai *et al.*, 1999; Dai *et al.*, 2003). Besides *H. irregulare*, the other species are known from East Asia, Australia and adjacent areas (Niemelä *et al.*, 1998). Currently, eight distinct taxonomic species have been described within the genus.

The virulence of this pathogen has been shown to be partly under mitochondrial control. Recent progress in the *H. irregulare* research include: possibilities for fruiting and classical genetics, publication of a genetic linkage map, expressed sequence tags (ESTs) compilation of more than 3,000 sequences are available, transformation system, knowledge of genetics of interspecific recognition, phylogeny of the species complex, pathogenicity factors in nucleus and mitochondria. The complete sequence of *H. irregulare* is the first plant pathogenic homobasidiomycete with comprehensive genome coverage; and this is challenging for investigations in many areas, including pathogenicity factors, interactions with host organisms, lignin degradation and bioremediation applications, and fungal biology and evolution. In this study, we used the available complete genome sequence of this organism and analysed its codon usage, aiming to understand the genetic organization of the *H. irregulare* genome.

The complete genome sequence and coding sequences of *H. irregular* was obtained from JGI (<http://genome.jgi-psf.org>). Sequences were initially checked for sequencing errors and finally 6,497 CDS sequences of *H.irregulare* having more than 100 codons were extracted directly to avoid sampling bias in calculations of codon usage (Wright, 1990). All the sequencing errors were removed using in-house developed perl/Bio-perl scripts and BioSeq modules have been used to prepare the final sequence dataset. All the popular indices of codon usage (Nc, RSCU) have been used to estimate the levels of biased usage of the codon. All the statistical analysis have been carried out using SPSS 17.0 (<http://www.spss.com/>) , Origin professional 8.0 (<http://www.originlab.com/>) and R (<http://www.r-project.org/>). All correlations are based on the Nonparametric Spearman's Rank Correlation analysis method.

Genome of *H. irregulare* is GC rich (52.0%), which predicts the biased usage of G- /or C- ending codons in the coding regions of this genome and the same was observed. Since there were wide variations in the GC bins of the genome, therefore to evaluate GC variation, G+C content was calculated at all the three codon positions. The results showed that the G+C contents at three positions of codon differ significantly, GC1 (58.37% ± 0.07); GC2 (47.5% ± 0.09); GC3 (60.9% ± 0.13), which explain occurrence of heterogeneity in the composition of the genes. To evaluate the compositional constraints, Nc plot was further analyzed. Nc plot revealed that a high number of the genes are lying below the expected curve and

poised towards GC3s, which clearly demonstrates that besides the compositional bias, there are several other factors also influencing the biased patterns of codon usage in this genome.

To further investigate the constraints on the codon usage biology we carried out correspondence analysis of the RSCU in a 59-dimensional hyperspace and the data was partitioned and analyzed across the first four major axes. We observed that the first major axis (axis1/COA; 20.4%) revealed the major inertia of 59-dimensional hyperspace and the subsequent axes/COA (2, 6.02%; 3, 4.7%; and 4, 3.5%) showed decreased variation subsequently.

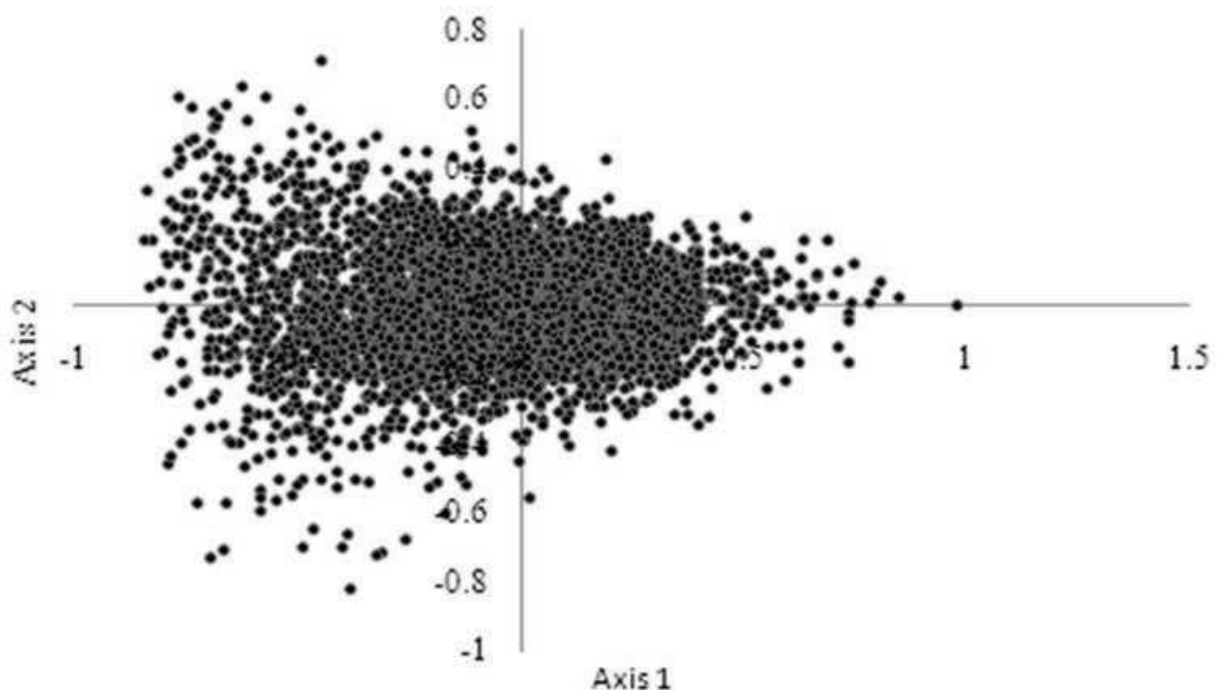


Figure 1. Plot for coordinates of first two major axes of correspondence analysis for *Heterobasidion irregulare*

To evaluate the levels of gene expression, codon adaptation index (CAI) was calculated; ribosomal proteins were systematically searched and used as reference sequences for calculating CAI (Gupta *et al.*, 2004). To evaluate the effect of gene expression level on synonymous codon usage bias spearman's rank correlation analysis was carried out between CAI and the coordinates of genes along the axis1/COA. A high negative correlation between these two variables ($r=-0.812$, $P<0.01$) was observed whereas high significant positive correlation was found between CAI and GC_{3s}, GC ($r=0.789$, $P<0.01$; $r=0.682$, $P<0.01$). These results are in complete agreement with the finding that the genes with higher expression level exhibit a greater degree of codon usage bias towards GC ending codons and prefer to use codons with C or G at the synonymous position as compared to weakly

expressed genes. In summary, this study has shown that the codon usage variation among the genes of *H. irregulare* is influenced by mutational bias and translational selection. As more complete genomes of this phylogenetic taxa being sequenced, different factors shaping the pattern of codon usage might be found.

References

- Dai Y.C., Korhonen K., 1999. *Heterobasidion annosum* group S identified in northeastern China. *European Journal of Forest Pathology* 29: 273-279.
- Dai Y.C., Vainio E.J., Hantula J., Niemelä T., Korhonen K., 2003. Investigations on the *Heterobasidion annosum* complex in central and eastern Asia with the aid of mating tests and DNA fingerprinting. *Forest Pathology*. 33: 269-286.
- Gupta S.K., Bhattacharyya T.K., Ghosh T.C., 2004. Synonymous codon usage in *Lactococcus lactis*: mutational bias versus translational selection. *Journal of Biomolecular Structure and Dynamics* 21: 527-536.
- Hershberg R., Petrov D.A., 2009. General Rules for Optimal Codon Choice. *PLoS Genet* 5: e1000556.
- Niemelä T., Korhonen K., 1998. Taxonomy of the genus *Heterobasidion*. In: Woodward S., Stenlid J., Karjalainen R., Hüttermann A. (eds.). *Heterobasidion annosum* Biology, Ecology, Impact and Control. CAB International, Wallingford, Oxon, OX10 8DE, UK, pp. 27-33.
- Wright F., 1990. The 'effective number of codons' used in a gene. *Gene* 87: 23-29.