Allele Intersection Analysis: a novel tool for multi locus sequence assignment in multiply infected hosts

a story about *Wolbachia*, cloning, sequencing and set theory written by

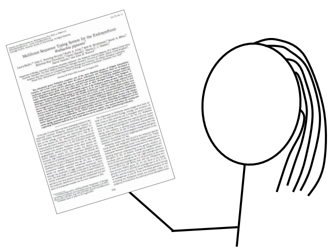
Wolfgang Arthofer[1], Markus Riegler[2], Karl Moder[3], Daniela Schneider[4], Wolfgang J Miller[4] and Christian Stauffer[5]

[1] Molecular Ecology Group, Institute of Ecology, University of Innsbruck, Austria
[2] Centre for Plants and the Environment, University of Western Sydney, Australia
[3] Institute of Applied Statistics and Computing, BOKU, University of Natural Resources and Applied Life Sciences, Vienna, Austria
[4] Centre of Anatomy and Cell Biology, Medical University of Vienna, Austria
[5] Institute of Forest Entomology, BOKU, University of Natural Resources and Applied Life Sciences, Vienna, Austria

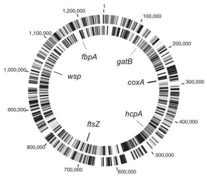
Correspondence: wolfgang.arthofer@uibk.ac.at http://www.uibk.ac.at/ecology/ http://www.peerart.at/aw

Back in 2006 ...

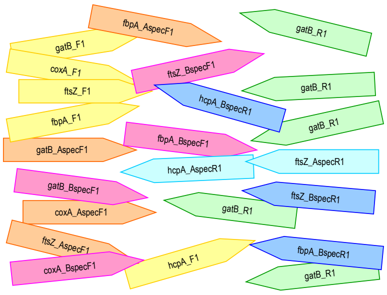
... Baldo et al. published a Multi Locus Sequence Typing (MLST) system for the endosymbiont *Wolbachia*.

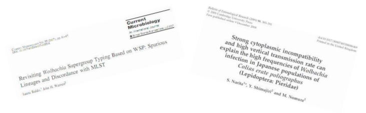


The MLST uses sequence information of 5 genes to unambigously define a *Wolbachia* strain. All information can be found in a web database.



To amplify the genes, universal and supergroup specific primers were introduced.

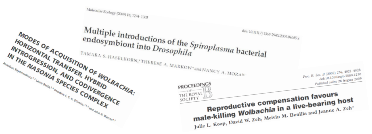




Current Microbiology Vol. 55 (2007), pp. 81–87
Current Microbiology An International Journal
Revisiting *Wolbachia* Supergroup Typing Based on WSP: Spurious Lineages and Discordance with MLST
Laura Baldo, John H. Werren

Bulletin of Entomological Research (2009) 99, 385–391
© 2008 Cambridge University Press
First published online 9 December 2008
Strong cytoplasmic incompatibility and high vertical transmission rate can explain the high frequencies of *Wolbachia* infection in Japanese populations of *Colias erate poliographus* (Lepidoptera: Pieridae)
S. Narita, Y. Shimajiri and M. Nomura

Today, MLST is the standard for describing a *Wolbachia* strain!

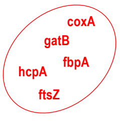


Molecular Ecology (2009) 18, 1294–1305
Multiple introductions of the *Spiroplasma* bacterial endosymbiont into *Drosophila*
TAMARA S. HASELKORN, THERESE A. MARKOW and NANCY A. MORAN

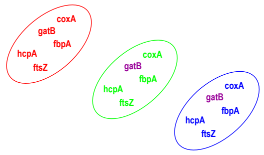
MODES OF ACQUISITION OF WOLBACHIA: HORIZONTAL TRANSFER, HYBRID INTROGRESSION, AND CODIVERGENCE IN THE NASONIA SPECIES COMPLEX
Rhitoban Raychoudhury, Laura Baldo, Deodoro C. S. G. Oliveira and John H. Werren

PROCEEDINGS OF THE ROYAL SOCIETY B
Proc. R. Soc. B (2009) 276, 4021–4028
doi:10.1098/rspb.2009.1230
Published online 26 August 2009
Reproductive compensation favours male-killing *Wolbachia* in a live-bearing host
Julie L. Koop, David W. Zeh, Melvin M. Bonilla and Jeanne A. Zeh

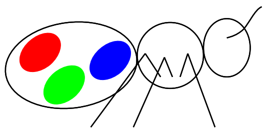
Consider a *Wolbachia* strain as a set containing MLST sequences as elements. And lets name this strain here simply ‚red'.



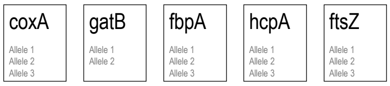
Different strains contain different elements – but maybe they even share some: here the blue and the green *Wolbachia* have the same gatB! All sequences of a distinct strain give the MLST sequence type, a unique strain identifier.



A multiply infected insect is again a set, containing different strains as elements. We call the taxative list of strains infecting one individual the ‚infection type'. The infection type of this insect is {red, blue, green}.



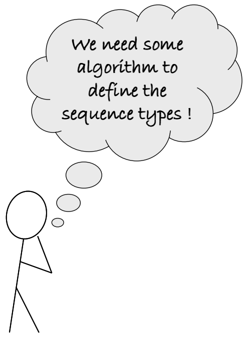
It's easy to clone and sequence the MLST amplicons of this insect ...

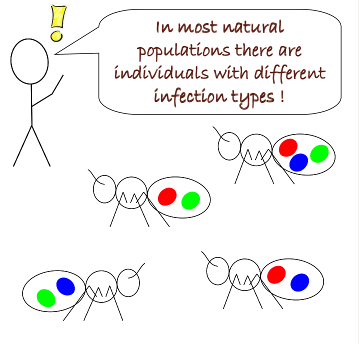


... and we know now that the insect was triple infected. Unfortunately, by cloning the alleles lost their ‚color' – we do not know which alleles belong together.

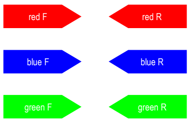
If there would be only one strain from the A anb B supergroup, specific primers would fix the problem ...

... but the cherry fruit fly, for instance, harbours 5 *Wolbachia* strains, and 3 of them are A-group !



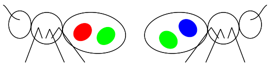

!
In most natural populations there are individuals with different infection types !

Let's build a diagnostic system that can easily identify an individual's infection type:

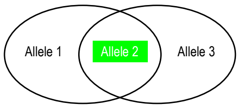


Strain specific PCR primers can do this job, targeting a highly variable gene – wsp could be a candidate, or a single copy VNTR, or ....

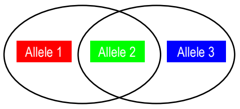
Now we search for two individuals that share only one *Wolbachia* strain.



We amplify, clone and sequence one MLST gene, and compare the alignments.



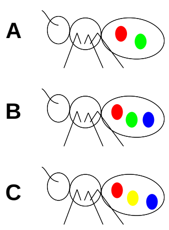
The allele found in both individuals must belong to the green strain !



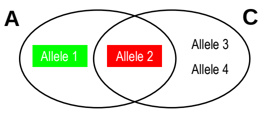
Furthermore, it's obvious that allele 1 comes from the red strain, and allele 3 from the blue one !

Any combination of individual infection types that allows the assignment of all strains is called ‚informative'.

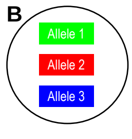
Look at this informative combination in a 4-fold infected species:



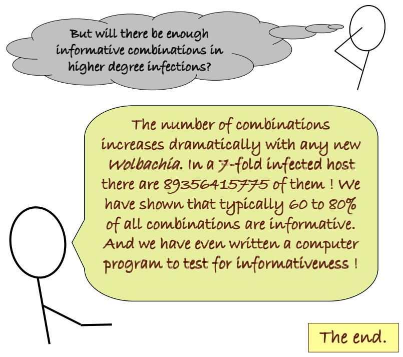
We will first resolve the red strain by intersecting the alignments of individuals A and C. The remaining sequence from A must belong to green.



From individual B we know now already all sequences except one. This must belong to the blue strain!



Finally, the last unknown sequence in C belongs to yellow.


But will there be enough informative combinations in higher degree infections?
The number of combinations increases dramatically with any new *Wolbachia*. In a 7-fold infected host there are 89356415775 of them ! We have shown that typically 60 to 80% of all combinations are informative. And we have even written a computer program to test for informativeness !
The end.