Genomic prediction in insects: a case study on wing morphology traits in the jewel wasp *Nasonia vitripennis*

Bart Pannebakker – Laboratory of Genetics





Shuwen Xia

BINGO-ITN



- Breeding Invertebrates for Next-Generation biOcontrol
- Innovative Training Network for Early Stage Researchers (2015-2020)
 - 13 PhD students, 13 partners from academia & industry

- Advance current knowledge on the use of natural genetic variation in biocontrol practice
- Enhance the application of (quantitative) genetic methods to invertebrate biocontrol
- Train young researchers in an extensive suite of interdisciplinary skills

Using genetics in biocontrol

• Selective breeding for biocontrol traits



BIOLOGICAL	Cambridge
REVIEWS	Philosophical Society
<i>Biol. Rev.</i> (2020), 95 , pp. 1838–1854. doi: 10.1111/brv.12641	1838

Next-generation biological control: the need for integrating genetics and genomics

Kelley Leung^{1*} •, Erica Ras², Kim B. Ferguson³ •, Simone Ariëns⁴, Dirk Babendreier⁵ •, Piter Bijma⁶ •, Kostas Bourtzis² •, Jacques Brodeur⁷ •, Margreet A. Bruins³, Alejandra Centurión⁴, Sophie R. Chattington⁴ •, Milena Chinchilla-Ramírez⁸ •, Marcel Dicke⁹ •, Nina E. Fatouros¹⁰ •, Joel González-Cabrera¹¹ •, Thomas V. M. Groot¹², Tim Haye⁵ •, Markus Knapp¹² •, Panagiota Koskinioti^{2,13} •, Sophie Le Hesran^{9,12}, Manolis Lyrakis^{14,15} •, Angeliki Paspati⁸ •, Meritxell Pérez-Hede⁸ •, Wouter N. Plouvier¹⁶ •, Christian Schlötterer¹⁴ •, Judith M. Stahl^{5,17} •, Andra Thiel⁴ •, Alberto Urbaneja⁸ •, Louis van de Zande¹ •, Eveline C. Verhulst⁹ •, Louise E. M. Vet^{9,18} •, Sander Visser^{19,20} •, John H. Werren²¹ •, Shuwen Xia⁶ •, Bas J. Zwaan³ •, Sara Magalhães²² •, Leo W. Beukeboom¹ • and Bart A. Pannebakker³ •

Aim

- Explore the potential for genomic prediction in insect biocontrol agents
- Wing morphology traits in Nasonia vitripennis



Nasonia vitripennis

- Gregarious parasitoid of blowfly pupae
- 20-40 offspring per host
- Males haploid, females diploid
- Parasitoid genetic model:
 - Studied since 1940s
 - Easy culturing
 - Visible mutants
 - Large set of genomic tools & resources
 - Annotated genome sequence
 - Fine-scale genetic maps
 - RNAi
 - CRISPR/Cas9
 - 4 interfertile sister species





Nasonia vitripennis – wings & biocontrol

• Sold as biocontrol agent of houseflies in stables





Dispersal

- Fitness of parasitoids related to wing size and wing shape:
 - Larger wings \rightarrow more dispersal \rightarrow more hosts parasitized
 - Optimum dispersal rate for biocontrol agents

Methods

- HVRx outbred laboratory population (N_e =236)
- *N*=1,248 females:
 - 720 females from generation 169 (G169)
 - 528 females from generation 172 (G172)
- For each female:
 - Front wing and hind tibia mounted on microscope slide
 - Body used for DNA extraction
- Host identity recorded only for G172, not for G169





Methods - phenotyping

- forewings and hind tibia mounted with Euparal
- photographed on a Zeiss scope at 2.5X magnification
- digitized landmarks using tpsDig software
- Traits measured:
 - Tibia length (µm)
 - Wing length (µm)
 - Wing width (µm)
 - Aspect ratio WL/WW (-)







Methods - genotyping

- DNA extraction
- Genotyping-by-Sequencing (GBS)
 - Reduced representation method:
 - Restriction enzyme (ApeKI)
 - Illumina HiSeq X Ten (150bp reads, paired end)
 - Reads mapped back on to reference genome



Methods - GEBV estimation

• GBLUP model:
$$y = \mu + Xb + Z_gg + \begin{pmatrix} 0 \\ Z_cc \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

- **G** matrix of genomic relationships
- Analyses done using ASREML 4.0 and Calc_grm

- y = vector of phenotypic records
- $\mu = intercept$
- X =design matrix
- $\boldsymbol{b} = \mathsf{fixed} \mathsf{effects} \mathsf{vector}$
- g = vector of random additive polygenic effects
- Z_q = incidence matrix for breeding values to animals
- $\mathbf{0} =$ vector of zeros for G169
- c = random host effects vector for G172
- $\mathbf{Z}_{c} =$ indidence matrix for host effects
- e_1 & e_2 = vector of random residuals for G169 and G172



Methods - validation

- Cross validation:
 - Across generations
 - Forward: G169 \rightarrow G172
 - Backward: G172 \rightarrow G169
 - 5-fold cross validation



- Accuracy: $\frac{cor(y,GEBV)}{\sqrt{h^2}}$
- Dispersion: regression of slope of phenotype on predicted GEBV

Results - phenotypes

	Mean (SD)				
Traits (units)	G169	G172	combined		
Tibia length (µm)	603 (42.2)	618 (30.1)	610 (38.3)		
Wing length (µm)	1921 (118.5)	2009 (88.7)	1958 (115.4)		
Wingwidth (up)	974 (57 7)	909 (44 0)	994 (52 6)		
Wing width (µm)	074 (57.7)	898 (44.0)	004 (55.0)		
Aspect ratio (-)	2.20 (0.04)	2.24 (0.03)	2.21 (0.04)		





Wing length





Aspect ratio



Results - genotypes

- 8,405,551 SNPs
 - Many off-target reads due to imperfect size selection step
 - Strict filtering was needed (on target sites)
 - 8,639 SNPs

- Slow decay of LD
 - Half-decay distance *r*²=34.1 kb



Distance between SNPs (kb)

Results – population structure

PCA plot of genotypes



Heatmap of **G** matrix



Results – validation of GEBVs

	Traits		across genera valida	ation forward	across generation validatio	backward n	5-fold cross	s-validation
		h^2	Accuracy	Dispersion	Accuracy	Dispersion	Accuracy	Dispersion
SIZE	Tibia length	0.21	-0.22	-0.21	0.21	2.21	0.52	0.96
	Winglength	0.22	-0.12	-0.12	0.17	1.72	0.60	1.18
	Wingwidth	0.20	-0.06	-0.06	0.22	1.62	0.67	1.19
	Aspect ratio	0.22	0.47	0.73	0.65	1.48	0.55	0.79

- Moderate h^2 for all traits
- Size traits small or negative accuracy across generations
- Accuracy higher for 5-fold cross validation
 - Aspect ratio higher accuracy & little dispersion

Discussion

- Promising accuracies for 5-fold cross validation and for wing aspect ratio across generations (0.47-0.67)
- Accuracy of prediction in range with beekeeper workability traits in honeybees (0.44-0.65, Bernstein *et al*. 2023)
- Low accuracies in size traits likely due to strong common host effect (Xia *et al.*, 2020):
 - ~50% of V_p for size-related traits
 - 8% of V_p for wing aspect ratio
- Common host effect is important



Outlook – genomic selection in biocontrol

- Insect biology
 - Small body sizes: destructive sampling
 - Fast generation times: fast genotyping needed
- Reliable and fast genotyping remains difficult
 - GBS not optimal
 - Other genotyping methods
 - SNP arrays for large scale production insects (e.g. BSF etc)
 - Low coverage (0.1x) WGS on MinIon platform
- Overall: genomic selection in insects works, but is technically challenging, likely more gain from phenotypic selection



Acknowledgements

- Shuwen Xia
- Gabriella Bukovinszkine Kiss
- Hendrik-Jan Megens
- Martien Groenen
- Bas Zwaan
- Piter Bijma













This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 641456







Paper: Xia et al., 2024 bioRxiv

Off diagonal values of G matrix

