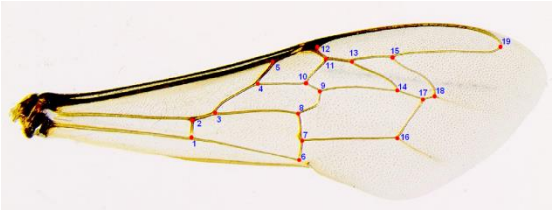
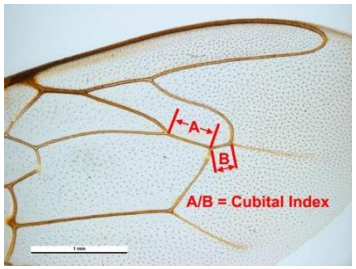
A novel SNP-based tool for estimating C-lineage introgression in the dark honey bee (*Apis mellifera mellifera*)

M. Alice Pinto, Dora Henriques, Keith A. Browne, Mark W. Barnett, Melanie Parejo, Per Kryger, Tom C. Freeman, Irene Muñoz, Lionel Garnery, Fiona Highet, J. Spencer Jonhston, Grace P. McCormack

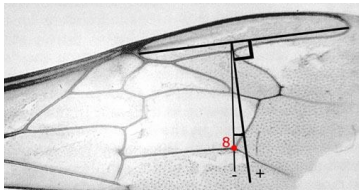
Tools for colony identification



@ Helena Ferreira



@ Simon Hinkley & Ken Walker

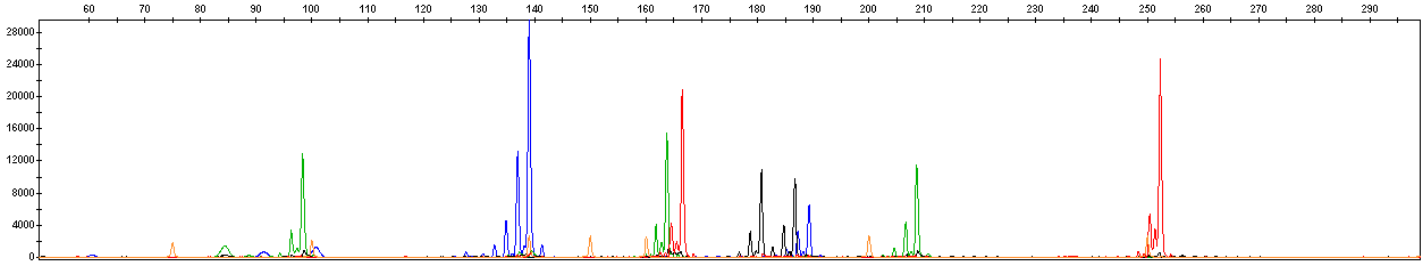


http://www.cybis.se/cbeewing/pertxt/index.htm

Morphometry

- ✓ ApiClass
- ✓ DrawWing
- ✓ BeeMorph
- ✓ CBeeWing

Microsatellites



Single Nucleotide Polymorphisms (SNPs)

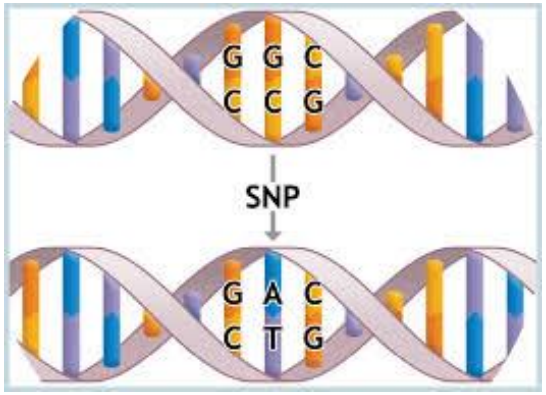
frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 21 December 2016
doi: 10.3389/fevo.2016.00140

Using Whole-Genome Sequence Information to Foster Conservation Efforts for the European Dark Honey Bee, *Apis mellifera mellifera*

Melanie Parejo[1,2], David Wragg[3,4], Laurent Gauthier[1], Alain Vignal[3], Peter Neumann[1,2,5] and Markus Neuditschko[1]**

> 3 million SNPs



Advantages of SNPs

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2016) doi: 10.1111/1755-0998.12637

SNPs selected by information content outperform randomly selected microsatellite loci for delineating genetic identification and introgression in the endangered dark European honeybee (*Apis mellifera mellifera*)

IRENE MUÑOZ,*,† DORA HENRIQUES,*,‡ LAURA JARA,† J. SPENCER JOHNSTON,§
JULIO CHÁVEZ-GALARZA,* PILAR DE LA RÚA† and M. ALICE PINTO*

- Powerful for estimating C-lineage introgression

PLOS | ONE

RESEARCH ARTICLE

Reduced SNP Panels for Genetic Identification and Introgression Analysis in the Dark Honey Bee (*Apis mellifera mellifera*)

Irene Muñoz[1], Dora Henriques[1], J. Spencer Johnston[2], Julio Chávez-Galarza[1], Per Kryger[3], M. Alice Pinto[1]*

1 Mountain Research Centre (CIMO), Polytechnic Institute of Bragança, Campus de Sta. Apolónia, Apartado 1172, 5301–855, Bragança, Portugal, 2 Department of Entomology, Texas A&M University, College Station, Texas, 77843–2475, United States of America, 3 Aarhus University, Department of Agroecology, Forsøgsvej 1, 4200, Slagelse, Denmark

* apinto@ipb.pt

OPEN ACCESS

Citation: Muñoz I, Henriques D, Johnston JS, Chávez-Galarza J, Kryger P, Pinto MA (2015) Reduced SNP Panels for Genetic Identification and Introgression Analysis in the Dark Honey Bee (*Apis mellifera mellifera*). PLoS ONE 10(4): e0124385. doi:10.1371/journal.pone.0124385

Academic Editor: Wolfgang Blenau, University of Cologne, GERMANY

Received: January 9, 2015

Accepted: March 10, 2015

Published: April 13, 2015

Copyright: © 2015 Muñoz et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Muñoz was supported by Fundación Séneca (Murcia, Spain) through the Post-doctoral fellowship 19149/PD13-N whereas Dora Henriques and Julio Chávez-Galarza were supported by Fundação para a Ciência e Tecnologia (Portugal) through the PhD scholarships SFRH/BD/84195/2012 and SFRH/BD/68662/2010, respectively. This research was funded by Fundação para a Ciência e Tecnologia and COMPETE/QREN/EU through the projects PTDC/BIA-BEC/099640/2008 and

Abstract

Beekeeping activities, especially queen trading, have shaped the distrib
(*Apis mellifera*) subspecies in Europe, and have resulted in extensive int
eastern European C-lineage subspecies (*A. m. ligustica* and *A. m. carni*
range of the M-lineage *A. m. mellifera* subspecies in Western Europe. A
replacement and gene flow between native and commercial populations
varying levels across western European populations. Genetic identificat
sion analysis using molecular markers is an important tool for managem
tion of honey bee subspecies. Previous studies have monitored introgre
microsatellite, PCR-RFLP markers and most recently, high density assa
cleotide polymorphism (SNP) markers. While the latter are almost prohit
the information gained to date can be exploited to create a reduced pane
most ancestry-informative markers (AIMs) for those purposes with very l
tion. The objective of this study was to design reduced panels of AIMs to
A. m. mellifera individuals and to provide accurate estimates of the level
gression into their genome. The discriminant power of the SNPs using a
and approaches including the Weir & Cockerham's F_{ST}, an F_{ST}-based c
formativeness (I_n), and PCA was evaluated. This study shows that reduc
sign individuals to the correct origin and calculates the admixture level w
accuracy. These panels provide an essential tool in Europe for genetic s
and estimation of admixture levels which can assist management strate
honey bee conservation programs.

SCIENTIFIC REPORTS

OPEN

High sample throughput genotyping for estimating C-lineage introgression in the dark honeybee: an accurate and cost-effective SNP-based tool

Received: 18 January 2018
Accepted: 17 May 2018
Published online: 04 June 2018

Dora Henriques[1,2], Keith A. Browne[3], Mark W. Barnett[4], Melanie Parejo[5], Per Kryger[6], Tom C. Freeman[4], Irene Muñoz[7], Lionel Garnery[8,9], Fiona Highet[10], J. Spencer Jonhston[11], Grace P. McCormack[3] & M. Alice Pinto[1]

The natural distribution of the honeybee (*Apis mellifera* L.) has been changed by humans in recent decades to such an extent that the formerly widest-spread European subspecies, *Apis mellifera mellifera*, is threatened by extinction through introgression from highly divergent commercial strains in large tracts of its range. Conservation efforts for *A. m. mellifera* are underway in multiple European countries requiring reliable and cost-efficient molecular tools to identify purebred colonies. Here, we developed four ancestry-informative SNP assays for high sample throughput genotyping using the iPLEX Mass Array system. Our customized assays were tested on DNA from individual and pooled, haploid and diploid honeybee samples extracted from different tissues using a diverse range of protocols. The assays had a high genotyping success rate and yielded accurate genotypes. Performance assessed against whole-genome data showed that individual assays behaved well, although the most accurate introgression estimates were obtained for the four assays combined (117 SNPs). The best compromise between accuracy and genotyping costs was achieved when combining two assays (62 SNPs). We provide a ready-to-use cost-effective tool for accurate molecular identification and estimation of introgression levels to more effectively monitor and manage *A. m. mellifera* conservatories.

PLOS | ONE

RESEARCH ARTICLE

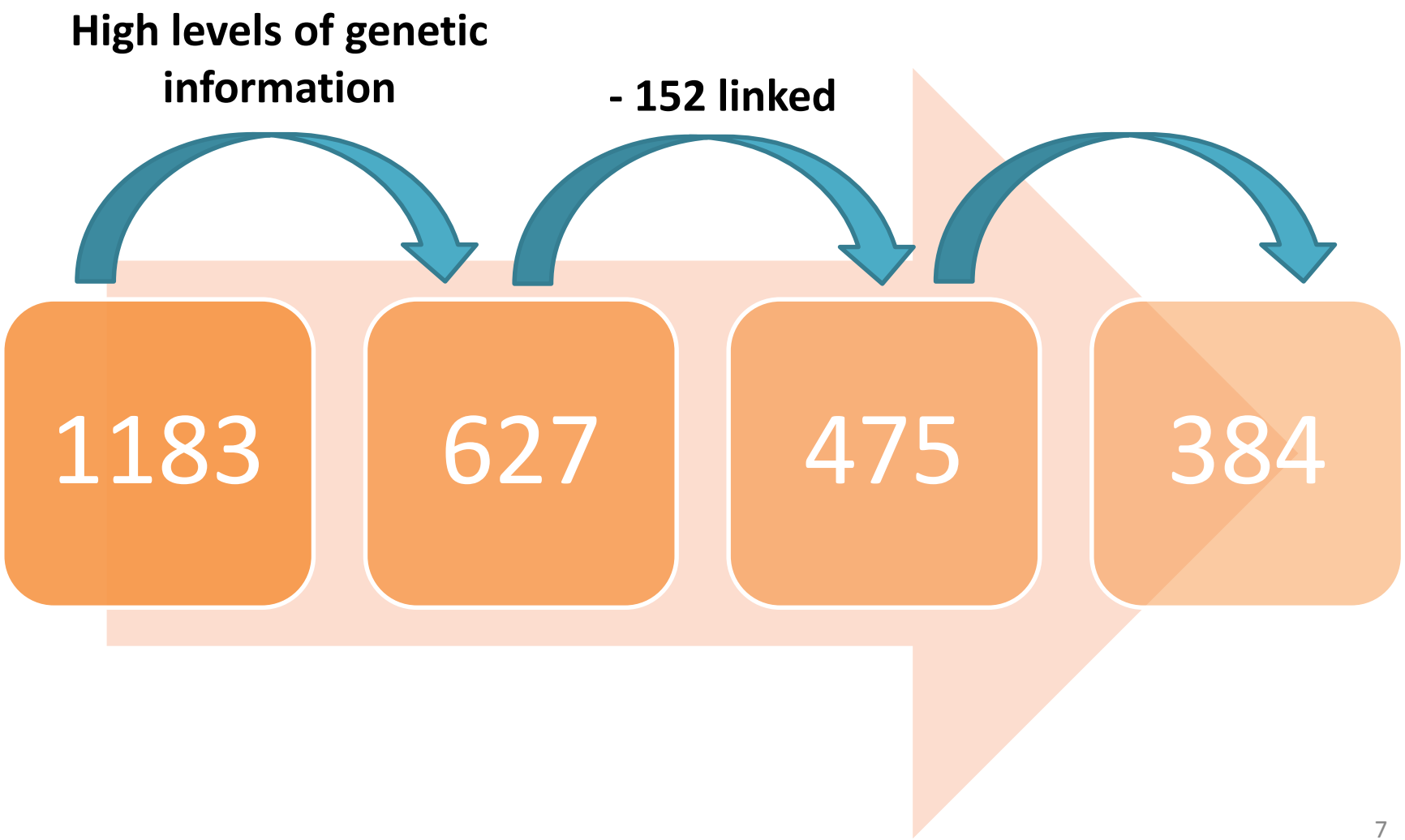
Reduced SNP Panels for Genetic Identification and Introgression Analysis in the Dark Honey Bee (*Apis mellifera mellifera*)

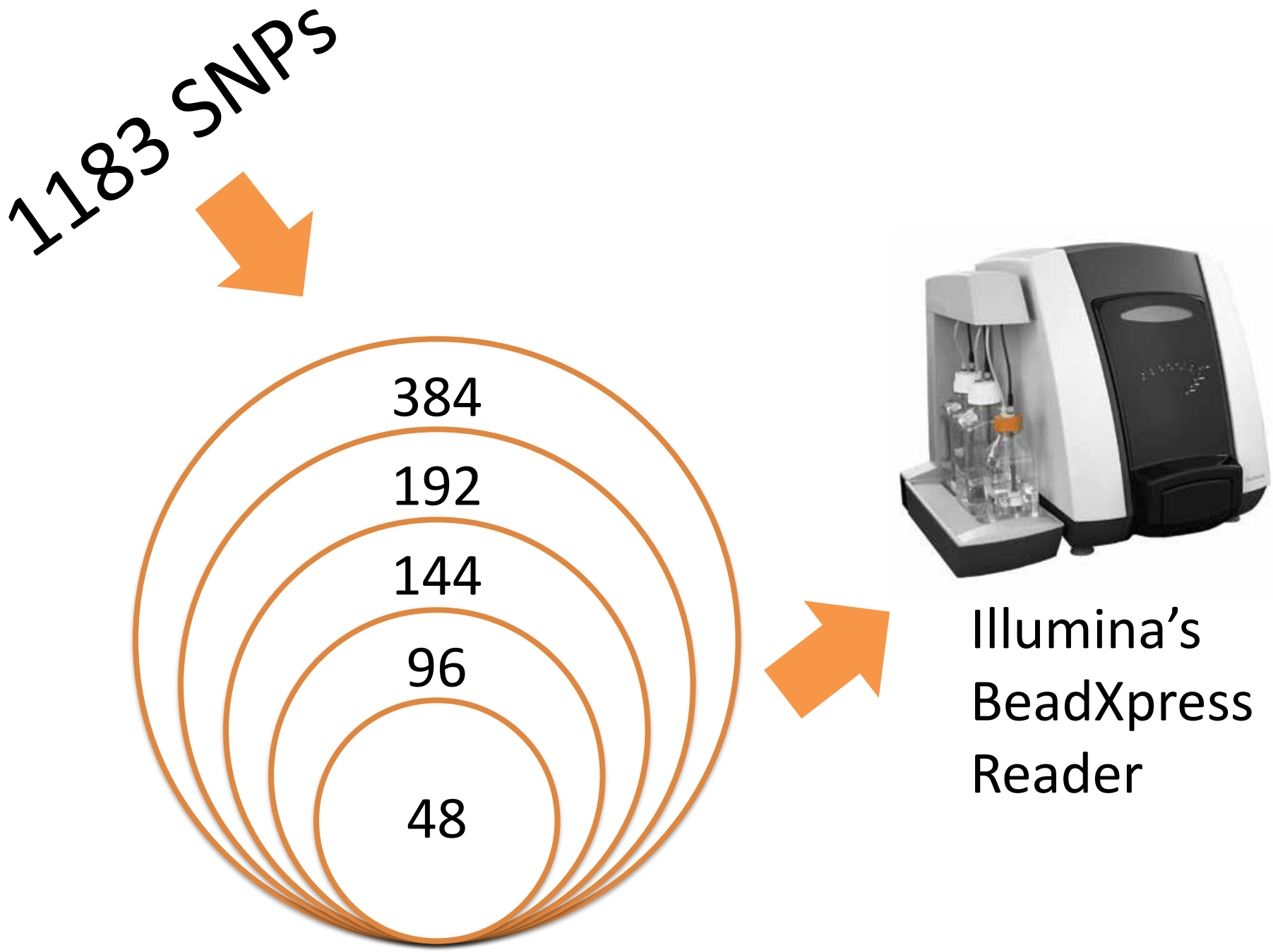
Irene Muñoz[1], Dora Henriques[1], J. Spencer Johnston[2], Julio Chávez-Galarza[1], Per Kryger[3], M. Alice Pinto[1]*

1 Mountain Research Centre (CIMO), Polytechnic Institute of Bragança, Campus de Sta. Apolónia, Apartado 1172, 5301–855, Bragança, Portugal, **2** Department of Entomology, Texas A&M University, College Station, Texas, 77843–2475, United States of America, **3** Aarhus University, Department of Agroecology, Forsøgsvej 1, 4200, Slagelse, Denmark

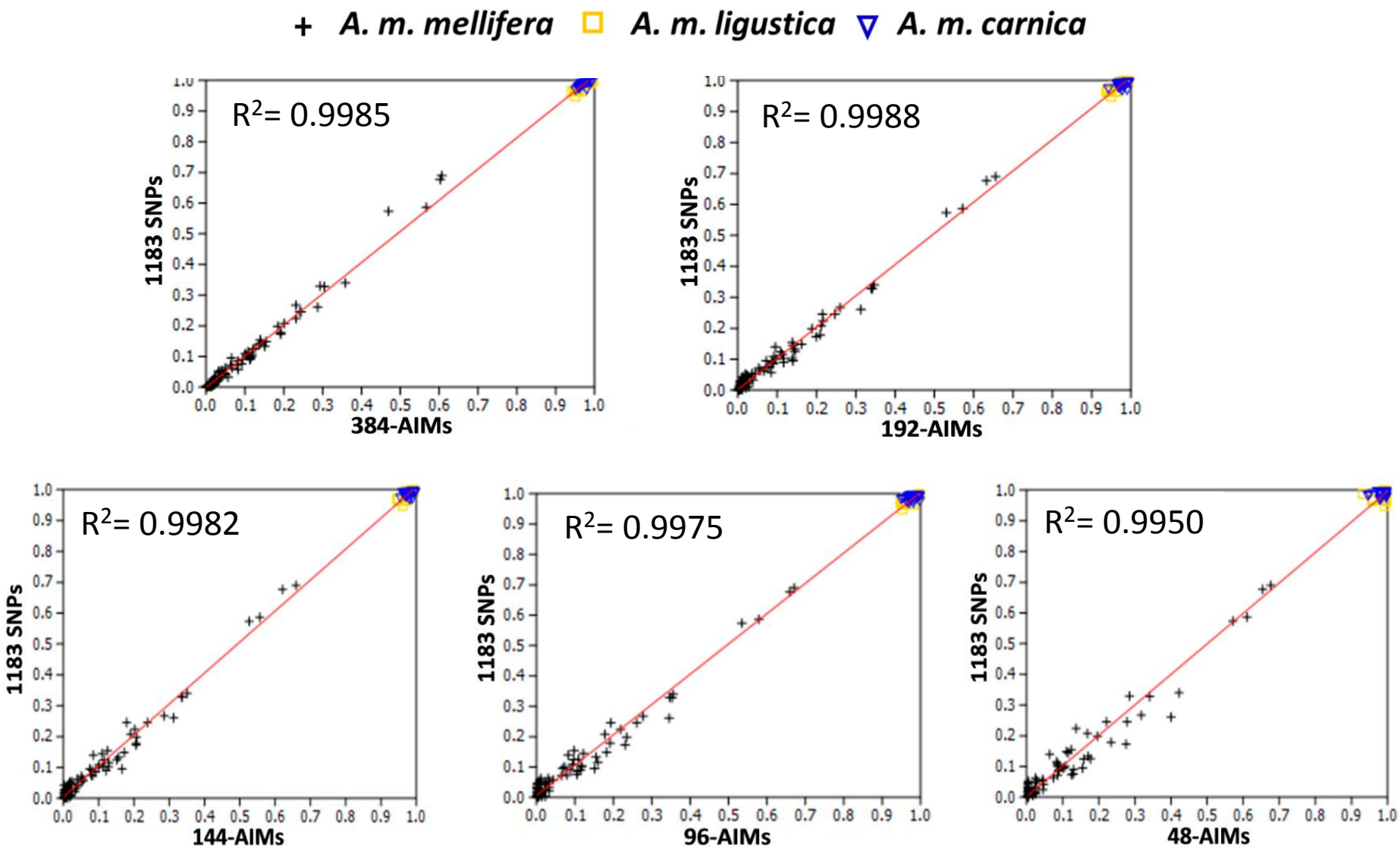
* apinto@ipb.pt

CrossMark
click for updates

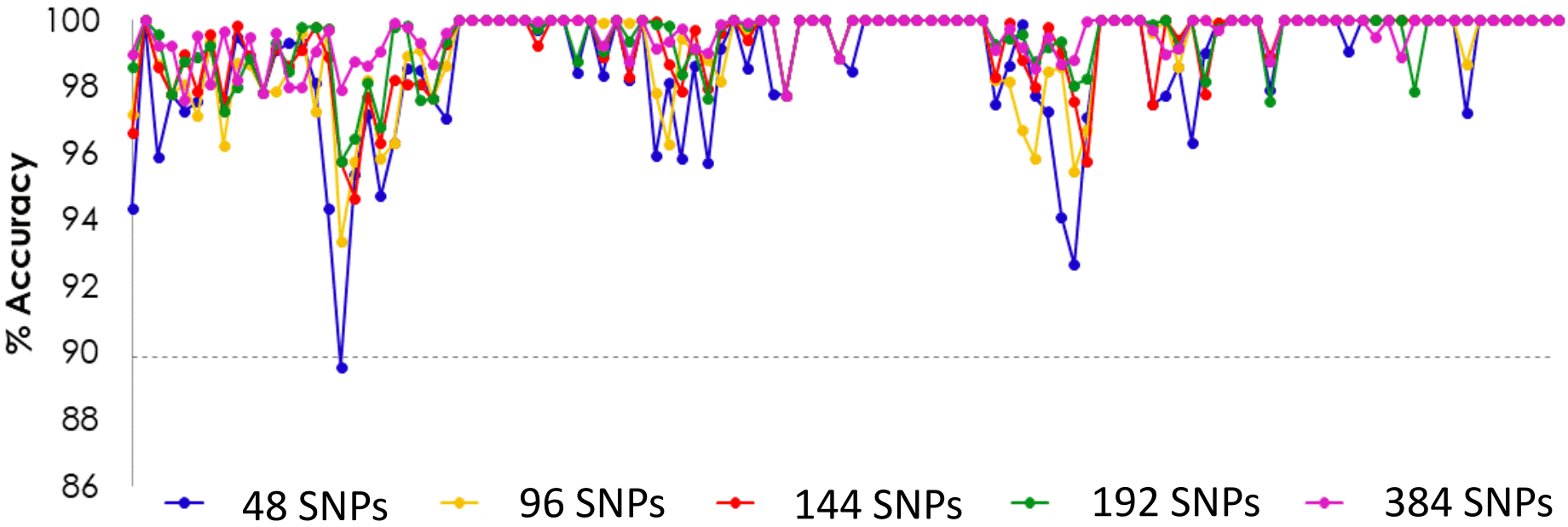

High levels of genetic information
- 152 linked
1183
627
475
384


1183 SNPs
384
192
144
96
48
Illumina's
BeadXpress
Reader

Precision


+ A. m. mellifera
A. m. ligustica
A. m. carnica
R2= 0.9985
1183 SNPs
384-AIMs
R2= 0.9988
1183 SNPs
192-AIMs
R2= 0.9982
1183 SNPs
144-AIMs
R2= 0.9975
1183 SNPs
96-AIMs
R2= 0.9950
1183 SNPs
48-AIMs

Accuracy



	48 SNPs	96 SNPs	144 SNPs	192 SNPs	384 SNPs
Correct Assignment	**98.2%**	**98.8%**	**99.0%**	**99.2%**	**99.4%**

Genotyping platform

Illumina's BeadXpress Reader

Agena iPLEX MassARRAY

1) Initial locus-specific PCR reaction

2) Single base extension using mass-modified dideoxynucleotide terminators of an oligonucleotide primer which anneals immediately upstream of the SNP

3) Using MALDI-TOF mass spectrometry, the distinct mass of the extended primer identifies the SNP allele

SCIENTIFIC REPORTS

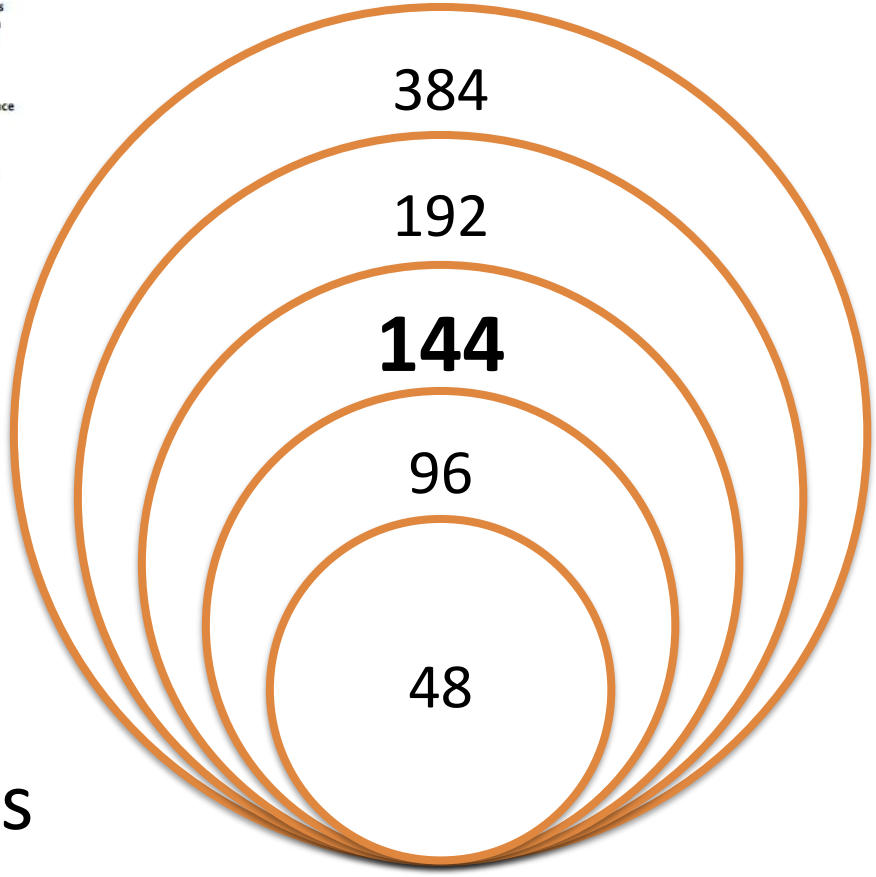
OPEN

Received: 18 January 2018
Accepted: 17 May 2018
Published online: 04 June 2018

High sample throughput genotyping for estimating C-lineage introgression in the dark honeybee: an accurate and cost-effective SNP-based tool

Dora Henriques[1,2], Keith A. Browne[3], Mark W. Barnett[4], Melanie Parejo[5], Per Kryger[6], Tom C. Freeman[4], Irene Muñoz[7], Lionel Garnery[8,9], Fiona Highet[10], J. Spencer Jonhston[11], Grace P. McCormack[3] & M. Alice Pinto[1]

The natural distribution of the honeybee (*Apis mellifera* L.) has been changed by humans in recent decades to such an extent that the formerly widest-spread European subspecies, *Apis mellifera mellifera*, is threatened by extinction through introgression from highly divergent commercial strains in large tracts of its range. Conservation efforts for *A. m. mellifera* are underway in multiple European countries requiring reliable and cost-efficient molecular tools to identify purebred colonies. Here, we developed four ancestry-informative SNP assays for high sample throughput genotyping using the iPLEX Mass Array system. Our customized assays were tested on DNA from individual and pooled, haploid and diploid honeybee samples extracted from different tissues using a diverse range of protocols. The assays had a high genotyping success rate and yielded accurate genotypes. Performance assessed against whole-genome data showed that individual assays behaved well, although the most accurate introgression estimates were obtained for the four assays combined (117 SNPs). The best compromise between accuracy and genotyping costs was achieved when combining two assays (62 SNPs). We provide a ready-to-use cost-effective tool for accurate molecular identification and estimation of introgression levels to more effectively monitor and manage *A. m. mellifera* conservatories.



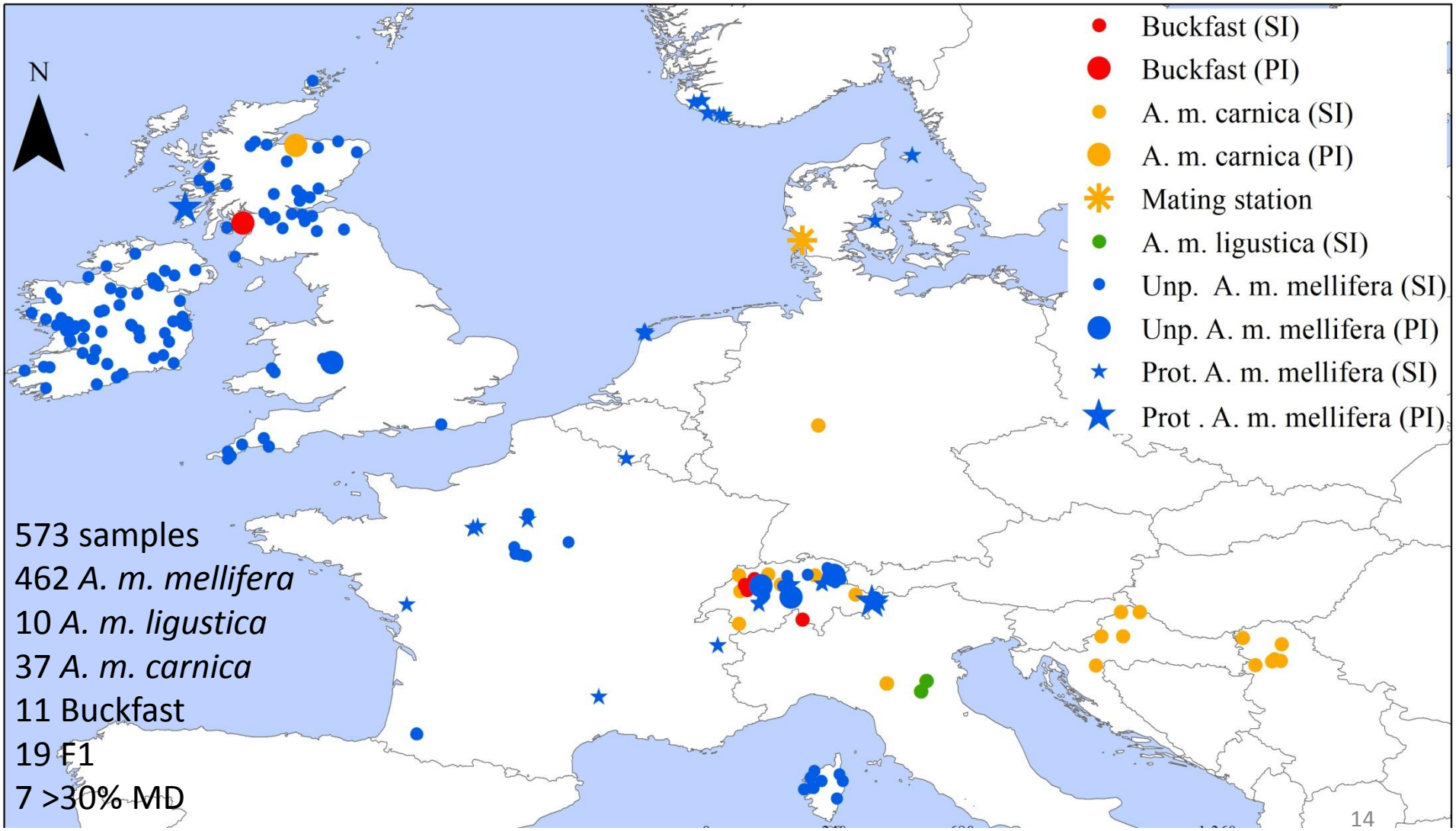
5 SNP panels

Assay design

144 SNPs → 127 SNPs → 4 SNP assays

• M1	• M2	• M3	• M4
38	35	30	24

Sample location



573 samples
462 *A. m. mellifera*
10 *A. m. ligustica*
37 *A. m. carnica*
11 Buckfast
19 F1
7 >30% MD

- Head
- Thorax
- Legs
- Thorax + abdomen

- Phenol-chloroform
- Magnetic beads/KingFisher™ Flex Purification System
- CTAB
- Omega bio-tek EZNA kit
- Qiagen EZ1 DNA tissue kit

- Pooled DNA
- Pooled individuals
- Single individuals

- Diploid
- Haploid

Assay sizes

144 SNPs → 127 SNPs → 4 SNP assays

Quality control

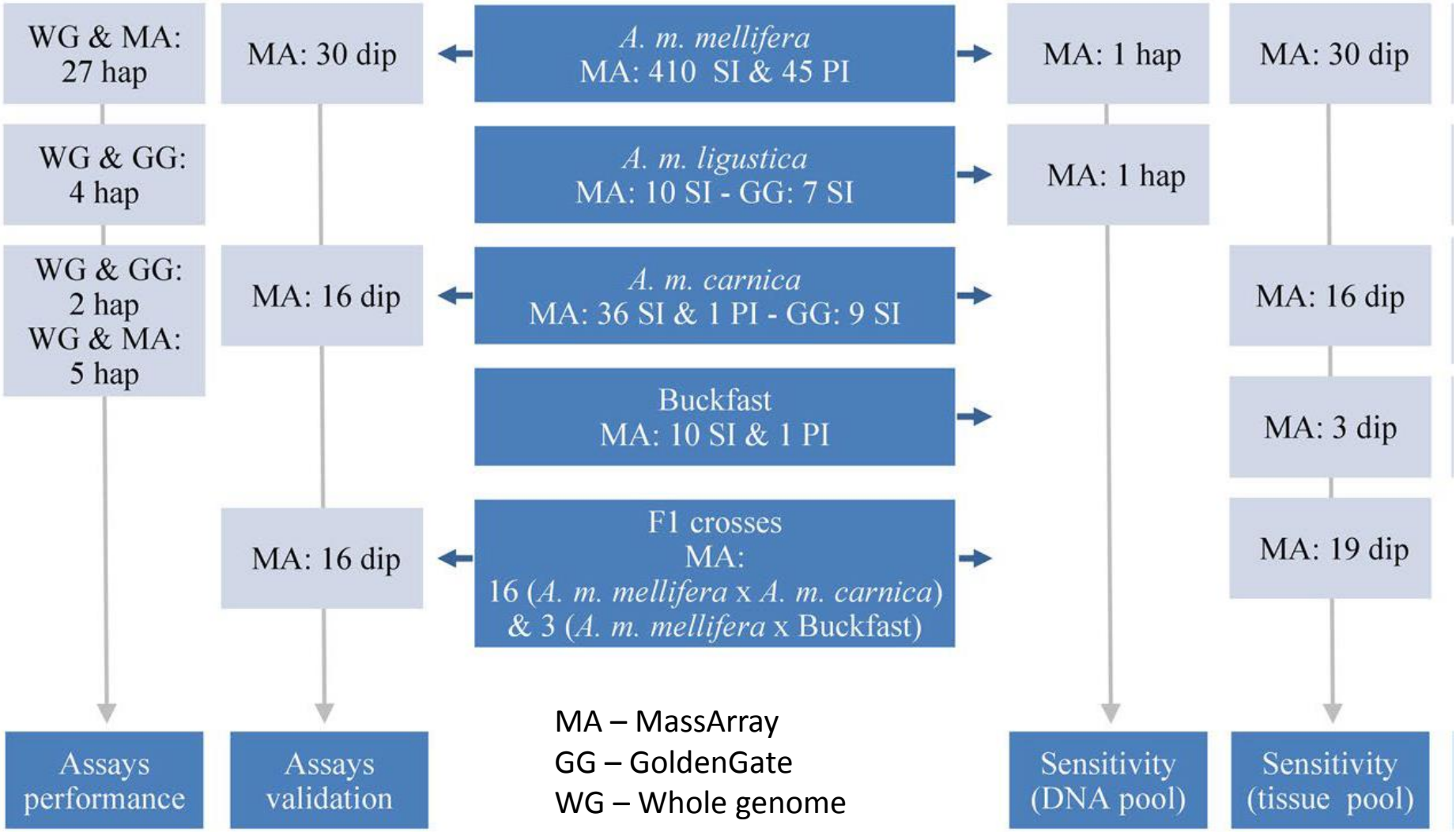
- Missing data (80%)
- Het. (>10% hapl ind.)

M1	M2	M3	M4
38	35	30	24

M1	M2	M3	M4
34	32	28	23

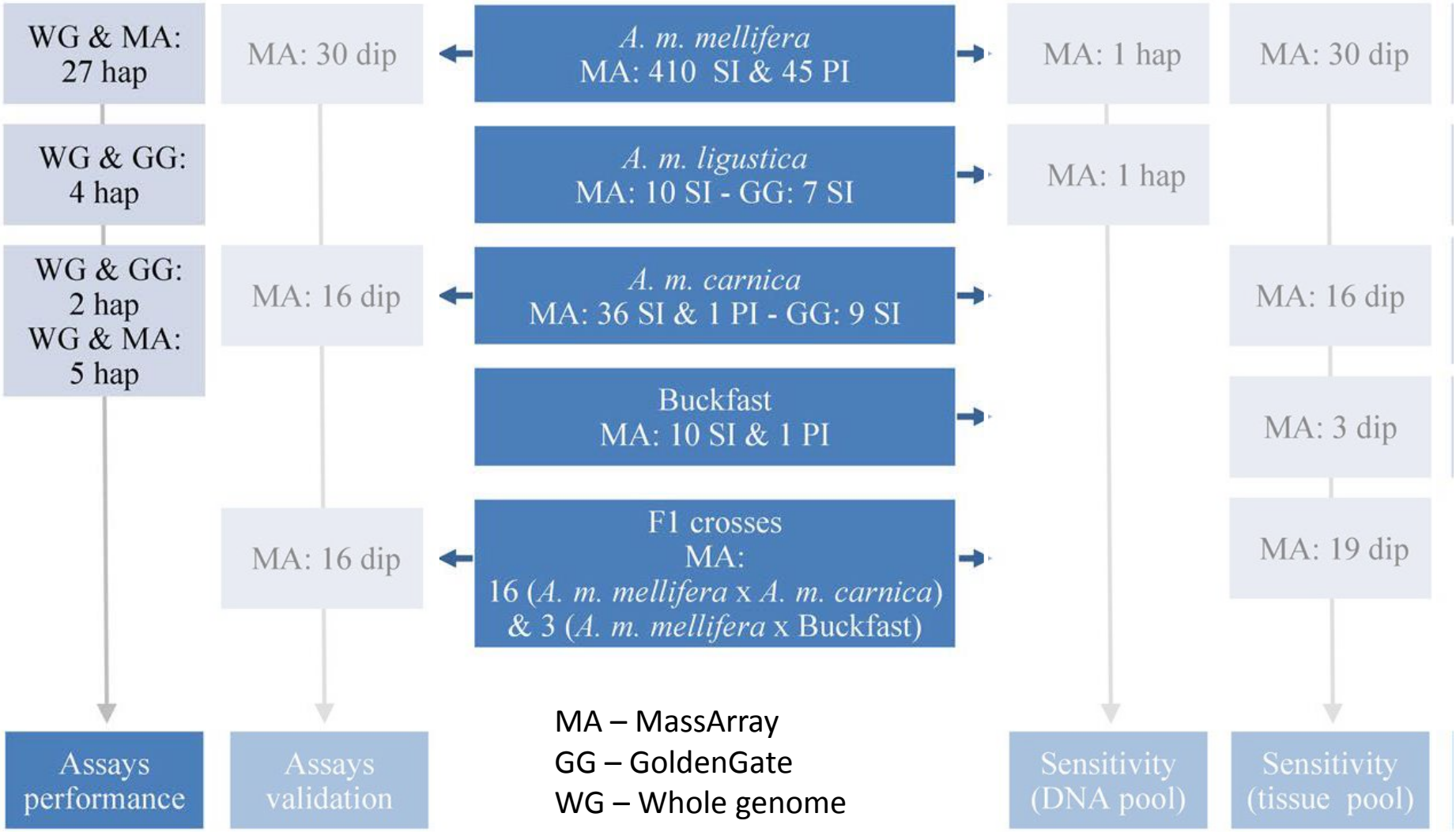
Total= 117 SNPs

Performance testing and validation of the four assays



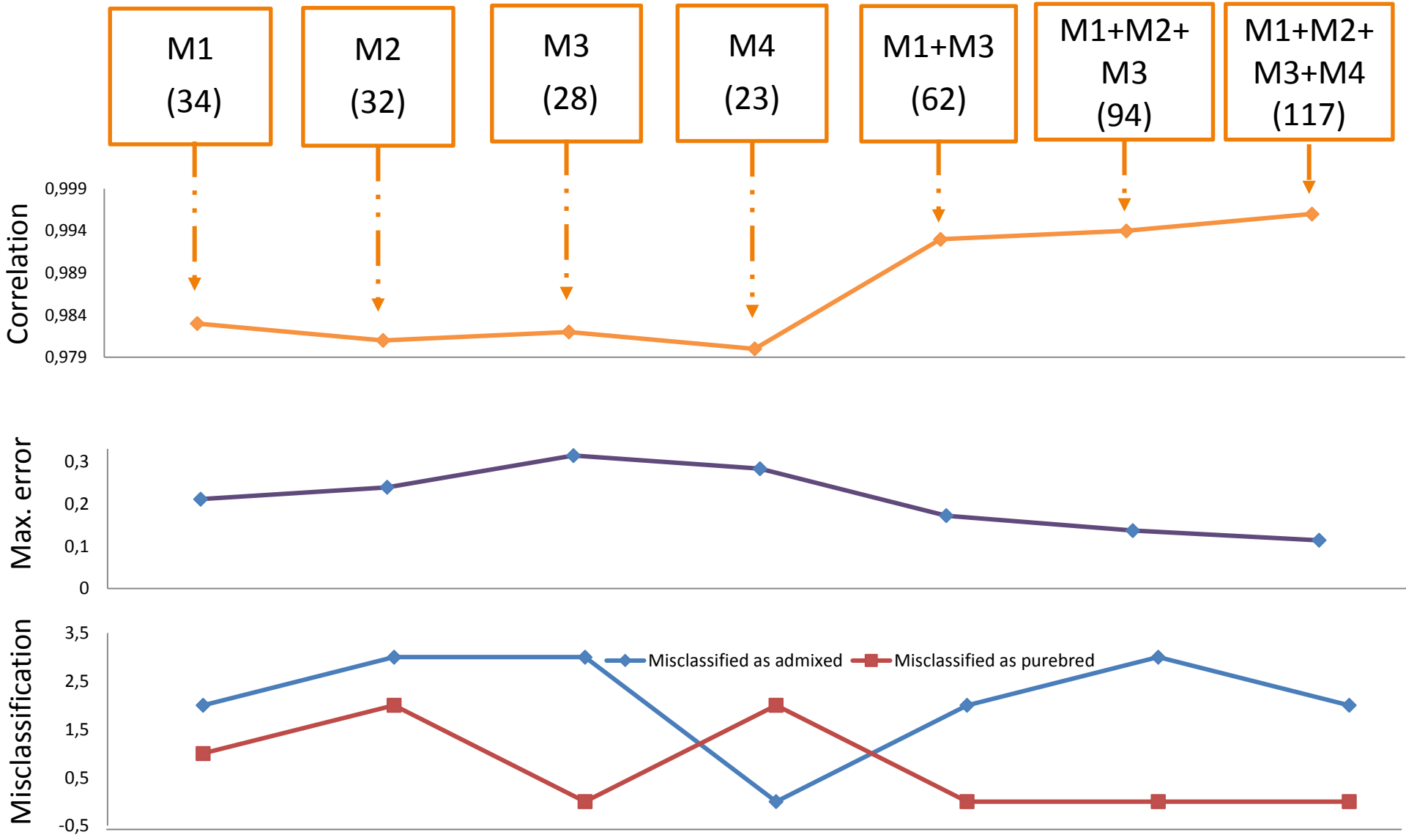
MA – MassArray
GG – GoldenGate
WG – Whole genome

Performance testing and validation of the four assays

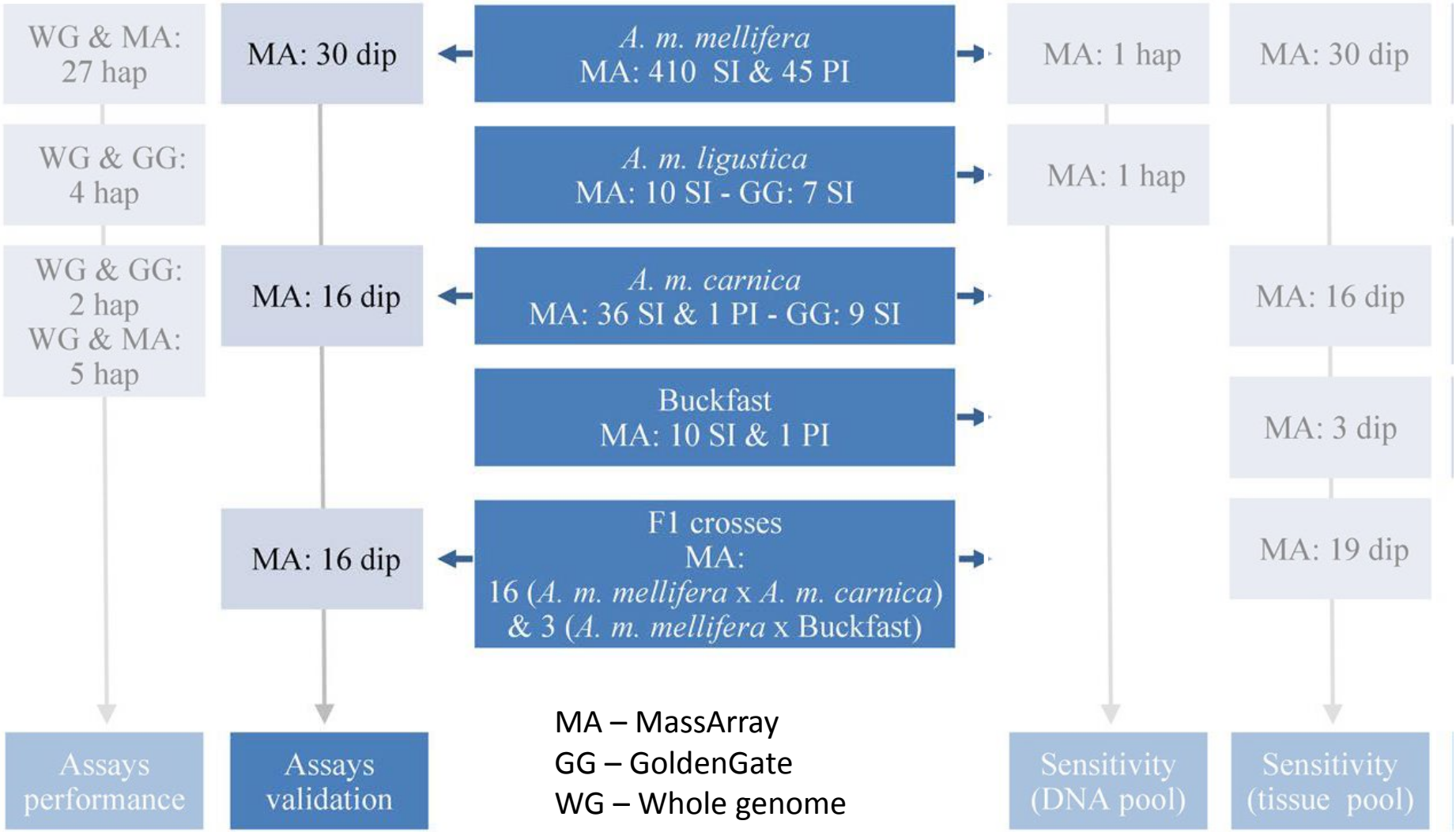


MA – MassArray
GG – GoldenGate
WG – Whole genome

Assays performance: comparisons with whole genomes



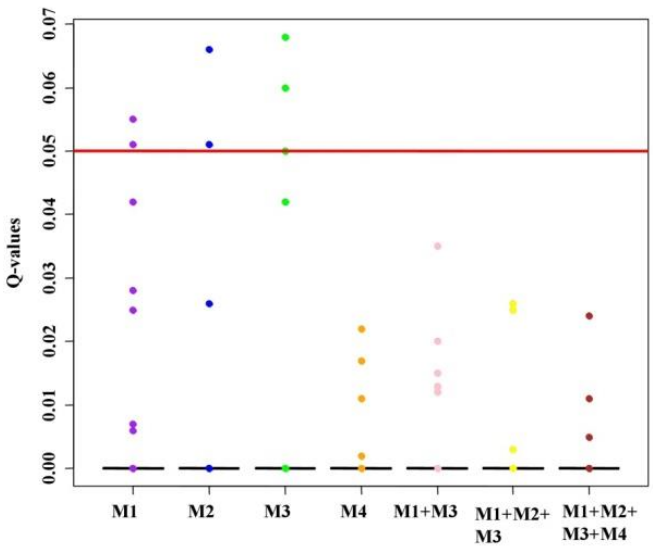
Performance testing and validation of the four assays



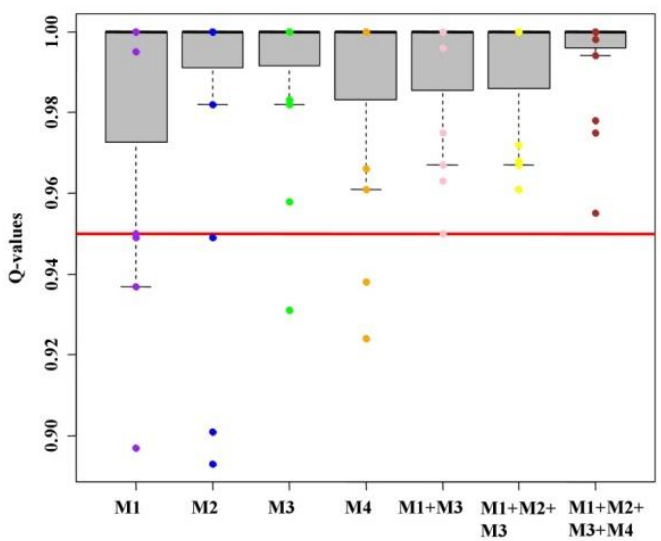
MA – MassArray
GG – GoldenGate
WG – Whole genome

Validating the SNP assays

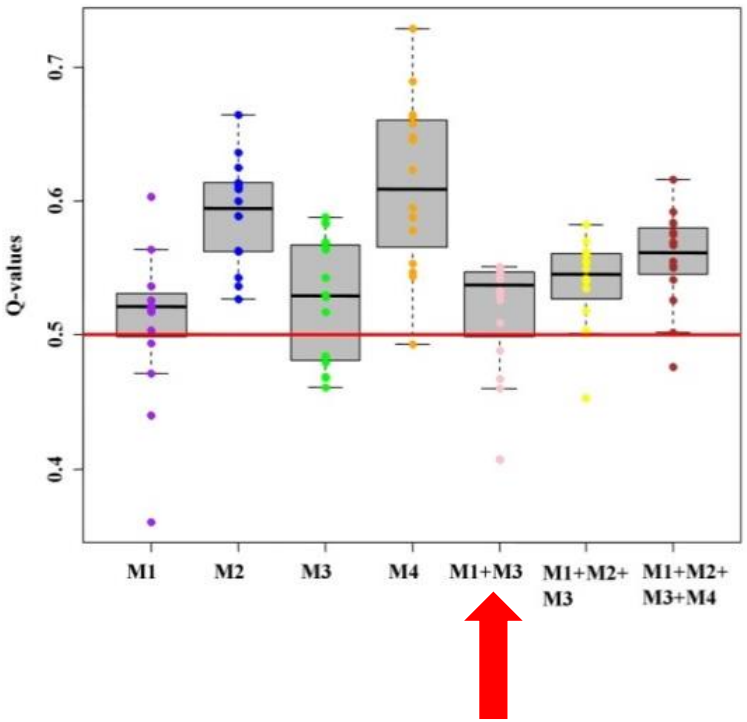
A. m. mellifera
expected Q-value <0.05



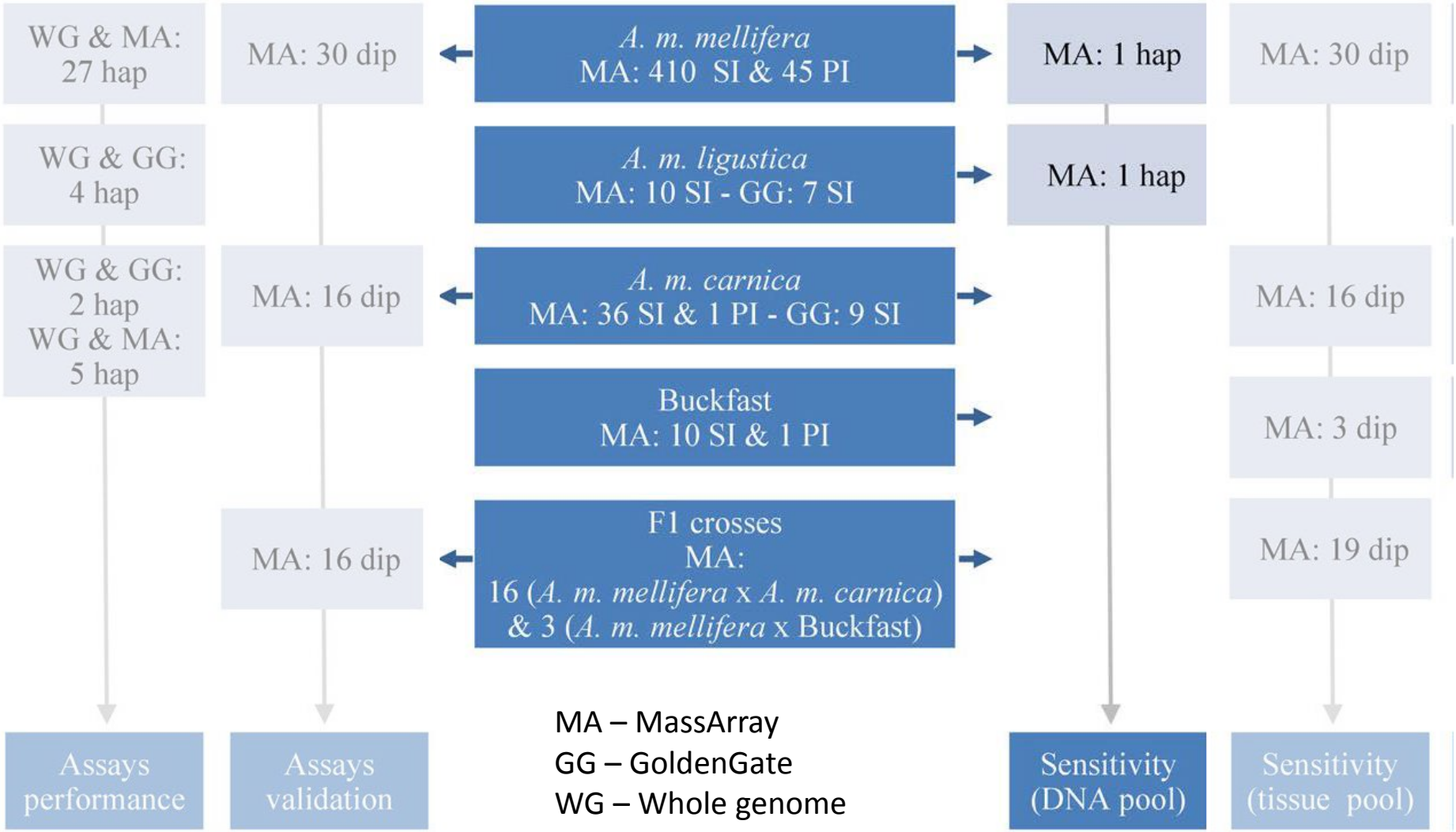
A. m. carnica
expected Q-value >0.95



F1 hybrids
(expected Q-value 0.50)



Performance testing and validation of the four assays

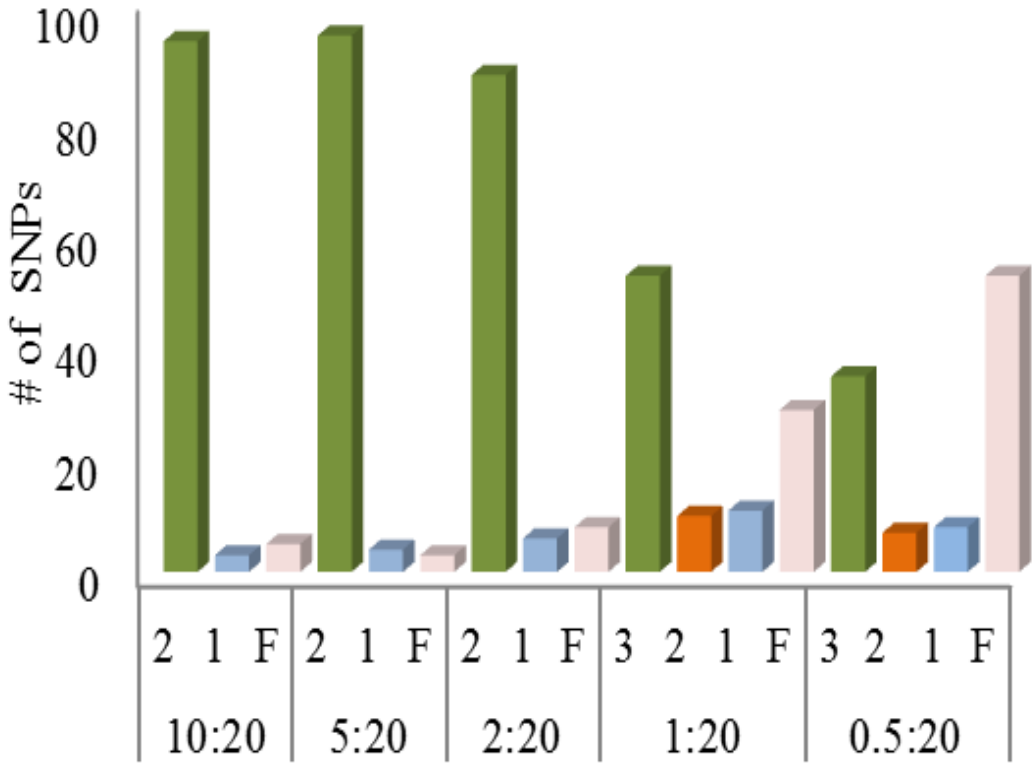


MA – MassArray
GG – GoldenGate
WG – Whole genome

Assessing sensitivity of the MassARRAY system in pooled DNA

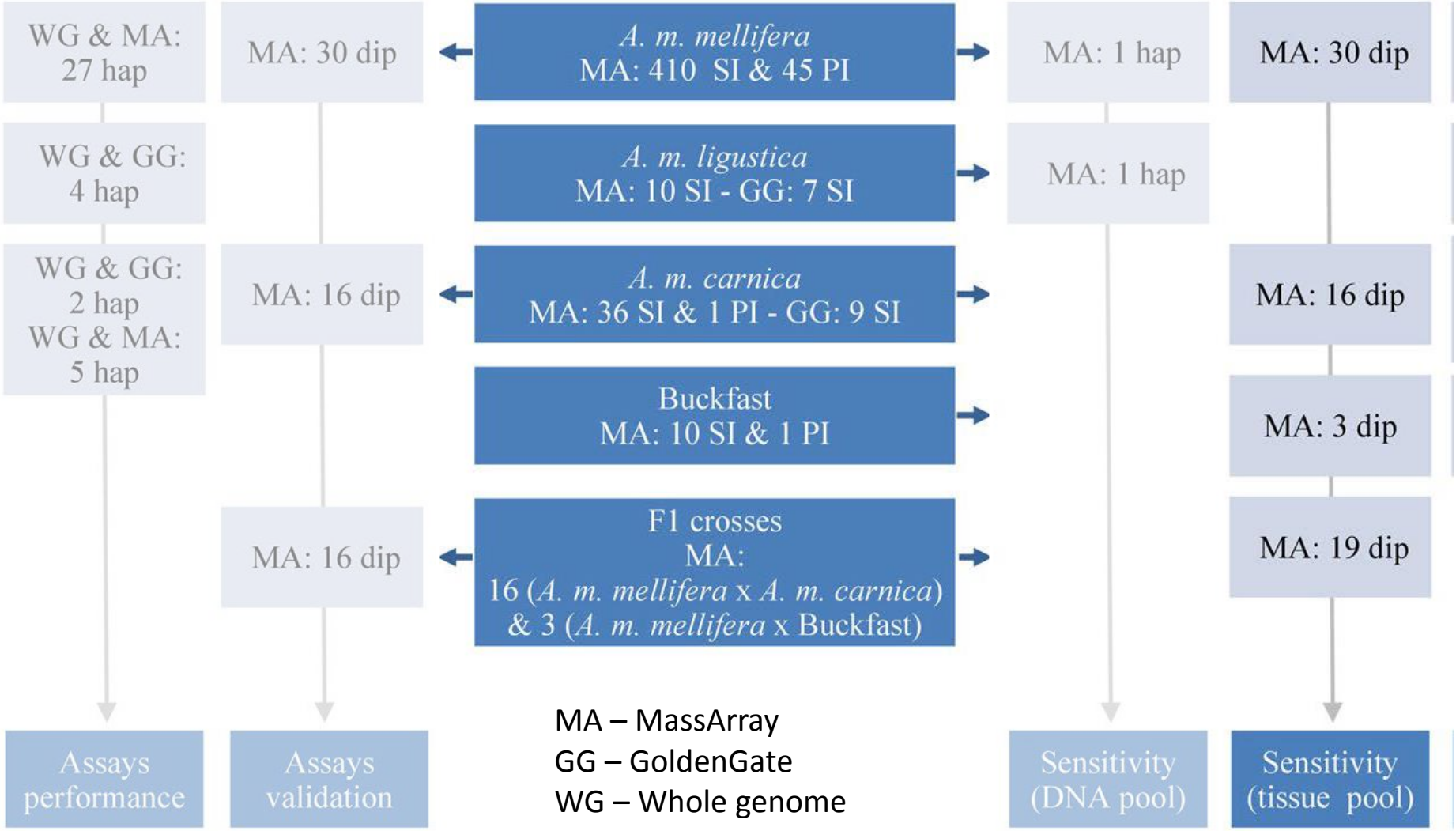
A. m. ligustica : A. m. mellifera

10:20	5:20	2:20	1:20	0.5:20



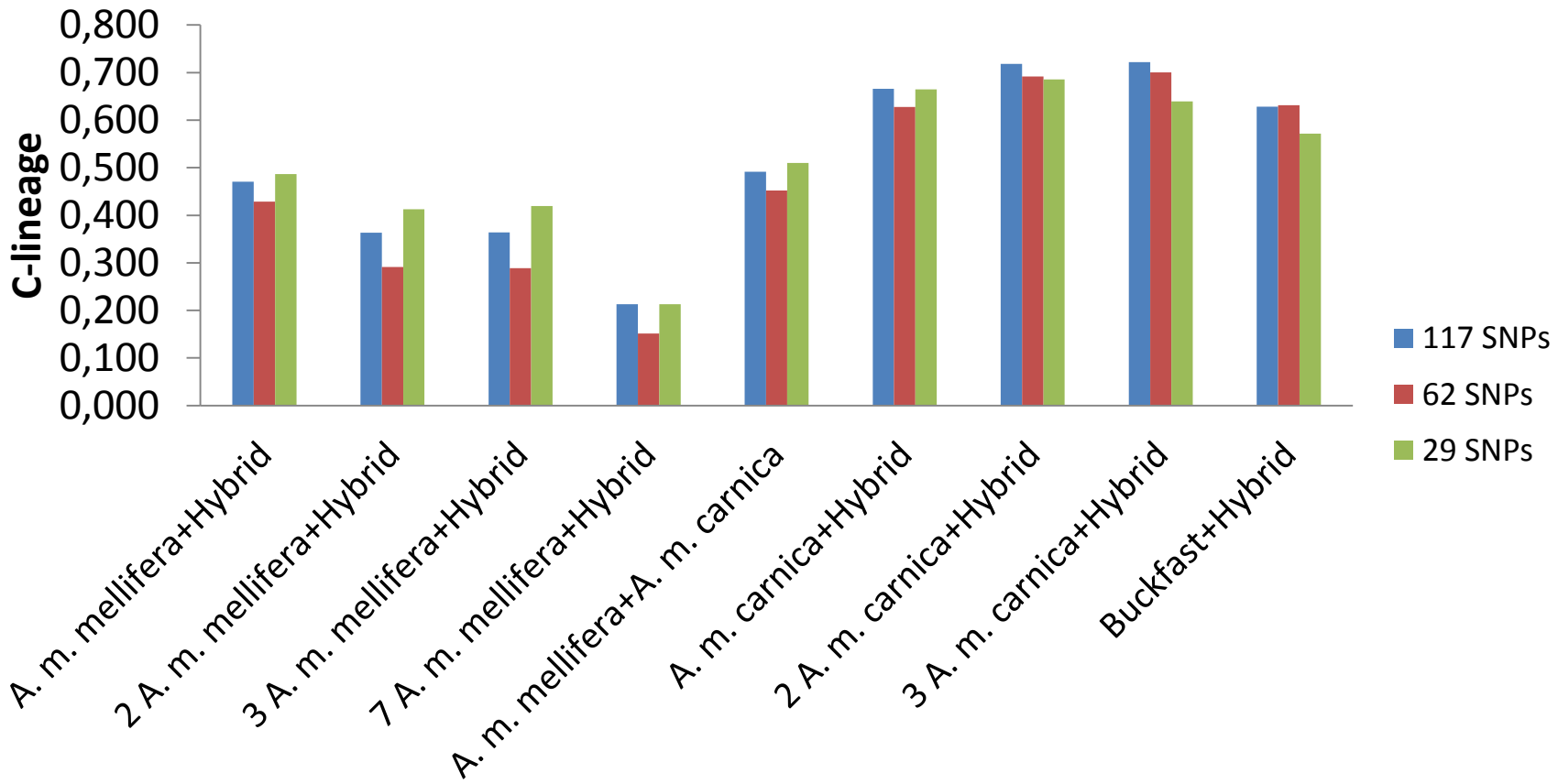
1,2,3 - # of r eplicates
F-Failed SNPs

Performance testing and validation of the four assays



MA – MassArray
GG – GoldenGate
WG – Whole genome

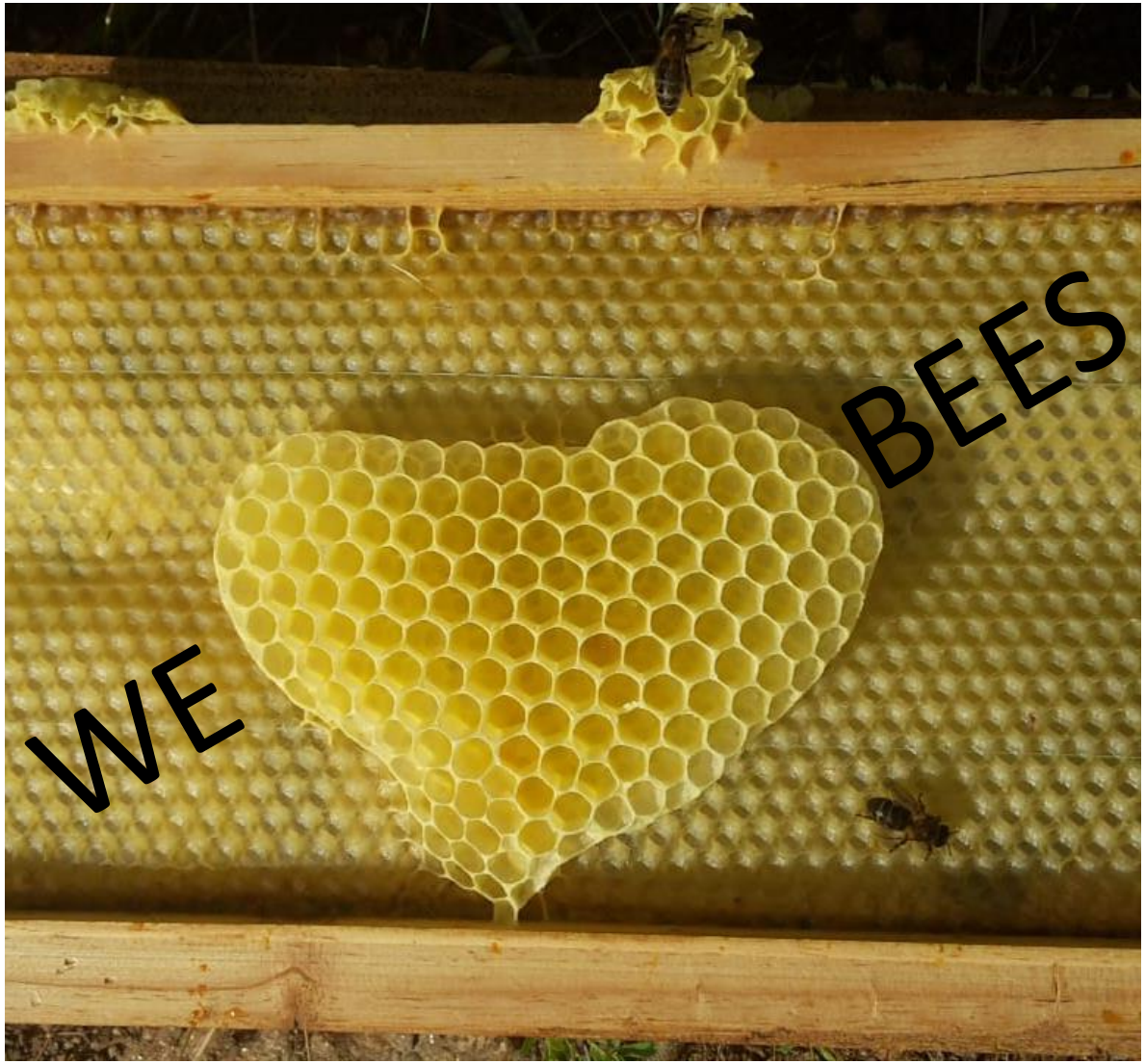
Assessing sensitivity of the MassARRAY system in pooled tissue



Concluding remarks

- ✓ The four assays and the genotyping platform worked well in DNAs extracted from different tissues and using a diverse array of protocols
- ✓ The assays had a high genotyping success rate and yielded accurate genotypes
- ✓ Performance assessed against whole-genome data showed that individual assays behave well, although the most accurate introgression estimates were obtained when the four assays wre combined (117 SNPs)
- ✓ The best compromise between accuracy and genotyping costs is achieved when combining assays M1 and M3 (62 SNPs), allowing genotyping twice as many samples for the same price

Thank you for you attention



Acknowledgements

- João Costa (Instituto Gulbenkian Ciência, Oeiras, Portugal), designed the multiplexes and genotyped the samples.
- José Rufino (IPB) provided computational resources for performing the analyses.
- Financial support was provided by 2013–2014 BiodivERsA/FACCE-JPI joint call for research proposals, with the national funders FCT (Portugal), “Agence Nationale de la Recherche” (France), and “Ministério de Economia y Competividade” (Spain).