Australian Government
Rural Industries Research and Development Corporation

RURAL INDUSTRIES
Research & Development Corporation

Genetic Evaluation of Australian Honey Bees using BLUP procedures

December 2015

RIRDC Publication No. 15/104

Australian Government

Rural Industries Research and
Development Corporation

Genetic Evaluation of Australian Honey Bees using BLUP procedures

by R. G. Banks

December 2015

RIRDC Publication No 15/104
RIRDC Project No PRJ-009276

© 2015 Rural Industries Research and Development Corporation.
All rights reserved.

ISBN 978-1-74254-841-8
ISSN 1440-6845

Genetic Evaluation of Australian Honey Bees using BLUP Procedures
Publication No. 15/104
Project No. PRJ-009276

The information contained in this publication is intended for general use to assist public knowledge and discussion and to help improve the development of sustainable regions. You must not rely on any information contained in this publication without taking specialist advice relevant to your particular circumstances.

While reasonable care has been taken in preparing this publication to ensure that information is true and correct, the Commonwealth of Australia gives no assurance as to the accuracy of any information in this publication.

The Commonwealth of Australia, the Rural Industries Research and Development Corporation (RIRDC), the authors or contributors expressly disclaim, to the maximum extent permitted by law, all responsibility and liability to any person, arising directly or indirectly from any act or omission, or for any consequences of any such act or omission, made in reliance on the contents of this publication, whether or not caused by any negligence on the part of the Commonwealth of Australia, RIRDC, the authors or contributors.

The Commonwealth of Australia does not necessarily endorse the views in this publication.

This publication is copyright. Apart from any use as permitted under the *Copyright Act 1968*, all other rights are reserved. However, wide dissemination is encouraged. Requests and inquiries concerning reproduction and rights should be addressed to RIRDC Communications on phone 02 6271 4100.

Researcher Contact Details

Name: R. G. Banks
Address: Animal Genetics and Breeding Unit
Animal Genetics and Breeding Unit
University of New England
Armidale NSW 2351

Phone: 02 6773 2425
Email: rbanks@une.edu.au

In submitting this report, the researcher has agreed to RIRDC publishing this material in its edited form.

RIRDC Contact Details

Rural Industries Research and Development Corporation
Level 2, 15 National Circuit
BARTON ACT 2600

PO Box 4776
KINGSTON ACT 2604

Phone: 02 6271 4100
Fax: 02 6271 4199
Email: rirdc@rirdc.gov.au.
Web: http://www.rirdc.gov.au

Electronically published by RIRDC in December 2015
Print-on-demand by Union Offset Printing, Canberra at www.rirdc.gov.au
or phone 1300 634 313

Foreword

All agricultural industries must address the challenge of prices rising faster for inputs or costs, than for outputs or products (this is commonly referred to as the cost-price squeeze). Continuous productivity improvement is needed to meet this challenge. A key strategy contributing to achieving continuous productivity improvement in all modern, viable industries is genetic improvement.

Identifying the individuals with the best genes is achieved in essentially all livestock industries by the use of advanced statistical methods to analyse pedigree and performance data. Identifying the individuals with the best genes is referred to as *genetic evaluation*.

This project explored the potential to apply such methods to breeding programs in the Australian honey bee industry. The results of this pilot project show that there is real potential for genetic improvement of production and health traits in Australian honey bees, and that there is scope for applying the advanced statistical techniques currently used in other industries.

The results highlight an exciting and valuable opportunity for Australian honey bee breeders. The report also flags ways implementation of genetic evaluation could be adapted to incorporate tackling major challenges such as Varroa.

This report is an addition to RIRDC's diverse range of over 2000 research publications and it forms part of our Honey Bee and Pollination R&D program, which aims to support RD&E that will secure a productive, sustainable and more profitable Australian beekeeping industry and secure the pollination of Australia's horticultural and agricultural crops.

Most of RIRDC's publications are available for viewing, free downloading or purchasing online at www.rirdc.gov.au. Purchases can also be made by phoning 1300 634 313.

Craig Burns
Managing Director
Rural Industries Research and Development Corporation

Contents

Tables

Figures

Executive Summary

What the report is about

This project explores the feasibility of applying modern techniques for statistical analysis (known as BLUP methods) to the genetic evaluation of honey bees.

Who is the report targeted at?

The report is most relevant to Australian honey bee queen breeders, and to industry and R&D managers.

Where are the relevant industries located in Australia?

Honey bee production is distributed across the agricultural regions of southern and eastern Australia (ie including SW Western Australia, and extending around the southern and south-eastern coasts up to southern Queensland).

Queen bee breeding is not concentrated in any particular region, but reflects the production areas.

The beneficiaries of the project, or more accurately of the implementation of the methods described, will be beekeepers across Australia, including those whose primary use of bees is for pollination services.

The principles and results outlined in this report are applicable to all queen breeding operations across Australia.

Aims/objectives

The aim of the project was to examine the potential for application of methods for data analysis used for genetic evaluation in most livestock (and increasingly, plant) industries world-wide.

Genetic evaluation – which in simple terms means identifying the individuals with the best genes for traits we are interested in, is the fundamental step in genetic improvement. The standard approach to this is now through use of advanced statistical methods which use pedigree and performance data, and which produce estimates of the value of the genes of each individual in the recorded population.

Assessing the scope for using these methods in the Australian honey bee population meant identifying queen breeders with appropriate data, and analysing that data using the appropriate methods. If the analysis is feasible, that means that there is scope to apply the methods in honey bee breeding in Australia.

Methods used

This project is very simple in design and conduct:

- Drawing on advice from industry experts, identify any queen breeders who may have suitable data for analysis: data with some form of consistent and systematic recording of performance of hives, and with some form of pedigree information collected across generations

- With the approval of the breeders, analyse the data sets to produce estimates of queens' genetic merit (termed Estimated Breeding Values, or EBVs)

- Communicate the results to industry, under the guidance of industry leaders and RIRDC

Results/key findings

Following the key steps outlined in Methods used:

- we were able to identify two breeders with datasets that could be analysed
- analysis was successful in both cases – Estimated Breeding Values for Queens could be calculated
- the results obtained, and the opportunities they generate, have been communicated to industry via the NSW Apiarists' Conferences in 2014 and 2015.

No additional findings were obtained, but the results do confirm the opportunity for considerable and valuable genetic improvement in Australian honey bees. This project essentially confirms the feasibility of applying modern methods for genetic evaluation and improvement in Australian honey bees, and that means that improvement can be made in a wide range of aspects of honey bee production and health.

The key benefit from such improvement would be to offset the cost-price squeeze. Typically, input costs grow 3-5% faster than product prices, meaning that the gross margin per hive could be declining by around 3.5% per year (or approximately $7.50 per hive per year, assuming 50kg per year honey production, and $5 per kg to the producer).

Implications for relevant stakeholders for:

- industry
- communities
- policy makers
- others where relevant

The main implication for industry is that it can be confident that genetic improvement of Australian honey bees is feasible. Achieving it will require that at least some queen breeders systematically collect data on queen pedigree and hive performance, and that the data collected is appropriately analysed. The range of traits in which improvement is achieved will depend primarily on what traits are recorded.

For communities, the main implication of the project is that industry, perhaps with the assistance of the broader community, can do a great deal to reduce the risks of:

a) declining viability of honey bee-based enterprises (honey production and pollination services) by breeding productive and healthy bees, and

b) of catastrophic decline in honey bee populations due to disease

These risk mitigation benefits of honey bee genetic improvement have both tangible financial benefits, as well as "sense of well-being" and welfare benefits. The financial benefits can be estimated, but this is not easy for the non-financial benefits, even though they are likely to be very strongly appreciated by the wider community.

For policy makers, the main implications are:

a) that provided that a sufficiently long-term view is taken, there is real cause for optimism about the future viability of Australian honey bee populations and production – genetic improvement can make a very significant contribution to continually improving productivity and viability.

b) At the same time, achieving these highly desirable outcomes will require some sustained investment in R&D and most importantly in education/extension programs for queen breeders and beekeepers. A planned and coordinated 10-15 year strategy for genetic improvement research, development and extension would be very valuable and should be considered.

Recommendations

There are two key recommendations of the report and their targets.

a) Queen breeders should adopt cost-effective systems for recording pedigree and performance of their bees, and have the data analysed to produce Estimated Breeding Values (EBVs) of the relevant traits. Use of the EBVs in selecting queen mothers will ensure genetic progress.

b) Industry and supporting agencies such as RIRDC should develop a strategy for genetic improvement of the Australian honey bee population, including sustained investment in both R&D and extension/education support for queen breeders and beekeepers. There is a strong case for significant ongoing community support for such a strategy, as there are very significant public goods that would be generated by genetic improvement of honey bees.

Introduction

Genetic improvement simply means breeding better and better bees over time. It depends on two key steps:

> **Genetic evaluation:** this depends on collecting records of performance and (usually) pedigree, and analysing that data to identify the queens with the best genetic make-up for the traits of interest.
>
> **Genetic improvement:** this means selecting the animals with the best genetic makeup, or estimated to have the best genetic makeup, as parents of the next generation. This means that each successive generation is genetically better than the previous one, and over time, performance levels for all recorded traits can be improved.

An important distinction – phenotype and genotype

What we see and can measure or score for a hive is called the phenotype of that hive.

The phenotype is determined by the genetic makeup of the queen, of the drones to which she was mated, and to the interaction between the genetic makeup of the queen, the drones, and the resulting workers, and their environment. Their environment includes things like quantity and quality of food sources, weather, and exposure to disease or predation.

Phenotype is not usually a perfect indicator of the genetic merit of the bees in the hive – not all the differences we see amongst hives are due to genetic differences. So for example, if we have 2 hives, and one produces 60 kg of honey and the other 50 kg, and we use queens from the 60 kg hive to breed the next generation, we should not expect to see a +10 kg lift in honey production in the next generation. How much we see is a reflection of the proportion of the observed or phenotypic differences that are genetic, and also to the random effects of year, site, disease etc.

As long as some proportion of the observed differences are due to genetic differences, and we can analyse performance data to identify the hives with the best genes, we can use the genetically superior queens as parents, and by focussing on queens with better genes, make genetic improvement.

This distinction between the phenotype and the genotype, or genetic makeup, is the core of genetic improvement.

Importance: all livestock enterprises have to earn income from products (such as honey produced) to offset fixed and variable costs. For most livestock industries, prices received do not increase in real terms (the price received adjusted for inflation), but both fixed and variable costs do, resulting in what is usually known as the "cost-price squeeze". The effect of this is that profit from an enterprise declines in real terms over time. The only way to offset this decline is to make constant productivity improvement(s). These may come from increasing scale, better feeding and/or health management, or genetic improvement.

Genetic improvement is attractive because its effects are cumulative – each year's improvement builds on the previous, and so provided that improvement is sufficiently rapid, the cost-price squeeze can be offset. This is now the case in industries such as pigs and poultry, and the Australian lamb industry. Livestock industries that are not making genetic improvement are quite literally inevitably going backwards.

The basic requirements for genetic evaluation are:

- **Records of performance**

  For honey bees, these can include measured data such as weight of honey produced in a season, hive weight at the start of winter, assessments of health and behaviour traits (such as assessments of rapid hygienic behaviour), and scored traits such as those described below for the Horners' data.

- **Genetic parameters for the trait(s) being recorded**

  Ideally, these will be calculated from the data itself. The data must have pedigree and some record of performance. The key parameters are heritability and genetic correlation.

  Heritability refers to the proportion of observed differences between hives that is due to genetic differences. (This is the same as the proportion of differences in phenotypes that is due to differences in genotypes).The remaining proportion of the observed differences will be due to non-genetic factors, such as location, season, non-genetic differences in health, and other often undefinable factors.

  Heritability is expressed on either a 0-1 scale, or a % scale. If the heritability is 0 (or 0%), then there are no genetic differences for that trait in the population – all the observed differences are non-genetic. If the heritability is 1 (or 100%), then all the observed difference is due to genetic differences. Typically, heritability of production traits such as kg honey will be in the range 0.25-0.40 (25-40%), and for health and behaviour traits around 0.50 (50%).

  If the heritability is low or zero then little or no genetic improvement is possible in that population. If it is higher, then genetic improvement is possible.

Table 1: Heritability estimates for key traits from overseas literature

Trait	Heritability range
Honey production	15-60% (30% is a likely estimate for Australian bees)
Hygienic behaviour	50-60%
Varroa resistance traits	Very wide range

  Genetic correlation: as well as investigating the degree to which genes affect differences in performance for an individual trait, it is possible to estimate the extent to which different traits are affected by genes in common – this is known as the genetic correlation. There are some limited reports of the genetic correlations amongst these traits in the literature (not shown here).

  These estimates of heritabilities and genetic correlations could be used initially in genetic analysis of Australian data, but local estimates should be generated as soon as data on ideally up to 5,000 queens with records has been collected.

Once genetic parameters and performance data are available, genetic analysis to estimate the value of each queen's genetic makeup is possible. The method of choice for genetic analysis is **BLUP (**which stands for Best Linear Unbiassed Prediction) is the method of choice for analysis of performance records for genetic selection in essentially all farmed species.

BLUP methods take account of factors such as year and location of the hives, and they allow use of information from all relatives, including those born and/or recorded in different farms and years. This makes the estimation of animals' genetic merit significantly more accurate across a population of many programs and hives, even within a single program, the ability to identify the best queens across years enables more intense selection of the best queens from all those available and hence significantly faster genetic improvement.

Provided that breeders have accurate records of performance and pedigree data, then BLUP genetic evaluation can be conducted.

The result or output from such analysis is a set of Estimated Breeding Values (or EBVs) for the queens with records and pedigree. EBVs describe queens' genetic merit for the recorded traits, in units of production – so that EBVs for honey production would be in kg. A queen with an EBV of +1 kg for honey has genes that are worth an extra 1 kg of honey per hive. Such an animal would pass on half this superiority to its progeny.

A BLUP analysis will also include taking account of, or adjust for, non-genetic factors, which for honey bees will include:

- Year of the record, and length of the production season
- Site of the hive

In the absence of EBVs based on good records, it is very difficult to make consistent genetic progress for any trait.

AGBU has available software to generate BLUP EBVs for honey bee data. This software can handle essentially any size of data set (we use similar software for analysis of datasets of several million animals in other species).

Consultation with overseas colleagues as well as review of the literature, indicate that the methods outlined here are beginning to be used in bee breeding in a small number of European countries. The small scale of current implementation of the methods and approach point to bee breeding in most countries being relatively traditional in its approach (this refers to the low level of systematic performance recording and genetic data analysis; use of advanced reproductive techniques such as AI is not uncommon), but that where it is being applied, it works.

Objectives

The project will generate 3 outputs and outcomes which will contribute to genetic improvement of honey bees in Australia:

a) Genetic parameters for a range of production and health traits, likely representative of parameters for the wider Australian population. These parameters will include heritabilities and genetic correlations, knowledge of which is essential to achieving rapid and appropriately balanced genetic improvement.

b) Estimated genetic trends for all traits measured in the program over at least 5 and up to 25 years. This information will reinforce confidence that genetic progress is possible, and will also provide the basis for estimating how much more rapid progress can be if modern genetic evaluation methods are applied.

c) A documented example of actual progress achieved and methods used to achieve it, to inform all Australian breeders. This will include simple, clear recommendations on what data should be collected and how to use the results of genetic evaluation in making rapid genetic progress.

Methodology

In this project, 2 datasets were identified:

a) Data collected by the Horner family, including records for 4 subjectively scored traits, for approximately 200 hives within 1 breeding line over 3 cycles of breeding and production (6 years). The traits recorded were:
   - CB: chalk brood (no variation was observed for this trait, and so no analysis is possible)
   - W: general rating of the "value" or work of the hive
   - Be: bee size
   - Br: brood viability

b) Data collected by Lindsay Bourke for the Australian Queen Bee Assessment Program, consisting of data on varying numbers of hives per line, for 14 lines, over 3 cycles of production. In a separate subset of this data, the hives in were scored for Rapid Hygienic Behaviour. This dataset was not made available for this project, but a brief analysis of the summary results is included.

In both datasets, analysis was conducted using BLUP software written by AGBU. The model for analysis fitted:

- Line or queen
- Year of the record

In the data provided by the Horner family, a queen pedigree was constructed across the years of data provided. In the data from the Australian Queen Bee Assessment Program, line was fitted, as well as sample (or hive) within line. No across-years trend in line merit was assumed at this stage – meaning that queens supplied from a particular line but in different years were assumed to come from the same genetic distribution.

Data collection and analysis

Data provided by the Horner Family

This dataset included records on approximately 200 hives (queens), scored by either Joe or Wayne Horner for Chalk Brood (CB), general Work Value of the hive (W), Bee Size (Be), and Brood Viability (Br). Each trait was scored on a 1-5 scale. All were from the "C" line maintained by the Horners.

Each hive scored is the phenotype of the queen in that hive. The phenotype is the performance that can be observed. In genetic evaluation, we are interested in understanding the genetic merit of the queen that underlies that phenotype.

Queen pedigree was constructed by tracing the queen mother of each queen back through the data. At this stage, data on the drone cloud mated with each queen has not been analysed (methods are available to account for multiple "sires" of a hive).

The first step in analysis was to estimate the heritability of each trait. As no variation was observed in chalk brood (CB) – all hives scored the same – no analysis is possible for this trait.

The heritabilities for the other 3 traits were:

- Work Value (W): 0.20 or 20%
- Bee Size (Be): 0.38 or 38%
- Brood Viability (Br): 0.40 or 40%

For all 3 traits, these results indicate that there is scope for selection to improve the trait.

From this analysis, Estimated Breeding Values (EBVs) for the queens can be calculated, for each of these 3 traits. A complete listing is provided in Appendix 3, but the following table summarises the results:

Table 2: Example Estimated Breeding Values (EBVs)

queen	EBV for Work Value	EBV for Bee Size	EBV for Brood Viability
C1	0.01	0.19	-0.37
C1a	0.27	1.46	1.20
C1b	0.33	-0.05	-0.06
C2c	0.41	0.41	0.44

C1 is genetically below average for Brood Viability (Br). It has an EBV of -0.37 for this trait, meaning that its genes are worth 0.37 less on this scoring scale than the average queen of all those scored and analysed. It will pass on half this EBV to its daughters (and sons).

C1a is genetically above average for Bee size (Be) and Brood Viability (Br). Its EBVs for these 2 traits are 1.46 and 1.20 respectively. Daughters (and sons) of C1a will produce hives with larger bees and higher brood viability than the average queen in this dataset.

C2c is genetically above average for Work Rating (W) with an EBV of 0.41.

The key messages from this analysis are:

- There is evidence for statistically significant genetic differences amongst queens in the C line for Work Value, Bee Size and Brood Viability.
- This provides scope for selecting to improve one or more of the traits.
- There was no clear evidence of any trend in the EBVs over time in this small dataset. No strong conclusions should be drawn from this as to whether genetic improvement is actually being achieved for these traits.

Comments:

The most important outcome of this small analysis is that we were able to construct queen pedigrees for the dataset, thus allowing a genetic analysis and producing EBVs.

The Horner family are implementing a system for capturing data on hive honey weight, which combined with data collected on the scored traits, will allow BLUP analysis of all their data.

Data provided by Lindsay Bourke, from the AQBBA Queen Assessment Project

Lindsay Bourke provided datasets on honey production and hive winter weight for lines being assessed through the AQBBA Queen Assessment project.

This data was analysed for the effect of line on honey production. This approach is based on treating the queens provided by the various programs as representatives of each particular line, and assuming no genetic improvement within lines across years.

This assumption must be made because in the absence of individual queen pedigree, it is not possible to disentangle the non-genetic effects of year from the genetic effects of queens within a line within a particular year. Also, we have no knowledge as to how queens supplied from each line were selected within those lines, so we are assuming a random sampling.

The analysis of the dataset showed very large differences between:

- years (a non-genetic effect)

- lines

- hives within lines

Table 3: Data provided by Lindsay Bourke, from the AQBBA Queen Assessment Project

Factor	Effect or range of effect	Comment
Overall average	84 kg per hive	
Season of production	2008: +41.5 kg 2010: - 21.8 kg 2013: -29.3 kg	Large differences between years
Line	From +5.8 kg honey to – 28.6 kg honey	Large range between lines in early years, smaller later – but some lines no longer represented
Hive	From +10 kg honey to – 15 kg honey	This is line + hive "EBV"

The effect of different years is shown by the "season of production" row – honey production was on average much higher in the 2008 season than either 2010 or 2013.

The effect of line is also large – a range of 34.4 kg after adjusting for season. It is clear from the dataset that lines which performed more poorly in early years were less likely to continue providing queens in later years.

Finally, within each line, a number of queens (or hives) were evaluated. After accounting for the effect of season and line, there are also differences between hives, which are partly genetic. The last row in the table above show the range observed in an approximate EBV for hives, which combines the estimated effect of their line, with their own effect, after adjusting for season.

The AQBBA also supported an assessment of Rapid Hygienic Behaviour carried out by Jody Gerdts (Bee Scientifics) and Lindsay Bourke, work supported by the Wheen Foundation and the Australian Honey Bee Industry Council.

Rapid Hygienic Behaviour is potentially important in disease resistance, and so it is of interest:

a) whether there is evidence for a genetic component to the trait, and

b) to explore the relationship between this trait and production. If high production was strongly correlated with poor RPH, this would make it very hard to select for productive, hygienic bees.

We sought permission to access the score data per hive, but this was not granted. The results presented in their reports suggest that there are line differences in RPH, and a suggestion that selection of queen mothers and drones from better performing lines resulted in some improvement (although it is not clear how the effect of year was accounted for in this). Together, these observations and reports from overseas literature suggest that there is almost certainly a genetic component to RPH – meaning that observed differences would be heritable and selection would be possible.

As a very simple assessment of any relationship between the honey-producing ability of lines, and their RPH, the EBVs calculated for honey weight can be plotted against the line means reported for RPH:

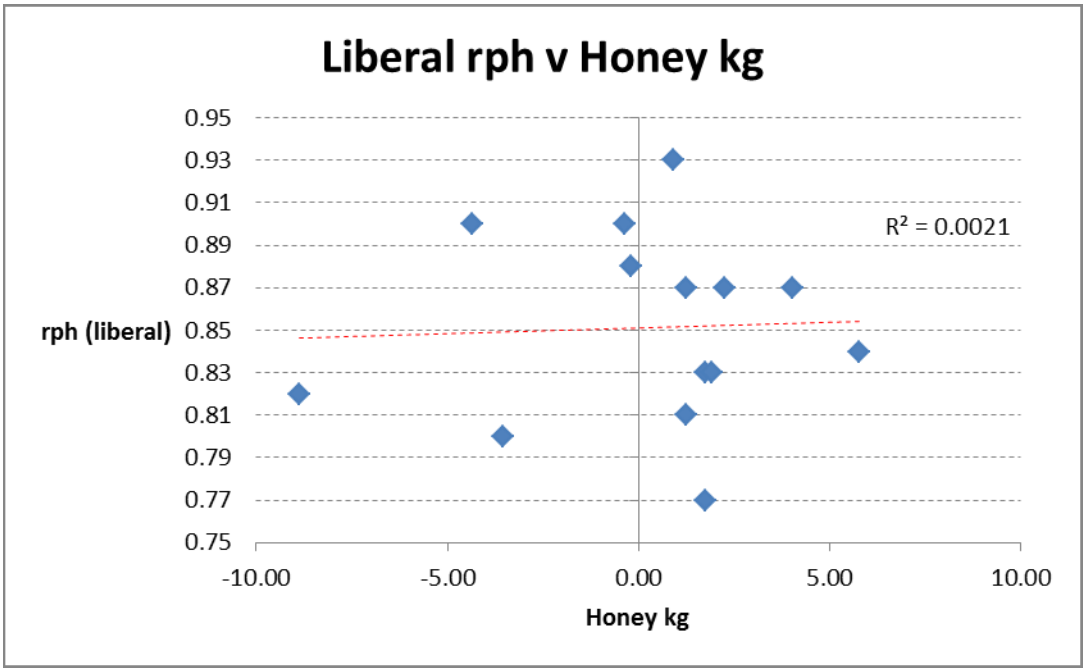

Figure 1: Data for RPH is extracted from the report "Hygienic Behaviour in the AQBBP June 2015 Update"

This chart suggests that there is no obvious relationship between line performance for Honey production and for RPH. This means that it should be possible to select simultaneously for honey production and RPH, as long as both traits are objectively scored.

Comments:

The AQBBA dataset provides <u>some</u> information on likely genetic differences between lines and hives within line. It does so because the hives are located and recorded in common locations.

At the same time, caution must be exercised in interpreting and using the results. We have no information on how the queens from each line were selected for entry into the test, and there are varying numbers of hives (queens) per line represented. The variation in representation can be handled statistically, but we have no way of knowing whether the queens supplied are statistically representative of their parent line.

Following on from this, there is a real challenge in judging what to do with the results that are obtained. The large range in hive performance within lines indicates that there is scope for selection within and between lines – using the best performing queens as parents (or those with the best EBVs), would allow real improvement.

The best approach would be to consistently record performance and pedigree data within the lines as well as in the AQBBA Queen Assessment program, and then combine the data from each line, with the results of the assessment program, to enable breeders to identify the genetically best queens across the programs.

The simple analysis of the relationship between honey production and RPH in the AQBBP suggests no unfavourable relationship between these two traits – meaning that selection for the two traits together is feasible.

Results

Results are summarised against each objective.

a) Genetic parameters for a range of production and health traits, likely representative of parameters for the wider Australian population. These parameters will include heritabilities and genetic correlations, knowledge of which is essential to achieving rapid and appropriately balanced genetic improvement.

The data sets obtained and used in the project were not large enough to estimate genetic parameters with great precision, but in both cases, the results of analysis allow some conclusions regarding these parameters.

In the Horners'data, the following heritability estimates could be calculated:

- Work Value (W): 0.20 or 20%
- Bee Size (Be): 0.38 or 38%
- Brood Viability (Br): 0.40 or 40%

While these estimates have large standard errors (meaning that the true value of the parameter in the population could vary widely around the estimates), these are consistent with values evident from the international literature.

Most importantly, the estimates support the hypothesis that there are genetic differences for these traits within this sample of the Australian honey bee population, meaning that there is scope for selection to improve these traits.

In the case of the AQBBA data, we were not able to access the full dataset which would have allowed us to estimate the proportion of variation in weight of honey or RPH that was due to line (which would indicate the extent of genetic differences between lines). However, the public reports from that work suggest that line differences were statistically significant, indicating significant genetic differences.

b) Estimated genetic trends for all traits measured in the program over at least 5 and up to 25 years. This information will reinforce confidence that genetic progress is possible, and will also provide the basis for estimating how much more rapid progress can be if modern genetic evaluation methods are applied.

In the Horners' data, Estimated Breeding Values (EBVs) were calculated for the 3 traits for which heritability was not zero. Within the EBVs, no clear trend across years was apparent. At face value, this would be interpreted as meaning that no genetic progress is being made for these traits in this program. However, this interpretation is not appropriate, as we simply do not have enough data to draw any statistically significant conclusion on this point. The availability of EBVs would certainly make it easier for the breeder to make progress.

Similarly, in the AQBBA data there is no clear evidence of genetic progress across time, but again, the dataset is not large enough to draw any statistically significant conclusion on this. In addition, without information on how the queens that were submitted were selected from within their parent lines, extreme caution should be used in drawing any conclusions about line merit.

These observations highlight a very important overall message from this project: while there is good evidence that there are genetic differences for a range of traits in the populations examined, the need for larger volumes of systematically recorded data is critical. Genetic improvement in any species depends on having enough data to reliably estimate genetic differences, and enough individuals to allow scope for selection. This need for data is almost certainly the most significant challenge the Australian industry must overcome if it is to exploit the very real opportunities for valuable genetic improvement.

c) A documented example of actual progress achieved and methods used to achieve it, to inform all Australian breeders. This will include simple, clear recommendations on what data should be collected and how to use the results of genetic evaluation in making rapid genetic progress.

From the data that we were able to obtain, a documented example of actual progress achieved is not possible at this stage. However, it is possible to summarise recommendations for recording and selection.

The steps in successful genetic improvement of honey bees

a) Recording:

The principal recommendation is simple:

1. To record pedigree (at least queen mother) for each hive:
- ID of the queen in that hive
- If possible, Queens that were parents of the drone cloud, or queen that was the mother of the drone if AI is used

2. Record the performance of each hive, including:
- ID (a number, ideally including the year, and the site of the hive if multiple sites are used in the breeding program, and the number for that particular hive)
- Date of putting the hive out
- Weight of honey produced (in kg)
- Days of honey production, or date at which hive or its honey is weighed
- Any scored traits (such as used by the Horner family)
- Any health or hygiene traits recorded (such as Chalk Brood, Rapid Hygienic Behaviour etc)

b) Genetic evaluation

Data collected in the form outlined above can be analysed to estimate genetic merit of the queens, as EBVs, for the recorded traits.

c) Selecting the best queens and drones

The EBVs should be used to identify the best queens to be parents of the next generation. If honey production is the only trait of interest, then simply pick the queens with the highest EBVs for honey production to be the queen (and potentially drone) mothers.

If more than one trait is being recorded and evaluated, a variety of ways of combining information across the traits is available. Outlining these in detail is beyond the scope of this report, but information can be provided on request to AGB U.

If pedigrees and EBVs are available, simple tools are available that can be used to actually design the breeding program to control inbreeding while maximising genetic improvement. Details are available upon request from AGBU.

Implications

Assessment of the impact of the outcomes on industry in Australia (where possible provide a statement of costs and benefits) This project is a pilot project, scoping the feasibility of applying modern genetic evaluation and improvement techniques in Australian honey bees. As such, any assessment of the impact of outcomes can only give an outline of what is possible, and provide some comment on what hurdles and risks might be involved.

The primary message of the project is that implementation of modern genetic evaluation and improvement methods are feasible. Can we estimate the benefits of such implementation?

To do this, a simple benefit: cost model can be developed, with the following assumptions (the model assumptions can be varied readily).

Assumptions:

- Number of honey-producing hives in Australia each year is 500,000
- Each hive produces on average 50 kg of honey, with a phenotypic standard deviation of 10% or 5 kg
- A kg of honey is worth \$4 to the producer
- The heritability of honey production is 25%
- Each commercial hive has only 1 year of production

To estimate the benefit-cost of genetic improvement, 3 steps are required:

a) **Calculate a realistic rate of genetic progress:**

The formula for rate of genetic progress is

Rate of progress = selection intensity x heritability x standard deviation

For honey production per hive per year, and assuming modest selection intensity, such that the queens selected to be queen mothers are from the best 33% of the population on EBV for honey production, this equates to

Rate of progress = 1.1 x 25% x 5 kg = 1.375 kg honey per hive, per year

(The value 1.1 is obtained from the normal distribution, or bell curve, and is the average deviation from the mean value of the best 33% of a population, expressed in standard deviation units)

This value is the amount by which we can be confident of being able to improve honey production per hive per year, through a sound genetic improvement program.

For this discussion of potential benefits from genetic improvement, we will focus solely on honey production. However, it is straightforward to extend the modelling, and a genetic improvement program, to other traits and to determine the overall economic value of improvement across the set of traits being tackled.

b) Model an adoption rate

It is implausible that all queens being bred would immediately be coming from breeding programs using modern genetic improvement methods. For this exercise, we will assume a 1% per year growth in the number of queens coming from such programs. Accordingly, over a 25-year period, if we start from 0 adoption, we will have reached 25% or 125,000 hives by year 25.

It is assumed that the commercial bee population lags 1 year behind the genetic merit of the queen programs.

c) Model the flow of benefits, incorporating discounting the time value of money:

In assessing any investment, including developing a genetic improvement program for honey bees, it is standard practice to discount future returns and costs to the present day. In Australian Rural R&D management, the standard discount rate for this purpose is 7%.

With these assumptions, we can model the flow of benefits from genetic improvement – these are shown in both a table and a chart below.

Table 4: Flow of benefits from genetic improvement

Year	Merit of Queens, in kg of honey per hive per year	Merit of Commercial Population	% Adoption	Value of Genetic Improvement	Discounted Value of Genetic Improvement	Cumulative Value
2016	0.00	0	0%	$0	$0	$0
2017	1.38	0.00	1%	$0	$0	$0
2018	2.75	1.38	2%	$55,000	$48,039	$48,039
2019	4.13	2.75	3%	$165,000	$134,689	$182,728
2020	5.50	4.13	4%	$330,000	$251,755	$434,484
2021	6.88	5.50	5%	$550,000	$392,142	$826,626
2022	8.25	6.88	6%	$825,000	$549,732	$1,376,358
2023	9.63	8.25	7%	$1,155,000	$719,276	$2,095,634
2024	11.00	9.63	8%	$1,540,000	$896,294	$2,991,928
2025	12.38	11.00	9%	$1,980,000	$1,076,989	$4,068,917
2026	13.75	12.38	10%	$2,475,000	$1,258,164	$5,327,082
2027	15.13	13.75	11%	$3,025,000	$1,437,156	$6,764,237
2028	16.50	15.13	12%	$3,630,000	$1,611,763	$8,376,001
2029	17.88	16.50	13%	$4,290,000	$1,780,197	$10,156,198
2030	19.25	17.88	14%	$5,005,000	$1,941,025	$12,097,224
2031	20.63	19.25	15%	$5,775,000	$2,093,126	$14,190,349
2032	22.00	20.63	16%	$6,600,000	$2,235,648	$16,425,998
2033	23.38	22.00	17%	$7,480,000	$2,367,976	$18,793,974
2034	24.75	23.38	18%	$8,415,000	$2,489,695	$21,283,669
2035	26.13	24.75	19%	$9,405,000	$2,600,561	$23,884,230
2036	27.50	26.13	20%	$10,450,000	$2,700,479	$26,584,708
2037	28.88	27.50	21%	$11,550,000	$2,789,476	$29,374,185
2038	30.25	28.88	22%	$12,705,000	$2,867,686	$32,241,870
2039	31.63	30.25	23%	$13,915,000	$2,935,326	$35,177,196
2040	33.00	31.63	24%	$15,180,000	$2,992,686	$38,169,882
2041	34.38	33.00	25%	$16,500,000	$3,040,111	$41,209,993

In the table:

- Merit of the queen population is in comparison to merit in 2015, ie it is the difference in genetic merit from the starting year.

- Similarly, merit of the commercial population is the difference from 2015, in extra honey per hive.

- Discounted value of genetic improvement is the total value of genetic improvement, expressed in 2015 dollars, using the 7% discount rate. The cumulative value of genetic improvement is the sum of each year's discounted value – in economic terms, this is the Present Value of the genetic improvement to that year.

The cumulative present value can be shown graphically:

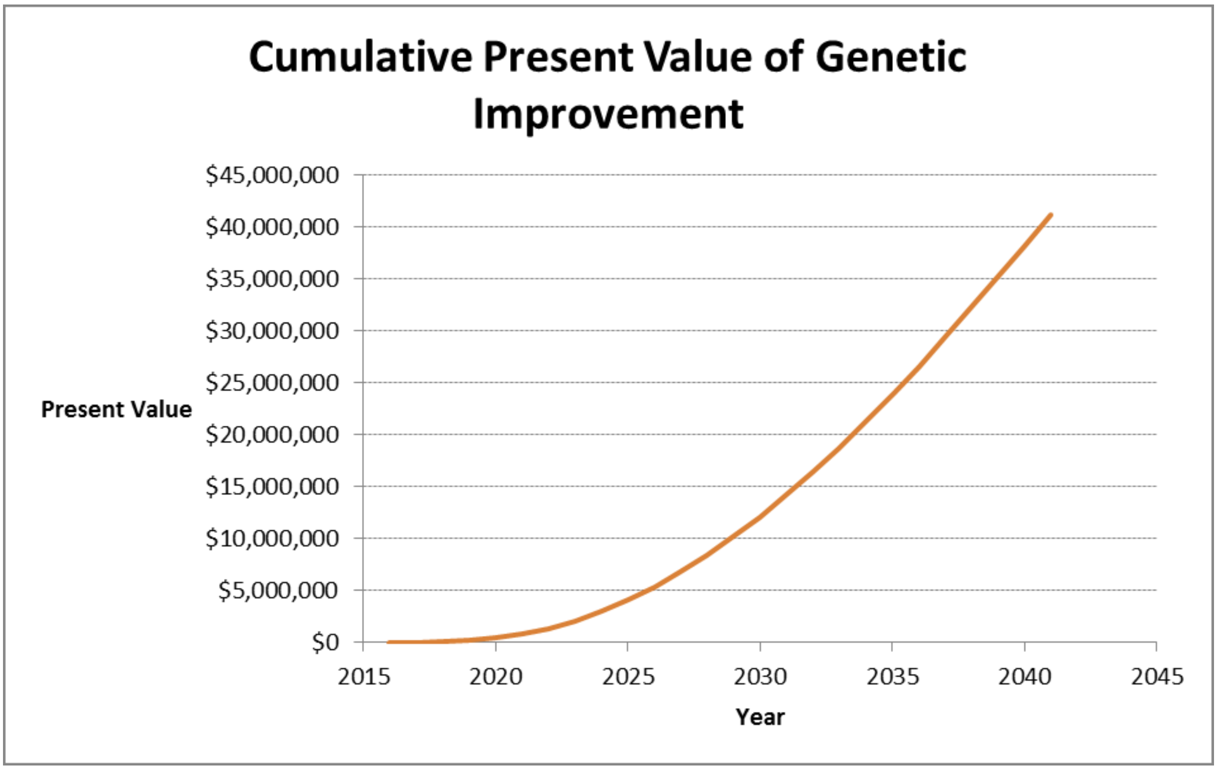

Figure 2: Cumulative Present Value of Genetic Improvement

This chart simply shows how much the genetic improvement that can be achieved each year into the future, is worth today. Using simple but realistic assumptions, the model shows that an industry genetic improvement program achieving realistic gains and with only very slow adoption, can generate **gains over a 25-year period worth approximately $41m in today's values.**

What are the risks and hurdles associated with this prospective return to industry and community?

a) Cost of implementation: there will be some costs of recording and analysis at the queen breeder level. Without detailed consultation with breeders, it is not easy to estimate these, but it seems unlikely that they would exceed $250,000 per year across all breeding programs – this estimate is based on extra time for the labour of recording, and for recording equipment, plus a modest amount for data analysis.

b) Costs of R&D: some investment in further R&D, and into extension and training for queen breeders and beekeepers, would be needed. The scale and nature of such investment would need to be defined by industry, but could be in the range $250-500,000 per year (note that this is not meant to imply that the honey bee industry could or should fund such investment on its own – rather, to provide an indicative estimate of an ongoing investment.)

If these two factors were included in the benefit-cost analysis model, at $250,000 for recording and analysis per year, and $500,000 for RDE per year, the cumulative net present value to 2041 becomes $32m rather than $41m.

This benefit-cost estimation is deliberately cautious in regards to both the rate of genetic improvement and the rate of adoption. Overall present value, both of benefits and also net of investments, is very sensitive to these 2 parameters. For both, it is very likely that industry could achieve better levels than are modelled here, which would in turn mean significantly larger returns to industry and community.

One point that reinforces this message regarding potential benefits is that the approach here does not include any benefit from mitigating losses from Varroa. It is certain that genetic improvement could assist in both preparing the honey bee population to be better placed to minimise losses from Varroa, and to recover more quickly should it arrive. If these benefits were included, the benefits of investment into genetic improvement would be many times larger than those estimated here.

There is an important final point for consideration around estimation of benefits. It is important for industry and the community to think carefully about who bears the risks (and costs) involved in implementing genetic improvement, and who gets the benefits. In most livestock industries, the breeding sector (in honey bees, this means the queen breeders) bear essential all the implementation costs, but only obtain a small share of the total benefits generated. This frequently generates a market failure in that breeders simply cannot afford to invest in the recording work required to maximise potential value of genetic improvement. This market failure is often tackled in part by having collective investment, such as through a Research and Development Corporation, assist with R&D costs, but increasingly industries are exploring ways to minimise the under-investment in recording by having some form of co-investment between breeders and others in the industry. Given the public good associated with having a healthy and productive Australian population of honey bees, there seems an overwhelming case for such assistance in this industry.

Recommendations

There is considerable value in encouraging and assisting the development of effective genetic improvement programs for the Australian honey bee industry. The project has demonstrated on a very small scale that performance data can be analysed to estimate genetic merit in Australian honey bee breeding programs.

It has proven difficult to obtain data – that provided by the Horner family and by Lindsay Bourke has been invaluable, but breeders' power to drive genetic progress would be enhanced if more data can be collected and analysed.

Based on the results obtained, there is no reason why breeders could not be aiming to improve average hive production by at least 1 kg per year indefinitely, while simultaneously making steady improvement in hygienic behaviour, and other traits such as temperament and hive over-wintering weight.

Keeping good records of pedigree will greatly assist in genetic improvement, but the most fundamental and essential resource is data. Without accurate and consistent records of hive performance, kept year to year, no significant improvement is possible.

It is strongly recommended that breeders and industry consider ways to facilitate recording, including possibly appointing someone to assist with training and possibly developing simple recording systems. Consistency between breeders in how records are kept, and in ID systems, would be advantageous.

Genetic analysis:

Despite the amount of work involved in collecting data on honey bees, the resulting datasets are not large (by comparison with those used in other livestock breeding).

This means that analysing data, as long as its format is understood and consistent, is straightforward and should impose no significant cost for an individual breeder or industry.

Further R&D

The most significant R&D opportunity for the industry is to identify ways to combine data from several breeding programs (ideally all breeding programs that are keeping records) for joint genetic analysis.

The relatively small size of the individual breeding programs (in terms of numbers of queens and queen mothers used per year), means that scope for selection within a line or program is limited, and that unless outside genetic material is introduced, there is some risk of inbreeding. Outside introductions are a simple means of managing inbreeding, but should ideally be made backed by some objective information (such as EBVs).

It is not clear how widespread or consistent is recording of pedigrees – it seems limited – so such collaboration across programs is inevitably constrained.

One R&D step that could help in this regard is to map the genetic relationships across programs through use of DNA analysis. This activity is outside the scope of this project, but a proposal for funding support is being developed.

It is not appropriate for this report to recommend such a proposal be supported, but exploring ways to enable across-program genetic analysis is strongly recommended.

Appendix 1: The steps in successful genetic improvement of honey bees

a) Recording:

The principal recommendation is simple:

1. To record pedigree (at least queen mother) for each hive:
- ID of the queen in that hive
- If possible, Queens that were parents of the drone cloud, or queen that was the mother of the drone if AI is used

2. Record the performance of each hive, including:
- ID (a number, ideally including the year, and the site of the hive if multiple sites are used in the breeding program, and the number for that particular hive)
- Date of putting the hive out
- Weight of honey produced (in kg)
- Days of honey production, or date at which hive or its honey is weighed
- Any scored traits (such as used by the Horner family)
- Any health or hygiene traits recorded (such as Chalk Brood, Rapid Hygienic Behaviour etc)

b) Genetic evaluation

Data collected in the form outlined above can be analysed to estimate genetic merit of the queens, as EBVs, for the recorded traits.

c) Selecting the best queens and drones

The EBVs should be used to identify the best queens to be parents of the next generation. If honey production is the only trait of interest, then simply pick the queens with the highest EBVs for honey production to be the queen (and potentially drone) mothers.

If more than one trait is being recorded and evaluated, a variety of ways of combining information across the traits is available. Outlining these in detail is beyond the scope of this report, but information can be provided on request to AGBU.

If pedigrees and EBVs are available, simple tools are available that can be used to actually design the breeding program to control inbreeding while maximising genetic improvement. Details are available upon request from AGBU.

Appendix 2: EBV results for C Line Queens, Horner Family records

a) For Working Value (“W”)

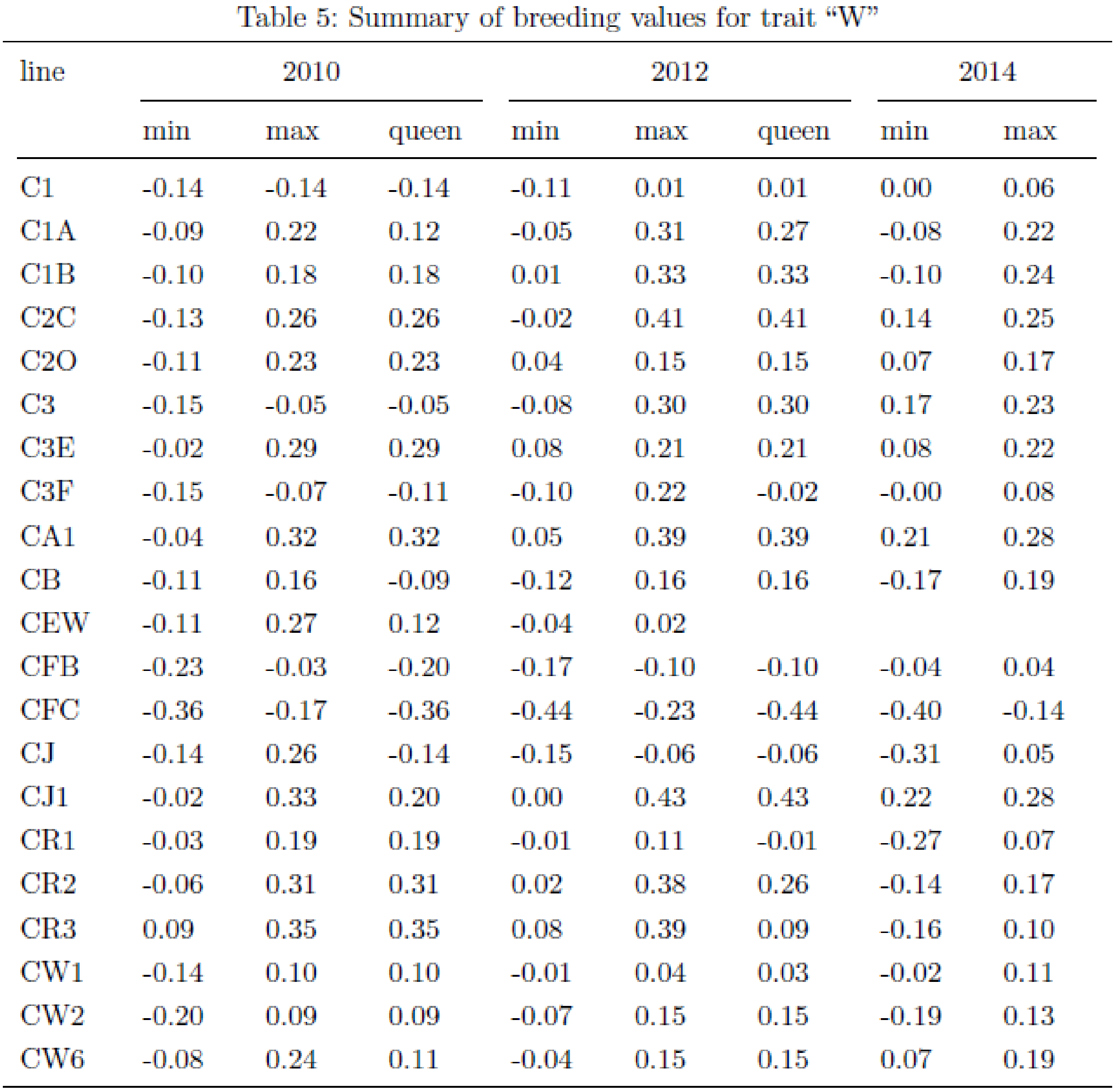

Table 5: Summary of breeding values for trait “W”

line	2010			2012			2014	
	min	max	queen	min	max	queen	min	max
C1	-0.14	-0.14	-0.14	-0.11	0.01	0.01	0.00	0.06
C1A	-0.09	0.22	0.12	-0.05	0.31	0.27	-0.08	0.22
C1B	-0.10	0.18	0.18	0.01	0.33	0.33	-0.10	0.24
C2C	-0.13	0.26	0.26	-0.02	0.41	0.41	0.14	0.25
C2O	-0.11	0.23	0.23	0.04	0.15	0.15	0.07	0.17
C3	-0.15	-0.05	-0.05	-0.08	0.30	0.30	0.17	0.23
C3E	-0.02	0.29	0.29	0.08	0.21	0.21	0.08	0.22
C3F	-0.15	-0.07	-0.11	-0.10	0.22	-0.02	-0.00	0.08
CA1	-0.04	0.32	0.32	0.05	0.39	0.39	0.21	0.28
CB	-0.11	0.16	-0.09	-0.12	0.16	0.16	-0.17	0.19
CEW	-0.11	0.27	0.12	-0.04	0.02			
CFB	-0.23	-0.03	-0.20	-0.17	-0.10	-0.10	-0.04	0.04
CFC	-0.36	-0.17	-0.36	-0.44	-0.23	-0.44	-0.40	-0.14
CJ	-0.14	0.26	-0.14	-0.15	-0.06	-0.06	-0.31	0.05
CJ1	-0.02	0.33	0.20	0.00	0.43	0.43	0.22	0.28
CR1	-0.03	0.19	0.19	-0.01	0.11	-0.01	-0.27	0.07
CR2	-0.06	0.31	0.31	0.02	0.38	0.26	-0.14	0.17
CR3	0.09	0.35	0.35	0.08	0.39	0.09	-0.16	0.10
CW1	-0.14	0.10	0.10	-0.01	0.04	0.03	-0.02	0.11
CW2	-0.20	0.09	0.09	-0.07	0.15	0.15	-0.19	0.13
CW6	-0.08	0.24	0.11	-0.04	0.15	0.15	0.07	0.19

Each row of this table corresponds to a sub-line within line C. For each year, the minimum and maximum breeding values for hives (queens) within that line is shown, and the columns headed “Queen” show the EBV of the Queen selected from within that line in that year.

For instance, using the second row, which is for sub-line C1A, in 2010, the lowest EBV hive had a value of -0.09 for W, the highest had an EBV of +0.22, and the hive selected to provide the queen for that subline for the next generation had an EBV of +0.12.

There are instances where the queen chosen had the highest EBV for the trait, and others where she didn’t: this reflects 2 factors:

- Selection is based on more than just performance for W

- Selection was not based on EBVs (which might have changed actual selections if available)

b) For Bee Size ("Be")

Table 6: Summary of breeding values for trait "Be"

line	2010			2012			2014	
	min	max	queen	min	max	queen	min	max
C1	0.01	0.05	0.05	-0.16	0.19	0.19	-0.13	0.80
C1A	-0.30	0.54	0.54	0.09	1.46	1.46	0.29	1.31
C1B	0.06	0.20	0.20	-0.05	0.03	-0.05	-0.59	0.72
C2C	-0.23	0.56	-0.23	-0.15	0.41	0.41	-0.05	0.91
C2O	-0.54	0.40	0.40	-0.01	0.76	0.76	-0.01	0.46
C3	0.07	0.51	0.28	-0.05	0.94	0.08	-0.56	0.19
C3E	-0.44	-0.07	-0.12	-0.20	-0.10	-0.20	-0.61	0.19
C3F	0.06	0.15	0.15	0.01	0.52	0.52	-0.05	1.05
CA1	-0.28	0.01	-0.28	-0.66	-0.14	-0.36	-0.48	0.56
CB	-0.33	0.41	0.21	-0.23	0.11	-0.23	-0.34	-0.28
CEW	-0.02	0.43	-0.02	-0.12	-0.02			
CFB	-0.54	-0.02	-0.08	-0.15	1.27	1.27	0.29	1.19
CFC	-0.08	0.12	-0.08	-0.22	-0.09	-0.22	-0.29	0.19
CJ	0.02	0.17	0.17	-0.14	-0.03	-0.14	-0.61	0.64
CJ1	-0.20	1.23	1.23	0.32	1.24	1.03	0.01	0.63
CR1	-0.49	0.36	0.36	-0.05	0.53	0.53	0.34	0.96
CR2	-0.68	-0.15	-0.25	-0.18	0.71	-0.04	-0.64	0.74
CR3	-0.78	-0.70	-0.78	-0.75	-0.29	-0.75	-0.84	-0.07
CW1	0.20	0.90	0.90	0.24	0.73	0.73	-0.32	0.44
CW2	-0.30	1.09	1.09	0.30	1.32	1.32	0.22	1.20
CW6	-0.57	-0.07	-0.20	-0.36	-0.13	-0.36	-0.83	-0.01

c) For Brood Viability ("Br")

Table 7: Summary of breeding values for trait "Br"

line	2010			2012			2014	
	min	max	queen	min	max	queen	min	max
C1	-0.28	0.11	0.11	-0.37	0.02	-0.37	-0.53	0.11
C1A	-0.01	0.74	0.74	0.26	1.20	1.20	0.68	0.80
C1B	-0.30	0.39	0.19	-0.13	0.02	-0.06	-0.31	0.22
C2C	-0.24	0.48	0.29	-0.03	0.44	0.44	-0.19	0.20
C2O	-0.33	0.21	0.21	-0.13	0.30	0.30	0.25	0.55
C3	-0.02	0.84	0.84	0.16	0.96	0.95	-0.01	0.69
C3E	-0.38	-0.02	-0.02	-0.48	-0.01	-0.48	-0.57	0.01
C3F	-0.39	-0.25	-0.25	-0.28	0.79	0.79	0.56	0.61
CA1	-0.21	0.51	0.51	0.11	0.83	0.83	-0.05	0.66
CB	-0.04	0.55	0.27	-0.26	0.25	-0.26	-0.52	0.19
CEW	-0.19	0.33	0.21	-0.07	-0.03			
CFB	-0.42	-0.21	-0.27	-0.29	0.40	0.40	-0.30	0.38
CFC	-0.53	-0.42	-0.42	-0.30	0.04	0.04	-0.33	-0.10
CJ	-0.33	0.24	0.24	-0.19	0.16	0.16	-0.20	0.40
CJ1	-0.33	0.38	0.05	-0.09	0.66	0.46	0.38	0.52
CR1	-0.57	-0.42	-0.57	-0.38	-0.08	-0.08	-0.37	0.24
CR2	-0.23	0.72	0.72	0.19	0.80	0.65	-0.17	0.41
CR3	-0.08	0.37	0.25	0.03	0.08	0.03	-0.38	0.30
CW1	-0.26	0.59	0.59	0.06	0.86	0.35	-0.31	0.50
CW2	-0.02	0.71	0.71	0.16	1.09	1.09	0.61	0.76
CW6	-0.48	0.36	-0.48	-0.40	-0.09	-0.17	-0.41	0.25

Glossary

EBV: **E**stimated **B**reeding **V**alue, an estimate from recorded data of the value of the individual's genes for the recorded trait(s).

The term Estimated refers to the fact that we cannot determine the value of the genes with 100% accuracy.

The term Breeding Value refers to the fact that the EBV tells us the value of each individual for breeding ie as a parent.

Genetic evaluation: the analysis of performance and pedigree data to calculate EBVs.

Genetic improvement: selecting better and better individuals as parents each generation and thereby breeding populations that are genetically better.

BLUP: **B**est **L**inear **U**nbiassed **P**rediction, is a method of statistical analysis that makes use of pedigree and performance data in calculating EBVs. It is the method of choice in all livestock and evaluation worldwide. "Best" refers to the fact that it is the best of a number of methods of calculating EBVs, in terms of its accuracy. "Linear" refers to the fact that it assumes a linear relationship between records and breeding value. "Unbiassed" refers to the fact that the method removes biases due to factors such as non-random mating and environmental differences among groups of recorded individuals.

Genotype: the genes of the individual

Phenotype: the observed performance of the individual. In simple terms, the phenotype is generated by the joint effects of the genotype (G) and the environment – things like nutrition, weather, disease etc (E), or in simple terms:

$$\mathbf{P = G + E}$$

DNA analysis: reading the DNA of the individual. There are a range of ways and levels of detail by which this can be done. DNA analysis gives more precise values for the relationships between individuals, and in some rare cases can help identify individual genes.

Genetic Evaluation of Australian Honey Bees using BLUP procedures

By R. G. Banks

December 2015

Pub. No. 15/104

RURAL INDUSTRIES
Research & Development Corporation

Phone: 02 6271 4100
Fax: 02 6271 4199
Bookshop: 1300 634 313
Email: rirdc@rirdc.gov.au
Postal Address: PO Box 4776,
Kingston ACT 2604
Street Address: Level 2, 15 National Circuit,
Barton ACT 2600

www.rirdc.gov.au