final report

Project code: L.GEN.1713

Prepared by: Russell Lyons and Imtiaz Randhawa
The University of Queensland

Date published: 25 March 2020

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

Improving the Australian Poll Gene Marker Test

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Abstract

Diagnostic testing for naturally poll cattle is becoming common practice in Australia because it helps avoid dehorning and disbudding in young calves. Selective breeding can rapidly increase poll frequency in herds, but it is confounded by concerns about the efficiency of poll testing assays, scurs inheritance, inbreeding and concerns of a loss in some production traits. Poll testing has evolved through microsatellites (MSAT) and single nucleotide polymorphism (SNP) genetic markers. Both common assays often give inconclusive outputs as “Not Determined” or “No Results”.

This project has investigated the poll testing assays across many common breeds in Australia, including both taurine and indicus cattle breeds, and developed a SNP based optimized poll testing (OPT) assay, which have increased the poll prediction efficiencies from 90.31% to 99.6% in the project data (n=20,636), and has a commercial poll testing success rate at 99.42% (n=70,031). The OPT has been rapidly adopted across the industry leading to greater accuracy and more confidence. Being compatible with genomic products, the test is also available at lower cost than the previous stand-alone tests.

Understanding of scurs genetics remains limited. While the project was unable to identify markers that are predictive of Scurs, and hence a test remains elusive, New insights suggest scurs development is influenced by a complex pathway involving non-genetic (sex and age of the animal) and genetic (poll alleles, polygenic and epistatic interactions) factors.

Comprehensive investigation of several production and reproduction traits was performed using 1,825,981 animals across multiple breeds that met the required criteria of having both EBVs and a published head status. Results confirmed that polledness will have no detrimental effects on herd productivity. Care is warranted to limit inbreeding effects for some breeds which have lower number of polled animals in breeding stocks.

Executive summary

Until recently, Australia utilised the microsatellite test for selection of polled breeding stock. This test was a world-leading technology and proved very successful in many breeds but was not without its limitations as highlighted in a recent MLA funded project L.GEN.1803 (2019) titled 'Poll DNA marker test improvements for industry sampling bias". Poll testing in Australia has undergone significant change in recent years, as the industry adopts new discoveries and technological innovations driven through the genomics revolution. Several different sequence variants of deoxyribonucleic acid (DNA) associated with polledness have been identified on chromosome 1 of the bovine genome, predominantly derived from taurine animals. Two large and unique mutations, named Celtic (Pc) and Friesian (Pf) respectively, are the predominant variants in most commercial cattle breeds globally. The diagnostic Single Nucleotide Polymorphism (SNP) markers for Pc and Pf have since been established and incorporated into poll tests to allow accurate poll testing using these SNP markers rather than the previous microsatellite markers. Advantages include no requirement for phenotypic data, increased accuracy of test results and potential cost savings for the industry through co-testing on compatible genomics platforms.

With the accumulation of sufficient data from a range of breeds, it was observed that the early iteration of the SNP-based poll test was very good for taurine breeds. However, it became evident that the test was less successful for Brahman and other largely indicus influenced breeds, with large numbers of unassignable animals being tested. This was deemed an impediment to the long-term utilisation of the poll test in Northern Australia. Other factors limiting the successful adoption of poll testing included uncertainly related to the scurs phenotype and existing dogma implicating polledness as associated with inferior genetics for reproductive traits.

The primary objective of this project was to develop an improved SNP-based poll test suitable for inclusion on future SNP chips and which would provide a cost-effective and reliable method for accurately assessing the poll status of all economically important cattle breeds in Australia. Furthermore, research would be undertaken to (1) identify markers associated with the scur phenotype in the hope of finding and incorporating a scurs diagnostic test for elimination or accurate prediction of carriers of the scur gene, and (2) determine whether there are negative effects of the poll gene on reproductive phenotypes.

This project identified mutations common in the Brahman, Brahman-influenced and Shorthorn breeds that prevented the standard SNP-based poll test used at that time from giving informative results for a subset (in some cases up to 20%) of animals tested. Once identified through sequencing, it was possible to amend the predictive algorithms to resolve this problem without the addition of new markers to the existing commercial products, making for a speedy implementation to the Optimised Poll Test (OPT) in a large commercial laboratory (Neogen Australasia). Importantly, assessment on a large number of animals across all common breeds (n = 20,636) that had been previously tested using SPTv1 demonstrated that reassessment with OPT did not change in the previous poll results (HH, HP or PP), ensuring no revision of historic results was needed. A total of 1999 previously unassigned (No Result) animals were retested using OPT, and 1990 (99.4&) were resolved. updated results were provided to the test submitters. The research findings of the OPT translation has been published (doi:10.1534/g3.119.400866) and is available for any service provider wishing to use this modified poll test.

Understanding of the scur phenotype in Australian herds has been improved through the current study. Relationships between scurs and sex have been confirmed, with scurs more likely in male animals. Heterozygous (HPc and HPf) animals have the highest likelihood of becoming scurred, although this is not a perfect association. Challenges identified during the project for this objective include difficulty in getting accurate and consistent scur phenotypes, in part due to the variable age of animals at first presentation of the phenotype and the heterogeneity of the phenotype which may

cause it to be easily confused with horns. For example, it was noted that phenotypes for some of the animals recorded as smoothly polled would later develop regular scurs with advancing age. Accurate phenotypes are a must for genome wide association studies (GWAS) and hence it was necessary to take a conservative approach, using only animals that had been confirmed as scurred by experienced staff using close inspection and palpation. The GWAS and Composite Selection Signals (CSS) analyses identified several regions of interest and several potential candidate genes for further investigation. There appears to be a statistically significant reduction in the rate of scurs for those animals carrying the Friesian (Pf) mutation versus the Celtic (Pc) mutation. Further research is recommended but this result, if confirmed, may provide important insight into the mechanism underpinning scurs development which remains to be fully understood.

Through comparison of large datasets of animals across multiple breeds obtained from BREEDPLAN comprising of fertility- and production-related EBVs, and also with known poll genotypes and concordant listed phenotypes, this study was able to confirm that there is no significant difference in fertility or other traits in polled versus horned animals. As such, polledness should not be considered as detrimental to breeding objectives across the industry.

Outcomes have the potential to significantly reduce costs of production associated with horned herds, including minimisation of production costs from dehorning and disbudding which is conservatively estimated at US$6/head (≈AUD$10) per head, elimination of post-treatment calf losses, reduction of animal and handler welfare risk and improved hide and meat quality.

Table of contents

1 Background

Cattle were historically naturally horned for defence from predators and to exert dominance for food and mating (Lundrigan 1996; Davis *et al.* 2011). Since domestication of cattle (*Bos taurus* and *Bos indicus*), the shape and size of horns have become a prominent feature in defining breeds, with different selection strategies leading to a greater diversity of phenotypes between the breeds (Zeder 2008; Ajmone-Marsan *et al.* 2010).

Generally, the head phenotype in cattle is classed into 3 categories:

- Horned when keratin coated permanent pointy protrusions anchored to the skull.
- Scurred when rudimentary horns are loosely attached to the skin rather than the skull.
- Polled when there is a complete absence of horns and scurs.

Absence and presence of horns or scurs is determined during prenatal development. However, the postnatal development of horns versus scurs is very difficult to differentiate at an early age because both present initially as free-floating horn-buds and subsequently, only the former fuses to the cranium (Dove 1935; Wiener *et al.* 2015).

With developments in the intensity and practices of raising beef and dairy cattle over time, the desirability of horns has changed (Zeder 2008; Ajmone-Marsan *et al.* 2010). Generally, in the modern commercial cattle industry, horned animals are less desirable because they pose potential hazards for other cattle and animals used for mustering (horses, dogs), feeding, handling and transport facilities, and farmworkers. Horns are associated with higher costs of on-farm and post-farm production and greater risk of reduced meat and skin quality (Bunter *et al.* 2013; Knierim *et al.* 2015; Schafberg and Swalve 2015).

While dehorning or disbudding is a common practice to reduce these issues on-farm, all surgical, chemical and cautery procedures are costly and cause varying degrees of pain and reduced animal welfare, morbidity, mortality and reduced productivity, with dehorning often resulting in exposure of the frontal sinus (Stafford and Mellor 2005; Neely *et al.* 2014; Knierim *et al.* 2015; Herskin and Nielsen 2018). As animal welfare becomes more important as production and consumer consideration, there is also economic and social license considerations. Therefore, polled cattle have become more desirable and an increasing trend of producing polled animals in various breeds in Australia have been noticed (Fig. 1)

Selective breeding for polled animals requires accurate early-in-life prediction of horn phenotype (Spurlock *et al.* 2014; Scheper *et al.* 2016). Horn phenotype is a qualitative trait controlled by genetics – 000483-9913 (OMIA 2019) – which has been mapped to bovine autosome 1 (BTA1) (Georges *et al.* 1993; Harlizius *et al.* 1997; Mariasegaram *et al.* 2012; Randhawa *et al.* 2016). The underlying genes and causal mutations for horns, scurs and polledness remain to be fully elucidated. Phenotypic penetrance suggests that the poll gene is dominantly inherited, i.e., PP (polled), pp (horned), with heterozygous animals usually polled but also commonly scurred (Pp).

The earliest Poll tests utilised in Australia were based upon combinations of 10 microsatellite markers (Fig. 2) mapped to the relevant region on a chromosome (Mariasegaram *et al.* 2012; Piper *et al.* 2014). Between 2010 and 2013 a single microsatellite marker (CSAFG29) was used as a commercial test for polledness, developed by CSIRO and the Beef Cooperative Research Centre (MLA Project B.AHW.0144 (2011); Henshall *et al.* (2011)). A second iteration utilizing all ten markers was implemented in 2014 to overcome issues including poor accuracy rates for *Bos taurus* breeds, such as Limousin and Brangus (39% and 38% accuracy respectively) (MLA Project B.AWW.0222 (2014); Piper *et al.* (2014)). This refinement was very successful in improving the accuracy of the test across a large range of breeds and became to standard Australian commercial test until 2018 and proved to be a world-leading initiative for the selection and breeding of increased polledness in the Australian herd for many breeds.

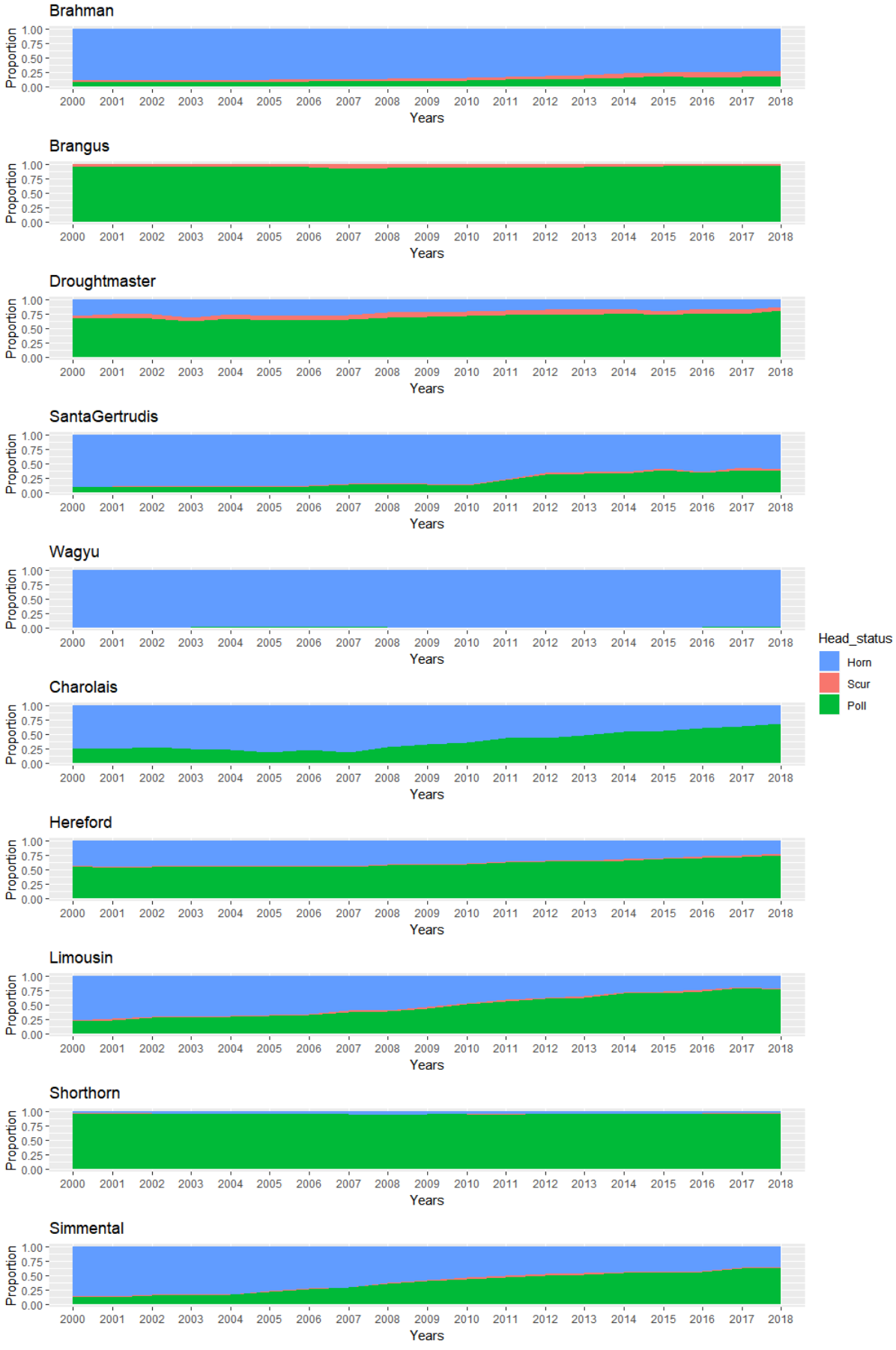


Fig. 1. Distribution and trends of head-status in 10 breeds of beef cattle during 2000-2018.

However, within-population instability and cross-population diversity of these obsolescent genetic markers make them vulnerable to diminishing accuracy over time and in non-ascertained populations. Data obtained from previous microsatellite testing of 20,534 animals (Fig. 3, Table 1) revealed that on average 11.7% of animals tested were unable to be assigned a poll status and were reported by UQ Animal Genetics Laboratory (UQAGL) as "Not Determined". While all breeds were affected, highest rates of Not Determined results were observed in Charolais (17%; n = 2,666), Shorthorn (16 %; n = 75)), Brangus (15.8%; n = 2,240) and Brahman (13.7%; n = 4,532). A small number of Wagyu cross animals were also tested (n=201) as some in the industry looked to the introduction of poll genetics into the herd, but the test proved uninformative for many animals with a Not Determined rate of 30.4%. This is unsurprising, as Wagyu were not part of the initial development. The success and limitations of this microsatellite test are well documented in MLA Project L.GEN.1803 (2019). As genome sequencing technologies have become more accessible and cost-effective, SNP-based testing has replaced microsatellite testing.

Through bovine genome sequencing initiatives, several different sequence variants of deoxyribonucleic acid (DNA) have been identified and shown to be associated with polledness in different breeds. Currently, these include Celtic (Pc), Friesian (Pf), Mongolian (Pm) and Guarani (Pg) mutations (Medugorac *et al.* 2012; Rothammer *et al.* 2014; Wiedemar *et al.* 2014; Medugorac *et al.* 2017; Grobler *et al.* 2018; Utsunomiya *et al.* 2019). Each mutation is a complex insertion-deletion of variable size on BTA1 (Fig. 2). None of these genetic mutations (Pc, Pf, Pm and Pg) are directly involved in gene coding, although putative causal effects have been reported by introgression of the Pc allele by gene-editing of bovine embryos that resulted in the birth of healthy and phenotypical unremarkable polled cattle (Carlson *et al.* 2016). The mutations may be involved in gene regulation and translation processes through unconventional mechanisms as speculated by the presence of antisense sequences caused by similar insertions disturbing normal function of horn growth-associated genes (Allais-Bonnet *et al.* 2013; Wiedemar and Drögemüller 2015). However, their association with polledness provides opportunities for genetically selecting animals to produce naturally polled cattle (Prayaga 2007; Spurlock *et al.* 2014; Windig *et al.* 2015). Notably, Pc and Pf are the most frequent mutations observed in most breeds in production systems globally, and hence are the focus of this study. The study has not investigated Pm and Pg as both are associated with breeds not represented in Australian herds (Nellore and Mongolian Yaks, respectively).

SNP-based poll testing has several advantages over the previous microsatellite poll tests. The test is not dependent upon a training or reference database of haplotypes and phenotypes. Diagnostic SNP are also compatible with recently developed SNP chips used for genomic evaluation, and as such could be introduced onto these chips allowing for co-testing of poll and other traits together with genomic evaluations. This strategy has been employed by commercial laboratories globally, and SNP Poll testing was already being investigated as an alternative test by UQ Animal Genetics Laboratory (UQAGL) in early 2017. The acquisition of UQAGL by Neogen in September 2017 resulted in the cessation of the microsatellite test as a standard offering, and replacement with the Neogen SNP poll test (abbreviated as SPTv1 for this report), either as a standalone test or as part of genomics bundles of products including poll, breed-relevant genetic defects and traits, parentage, as well as genomic data, delivered via breed societies for current or future BREEDPLAN evaluations incorporating genomic data. Coincident with the increased industry uptake of genomics, these bundles had a desirable effect of providing poll genotypes for a significantly larger number of animals across various breeds.

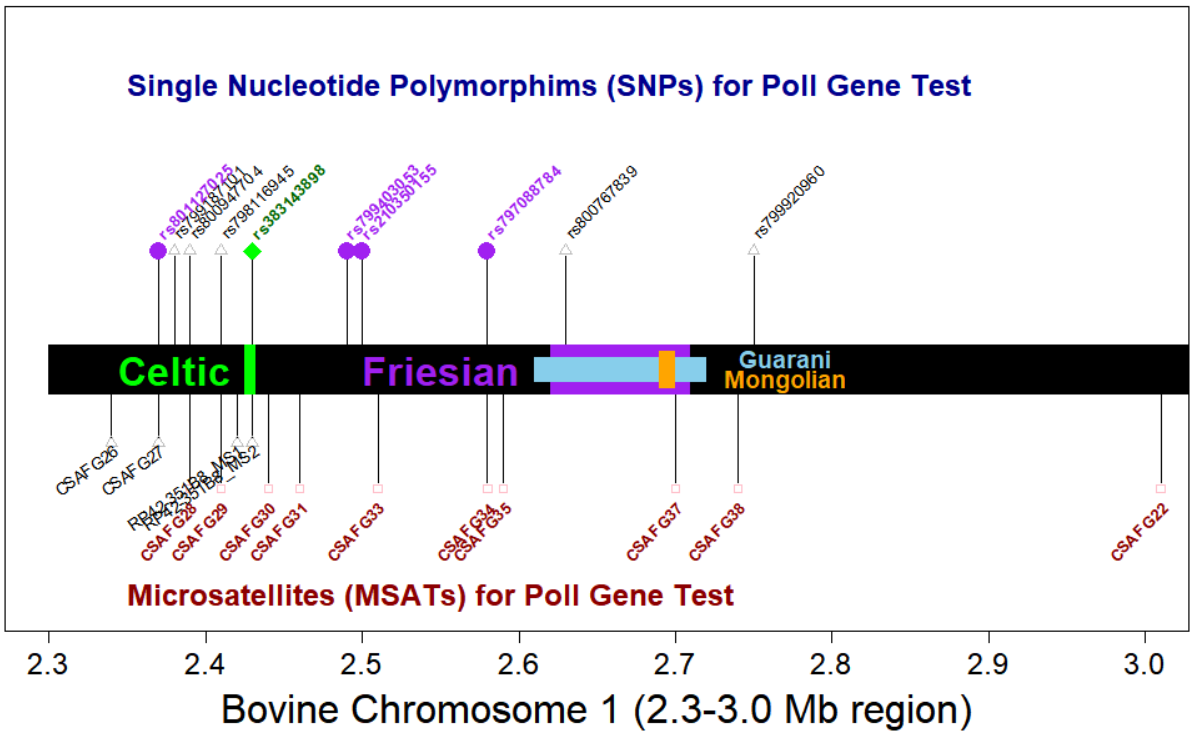


Fig. 2. Map of bovine chromosome 1 POLL region (Bovine assembly: ARS-UCD1.2) showing locations of 4 known insertion-deletions (Celtic, Friesian, Mongolian and Guarani) associated with polledness across various breeds of cattle. Top half shows 10 SNPs and bottom half shows 14 microsatellites, both types are used as genetic markers in various Poll Gene Tests. The Optimized Poll Test (OPT) developed by this project is based on the 5 coloured SNPs (1-green to predict Celtic and 4-purple to predict Friesian mutations). The MSAT-based haplotype poll test used the 10 red-coloured markers (see Appendix 9.2 for genomic positions).

The introduction of the SPTv1 test showed immediate reductions of undetermined results (listed as No Result or Not Determined for SNP and MSAT tests, respectively) across most taurine breeds including Wagyu (Fig. 3). Increases in the frequency of HH (horned animals) is due to the fact poll status is being captured on a larger, more representative proportion of herds, not only those that are phenotypically polled. As such the SPTv1 could be considered a more accurate representation of the frequency of horns and poll in the different breeds. However, with the accumulation of enough data, it became clear that the SPTv1 test did not have the same impact for Brahman or *Indicus*-influenced animals. For Brahman, the rate of undetermined (No Result) increased to 18.2%. For all other breeds the overall percentage undetermined decreased, and in many taurine breeds including Hereford, Charolais, Limousin and Wagyu the problem was completely resolved. However, Brangus (14% No Result), Droughtmaster (2.4% No Result) and Santa Gertrudis (3.8% No Result) all had some undetermined animals, while Shorthorn (7.5% No Result) stood out at the only Taurine breed with ongoing issues. Given the initial bovine sequencing efforts involved a limited number of taurine cattle breeds (Grobler *et al.* 2018), and poll markers identified were based on this taurine data, it was hypothesised that genetic differences in *indicus* breeds and shorthorn were contributing to the problem. As these No Results were unsatisfactory to the breeds and would undermine confidence in the test across the industry, further research was required as it appeared there was some genetic differences in the *indicus* breeds that are prohibiting the test from being fully effective.

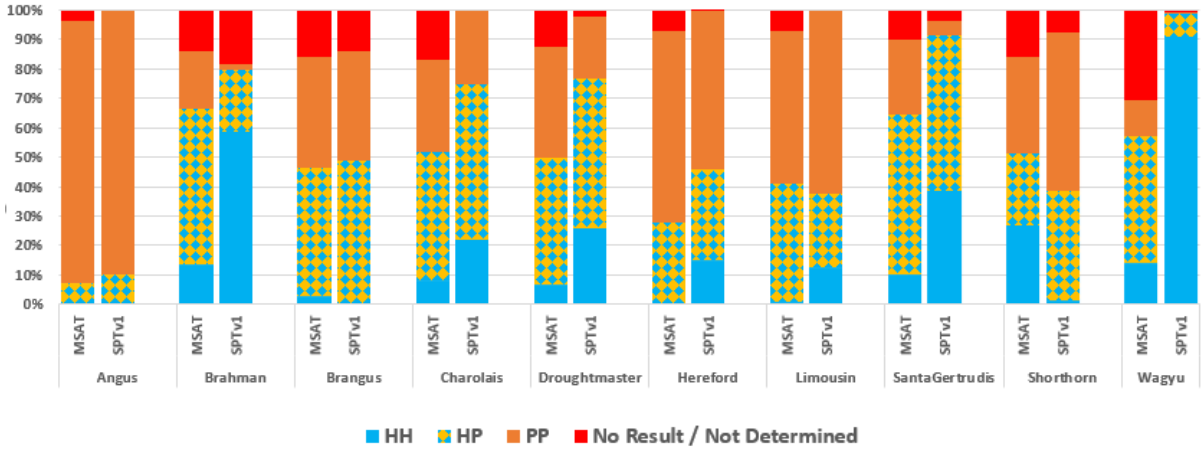


Fig. 3. Percentage of animals per breed genotyped as HH, HP, PP or undetermined (red) using the microsatellite poll test (MSAT) or the SNP Poll test (SPTv1).

A secondary limitation to the success of all previous and current poll tests is the inability to predict scurs formation. The genetic basis of scurs remains unclear although evidence suggests the condition is genetically complex and affected by polled status as well as sex of individuals (Capitan *et al.* 2009; Mariasegaram *et al.* 2010; Capitan *et al.* 2011; Tetens *et al.* 2015). Age of scurs development is also variable, and, unlike horns, can occur later in life (Capitan *et al.* 2009). In addition, an as yet unidentified African horn gene has been speculated as a possible explanation for the epistasis-like complexity in the horn inheritance in several breeds (White and Ibsen 1936; Long and Gregory 1978; Prayaga 2007). However, no empirical evidence has been presented to confirm its existence to date (Grobler *et al.* 2018). While one can argue that scurs is less of an economic and welfare issue within herds, the presentation of scurs is often indistinguishable from horns at an early age and industry practice would be to remove (dehorn or disbud) scurs at earliest convenience. Further research to understand the frequency and heritability of scurs in Australian herds will hopefully better predict scurs within herds will benefit the industry.

Poll is in some farmers' minds linked to reproductive deficiencies. Previous research has shown no relationship between Polledness and reproductive disorders such as poor sheath structure, premature spiral deviation of the penis (PSDP) or reduced libido, this same study found that PSDP was more prevalent in polled cattle (MLA project AWW.0187 (2008)). The authors concluded that an increased prevalence of the polled condition would have no or minimal negative influence on bull fertility. Despite this, there remains a resistance by some in the industry to use polled bulls due to a belief of inferior reproductive or production genetics being introduced to their herds. A comprehensive analysis of performance using existing phenotypic and genomic data will hopefully help address these ongoing concerns.

The over-arching objective of this project was to develop an improved SNP-based poll test suitable for inclusion on future SNP chips and which would provide a cost-effective and reliable method for accurately assessing the poll status of all economically important cattle breeds in Australia. At the same time, the project investigated solutions to other potential problems (scurs, perceived negative effects on fertility and beef traits) to remove potential barriers to adoption of poll testing in the Australian beef sector.

2 Project objectives

The main research objectives are;

1) A more cost-effective, SNP-based test based on haplotypes, which can be incorporated into multi-purpose SNP panels,
2) Elimination or accurate prediction of carriers of the scur gene, and
3) Determining whether there are negative effects of the poll gene on reproductive phenotypes.

This work will allow;

a) Greater uptake of breeding for poll as both testing costs and uncertainty relating to the phenotypes of heterozygous animals be reduced,
b) Potentially, the identification of bulls that, when used over horned cows, will produce "clean" progeny that do not require dehorning; and
c) Certainty in advocating the use of polled bulls where concern currently exists about negative pleiotropic effects on fertility and growth. These traits, and the relationship with polled trait, will have been examined and understood. This will provide breeders with accurate and quantified data to assess the benefits or risks of selection of polled bulls in their breeding program.

In addressing these industry-wide concerns that are recognised handbrakes to the wider application of poll testing in Australian herds, particularly Northern Australia, it is anticipated that these outputs will result in a reduction in the proportion of calves that need dehorning. There are clear benefits from both the social (handler and animal welfare) and economic (reduced calf mortality, reduced labour costs) in adoption of naturally polled herds. This project intends to remove impediments to the widespread uptake of the technology to achieve that goal.

3 Methodology

3.1 Genomic and phenotypic data to improve poll testing (Objective 1)

3.1.1 Ethics statement

Ethics approval for the pre-tested animals was not required, as these results were generated under commercial services using microsatellites and SNP based poll testing. However, collection of tail hair and blood samples, head phenotyping and generation of genotyping and targeted and whole-genome sequencing for this project were approved (Animal ethics approval numbers SVS/301/18, SVS/465/18 and SVS/ANRFA/397/19).

3.1.2 Genotypic and phenotypic records

A total of 39,943 animals from several of Australian beef cattle were used to investigate the previously available genomic records and test the new SNP-based poll testing assay (Table 1). The available records consisted of 20,534 animals tested with microsatellites (MSAT) based poll testing assays. Genotypic records of SNPs within the poll-locus on BTA1 and previously predicted results based on SNP-based poll testing assays were available for 20,636 animals (Appendix 9.1). Of these, 1533 animals were tested by both MSAT and SPTv1. The previously genotyped animals include 1,999 animals which were initially recorded as "No Results" with SPTv1 testing assay.

The available genomic records were without any phenotypic data. However, it was important to investigate the efficiency of new poll testing for phenotypic concordance. Therefore, 20,333 genotyped animals having identification of various breeds with phenotypic records were acquired

through the online BREEDPLAN database search (http://breedplan.une.edu.au/index.php). These animals were defined for their head-status as Horn, Scur or Poll.

3.1.3 Validation population

The Droughtmaster herd at the University of Queensland (UQ) Gatton campus was established as a validation resource population for this project to collect phenotypes and genetic data. The herd consists of 150 animals of all ages, of which most mature breeding cows and bulls were individually mustered by the project team and farm personnel to accurately record the head-status of each of these animals. Each animal was photographed for head-phenotype records, and tail-hair and blood samples were collected for poll-testing and future usage in the project to investigate scur genetics.

Another UQ (Pinjarra Hills) herd consisting of Brahman, Brangus and crossbred cattle was also included in the validation population. In addition, collection of tail hair samples and phenotypic records from various private cattle properties (Brangus, Droughtmaster, Hereford and Santa Gertrudis) were used to check the efficiency of the newly developed assay and investigate scurs genetics and phenotypic correlations. In total, 393 animals form the validation population. All were carefully phenotyped for head-status and tested with the newly developed OPT (Table 1).

Table 1. Number of samples of various breeds tested by Microsatellites and SNPs based assays.

Breed	Tested samples	Tested by MSATs	Tested by SNPs	Validation samples
Angus	1630	28	1602	
Brahman	7160	4532	2819	82
Brangus	806	745	72	18
Charolais	3148	2666	900	
Droughtmaster	2718	2223	558	162
Hereford	6424	3485	3341	66
Limousin	2193	2124	207	
Santa Gertrudis	4456	4306	136	29
Shorthorn	316	224	67	
Wagyu	9182	201	9050	
Other breeds [A]	1910	-	1884	36
Total	39,943	20,534	20,636	393

[A] List of other breeds are provided in Appendix 9.1.

3.1.4 Genetic markers (microsatellites and single nucleotide polymorphism)

A total of 14 microsatellite (MSAT) markers located between positions 2,341,080 to 3,014,463 on BTA1 have strong associations with polledness across different populations (Fig. 2, Appendix 9.2). The MSAT based poll testing used a set of 10-MSAT markers (Mariasegaram *et al.* 2012; Piper *et al.* 2014), which have been provided in the final report of the MLA project B.AWW.0222 (2014).

A set of 10 SNPs was identified through literature, as these SNP markers have shown strong linkage disequilibrium with the known genetic variants strongly associated with poll status (Fig. 2, Table 2). It is noteworthy to point out that, to date, a total of 4 genetic variants (technically called InDels = Insertion-Deletions) within the poll-locus on BTA1 have been associated with naturally polled status. These InDels have been named; Celtic (Medugorac *et al.* 2012), Friesian (Rothammer *et al.* 2014), Mongolian (Medugorac *et al.* 2017) and Guarani (Utsunomiya *et al.* 2019), representing the geographic origin of the discovery breeds. However, the Australian beef cattle breeds have their ancestry carrying predominantly the Celtic and Friesian InDels. Hence, the selected set of 10 SNPs

were located surrounding these two InDels and their alleles presumably causing poll are denoted as Pc and Pf for Celtic and Friesian, respectively. Out of these ten SNPs, the SNP-based poll testing (SPTv1) assay used 8 SNPs in the predictions.

Table 2. List of single nucleotide polymorphisms (SNP) on autosome 1 known for strong linkage disequilibrium (LD) with Celtic (Pc) and Friesian (Pf) mutations and their call rate in cattle breeds.

SNPs	ARS-UCD1.2 positions	Mutations	Poll alleles	LD with	Call rate (%) [A]
rs801127025	2,372,456	P_{5ID}	T	P_F	99.47
rs799187101	2,377,687	G>A	A	P_F	99.98
rs800947704	2,378,745	C>T	T	P_F	96.32
rs798116945	2,407,338	G>C	C	P_F	99.99
rs383143898	2,429,319	P_{202ID}	T	P_C	99.92
rs799403053	2,486,811	T>C	C	P_F	99.80
rs210350155	2,491,161	C>A	A	P_F	99.23
rs797088784	2,578,598	G>A	A	P_F	99.44
rs800767839	2,629,115	T>A	A	P_F	99.94
rs799920960	2,748,715	C>G	G	P_F	100

[A] indicates mean call rate (call rate = percentage success of returning an informative genotype for the specific marker). Further analysis was undertaken to assess Breed-wise call rates (Appendix 9.3).

3.1.5 Sequencing samples, primers and polymerase chain reaction (PCR)

Having notes an unusually high fail rate for one SNP marker (rs800947704) and given the fail rate was highest in Brahman (14.4%, Table 6), a set of 60 Brahman samples were selected from the available DNA resource bank to sequence the genomic region around SNP rs800947704. The selected SNP rs800947704 not only had highest failing rate, but also failed the Hardy-Weinberg equilibrium (HWE) test and was one of the major causes of haplotype distortions in the predictions of SNP-based poll testing. Hence, this marker warranted a detailed investigation by sequencing of the target region. Primers used are below.

- Forward primer: 5'-TCCCTCTGCTGTGATAAACACC-3' (primer length: 22 bp)
- Reverse primer: 5'-GTTTGGCCTTGGTTTGTGGT-3' (primer length: 20 bp)
- Internal primer: 5'-TCCAATGAACACCCAGGACT-3' (primer length: 20 bp)

DNA extraction from all samples used for targeted DNA sequencing and SNP genotyping was performed by using the standardized protocol at the genetic testing laboratory of Neogen Australasia. DNA fragments of 1,098 bp (2,377,810-2,378,907) harbouring a targeted genotype rs800947704 (g.2378745G>C) were amplified by using the Forward and Reverse primers. PCR reactions containing 15-20 ng DNA, 10 µM forward and reverse primers, 0.12 µl taq in a total volume of 25 µl and employed standard PCR cycling conditions on an Applied Biosystems thermocycler. An internal primer was then used to generate 665 bp (2,378,243-2,378,907) sequencing reads with rs800947704 flanked by approximately 300 bp upstream and downstream of the marker. Unused dNTPs and primers were removed using ExoSAP-IT® (USB Corporation distributed by GE Healthcare Bio-Sciences, Rydalmere, Australia). Sequencing was performed using an ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit Version 3.1 (PE Applied Biosystems, Foster City, USA) following the instructions supplied with the kit. Sequencing separation was performed on an ABI 3730xl automated sequencer. Forward and reverse sequences were aligned and edited using ChromasPro (Technelysium Pty Ltd, Tewantin, Australia).

3.1.6 Prediction of Celtic (Pc) and Friesian (Pf) types of poll associated alleles

Prediction of polledness-associated Pc and Pf mutations were generated using the SNP markers (Fig. 2, Table 2) available on commercial bovine BeadChip assays (Illumina) including Neogen's proprietary GGP-LDv4, GGP Taurus 50K or GGP Indicus 35K assays (Neogen Corporation, Lincoln, NE). The Celtic (Pc) allele is predicted by translating a single SNP marker rs383143898 based on its horn or poll allele (Table 2). The Friesian (Pf) allele is predicted based upon haplotype associated with multiple markers in LD with Pf (Table 2). Note that the combination of Pf-markers used varied with different versions of the Poll Test being assessed in this study. Results represent reconciled outcomes from both predictions to generate allele-pairs (genotypes) such as HH, HPc, HPf PcPc, PcPf or PfPf. However, if the Pc-associated SNP or more than two Pf-associated SNPs fail during genotyping, or two or more SNPs differ in predicted genotype (H versus Pf) then the result is considered ambiguous and termed as a "No Result". The optimized poll testing assay remained identical for Pc, while genotyping failure or contrasting prediction were restricted to only one differing or missing Pf SNP.

3.2 Whole-genome genotyping and sequencing for scurs gene (Objective 2)

3.2.1 High-density SNP genotypes (770K) for Scur analyses

A total of 200 animals were selected to investigate scurs gene in three breeds (Brahman, Droughtmaster and Hereford). Of these, 197 animals passed the DNA quality and genotyping assay requirements to generate high-density 770K data (Table 3). The obtained genotypes were investigated to find genomic regions underpinning scurs development by using two types of approaches, genome-wise association (GWAS) and composite selection signals (CSS). Each approach was applied on two subsets of data based on contrasting categories for head-status. The control dataset consisted of pure horned (HH genotypes) and pure Polled (PcPc, PcPf or PfPf) animals, so that the known genomic region underlying polledness can be detected as a validation process for each approach. The discovery dataset consisted of animals which were genetically similar (heterozygous, HPc or HPf) at the poll locus while phenotypically different for head-status, i.e., scurred and polled cohorts. It was expected that any genomic regions underlying scurs development would be detected, excluding the poll locus on BTA1 which would be neutralized by the shared genetic patterns in both cohorts. These approaches were applied to the control and discovery datasets within each of the three breeds. Given that the selected approaches can be affected by lower sample size, all breeds data were also analysed as a combined set.

Table 3. Breed-wise genotyped animals for genome-wide analyses for scurs.

Breeds	Control (Horn vs Poll)		Discovery (Poll vs Scur)		Total
	Horned (HH)	Polled (PP) [A]	Polled (HP) [B]	Scurred (HP) [B]	
Brahman	8	9	25	24	66
Droughtmaster	8	8	25	25	66
Hereford	8	8	24	25	65
Combined	24	25	74	74	197

[A] PP animals carry any two copies of Polled alleles with genotypes: PcPc, PcPf or PfPf.

[B] HP animals carry only one of the two polled alleles with genotypes: HPc or HPf.

3.2.2 GWAS and CSS analyses of genome-wide genotypes for scurs gene discovery

Genome wide association (GWAS) analyses

The genome-wise association (GWAS) analyses were performed using the *qtscore* function (with parameter: trait = "binomial") in R-package: GenABEL (Aulchenko *et al.* 2007). Genome-wide *p*-values were converted to –log10(*p*-values), which were used to generate Manhattan plots in R program. A significance threshold of $p < 1.0\text{-}6$, *i.e.,* –log10(*p)* = 6 was used to detect putative SNPs underlying the target phenotypes in the proposed analyses.

Composite selection signal (CSS) analyses

Composite Selection Signal (CSS) statistics is a recently established approach to find genomic regions underlying a specific trait (assumed experiencing selection pressure) by comparing two populations (cohorts) with contrasting phenotypes of that traits (Randhawa *et al.* 2014). Hence, CSS analyses considered two contrasting cohorts (i.e. cohorts *i* and *j*) based on head-status for the control (i.e., horned and polled) or discovery (i.e., scurred and polled) dataset in a single pairwise comparison. First, the constituent selection tests, F_{ST} (Weir and Cockerham 1984), XP-EHH (Sabeti *et al.* 2007) and ΔDAF (Grossman *et al.* 2010), were computed in each analysis by following their standard procedures. Then, the CSS statistics were computed at each locus by combining the results from three constituent selection tests such that, for each constituent method implemented to compare cohorts *i* and *j*, test statistics were ranked (1, ..., *n*) genome-wide on *n* SNPs. Ranks were converted to fractional ranks (*r*′) (between 0 and 1) by 1/(*n* + 1) through *n*/(*n* + 1). Fractional ranks were converted to $Z = \Phi^{-1}(r')$ where $\Phi^{-1}(\cdot)$ is the inverse normal cumulative distribution function (CDF). Mean *Z* scores were calculated by averaging *Z*-values across all constituent tests at each SNP *x* position. The pairwise CSS statistics were equal to $-\log_{10}$ of *p*-values of mean *Z* distribution directly obtained from the distribution of means from a normal distribution. The CSS scores were smoothed by averaging raw-CSS of SNPs within 50kb overlapping sliding windows along the length of each chromosome. The smoothed CSS scores were used to capture the putative genomic regions using the top 0.1% threshold.

3.2.3 Discovery of candidate genes within GWAS and CSS regions

The putative regions detected by the GWAS and CSS analyses were used to discover the candidate genes to be associated with scurs development and polledness. Note that high-density genotypes are expected to harbour a cluster of multiple SNPs above the significance threshold at any region of interest. Hence, regions with a single SNP above the threshold were considered spurious. The boundaries of remaining multi-SNPs putative regions were defined by the first and last significant SNPs, which neighboured as a cluster within 1 Mb span. These regions were searched for candidate genes on the bovine genome assembly ARS-UCD1.2 (Ensembl Genes 99). Genomic linkage and hitch-hiking effects can cause a few significant SNPs to be located in the neighbourhood of a candidate gene. Hence the putative regions were expanded 1 Mb (0.5 Mb in both directions).

3.2.4 Whole genome sequencing for scur analyses

Genome mapping approaches, including GWAS and CSS, potentially highlight a chromosomal region to associate with a candidate phenotype. Although using high-density (e.g., 770K) SNPs genotyping assays provide opportunity to localize a region of interest, they are less likely to provide insights about the underlying genetic elements (mutations, deletions-insertions etc.). Whole genome sequencing (WGS) on the other hand can detect any structural variation and has become accessible with reduced costs and high-throughput efficiency of the next-generation sequencing technologies.

The genetic variants associated with horn development, especially polledness, have been found as large structural variations, making it difficult to find those types of large mutations (over 80kb size)

with common short-read sequencing technologies. Long-reading sequencing technologies such as portable MinION device (Oxford Nanopore Technologies, https://nanoporetech.com) have enabled greater success in the detection of large structural variants in the genome, as well as providing rapid assembly and analysis opportunities (Lu *et al.* 2016).

This project has sequenced seven animals (Brahman = 5, Droughtmaster = 2) with different phenotypes and genotypes to capture any novel structural variants in the previously identified candidate genomic regions underlying scurs development that can potentially expand the research towards development of informative genetic markers for prediction of scurs in cattle. To increase the depth of genome coverage, five of the selected animals were sequenced twice, hence a total of 12 flow-cells were used on the MinION device.

Samples chosen for sequencing were grouped by phenotypic and genotypic categories to compare against each other, as well as for a comparison with WGS data from contemporary research groups working on bovine genomics, including MLA project P.PSH.0868 (2019), where horn-poll phenotypes and genotyped were known. Animals sequenced fell into the following categories:

- Horned with homozygous genotypes (HH)
- Clean polled with homozygous genotypes (PcPc)
- Clean polled with heterozygous genotypes (HPc)
- Scurred with heterozygous genotypes (HPc)
- Horn-like scurred (Fig. 4A) with heterozygous genotypes (HPc)
- Clean polled with homozygous genotypes (PcPc)
- Scab-like scurred (Fig. 4B-C) with homozygous genotypes (PcPc)

The last 2 animals were half-sibs, sharing the same sire and genotype, yet displaying distinct phenotypic differences at two years of age. As scurs in a homozygous polled (PcPc) animal is unexpected but identified as possible (Figure 4 B-C) it is hoped that this will give additional insight into the mechanism of Scurs development.

High-quality and yield DNA of these animals was extracted from approximately 5ml of whole blood using the Gentra Puregene Blood Kit. Each sample was QC evaluated using Qubit and Nanodrop spectrophotometers, and by gel electrophoresis. Sequencing libraries were prepared and loaded onto the Flow-cells by following the proprietary procedures and kits from Oxford Nanopore Technologies and New England Biolab. Sequencing runs were completed in 72-96 hours. The sequence data were then analysed by comparing them against reference genomes and each other to identify genetic variants for each category of sample phenotypes and genotypes (poll-locus).

A **B** **C**

Fig. 4. Photographs of some uniquely identified animals selected for whole-genome sequencing.
A: Brahman cow with HPc genotypes and poll-shaped head with fixed scurs.
B-C: Droughtmaster bulls with PcPc genotypes and poll-shaped head with scab like scurs.

3.3 Phenotypic data for genetic effects of polledness (Objective 3)

For comparison of genetic merit between animals, evaluation of genetic effects of a trait is more practical by substituting estimated breeding value (EBV) for phenotypic values. EBV is a tool of genetic evaluation between animals for a particular trait by accounting for heritability and fixed effects (see Supplementary file for further explanation of EBVs). EBVs for a quantitative trait capture the aggregate additive genetic value by using phenotype of an animal together with phenotypes of its relatives (Henderson 1975). EBVs denote how an animal's genetics is different than the genetic base or breed averages. Accuracy of EBV predictions increases as more information becomes available for animal's direct performance, pedigree and progeny. BREEDPLAN (http://breedplan.une.edu.au/index.php) is an advanced genetic evaluation system implemented for national beef recording scheme in Australia to compute EBVs, which can be used to highlight the genetic differences in various production, reproduction, carcase, behaviour and feed efficiency traits between various head-status cohorts. The aim of this investigation was to compare the genetic merit of naturally polled and horned animals for a range of economically important traits within the major breeds of Australian beef cattle.

3.3.1 BREEDPLAN phenotype data

The BREEDPLAN database provides access to registered animals of different breeds, including the head-status (generally Poll, Horn or Scur) and EBVs of several phenotypic traits in categories such as growth (weight), fertility, carcase, behaviour and feed efficiency (Table 4). Following preliminary assessment and QC of 2,132,525 eligible animals across ten Australian breeds (Table 5), and removal of incomplete or unsuitable data (e.g. young animals with low accuracies), phenotypic data was obtained from the BREEDPLAN database for 1,825,981 animals that met the required criteria of having both EBVs and a published head status. As recording of traits is breed dependent, and trait-specific EBVs depends on how commonly the trait is recorded in a breed, higher numbers of EBVs were available for commonly recorded traits affecting growth and fertility. Efforts were made to ensure a mix of animals from across different herds.

BREEDPLAN EBVs are classified for interpreting accuracy, such that less than 50% = preliminary, 50-74% = medium, 75-90% = medium-high, and above 90% = high accuracy estimates of the animal's true breeding value. Therefore, EBV records were obtained for all traits along with their EBV accuracy records. Note that accuracy of an EBV increases as more information such as repeated measures and relatives (pedigree, siblings and progeny), and is included into the prediction models, meaning that more animals have their EBVs within low-medium range than medium-high and high categories.

The phenotypic records for head-status obtained to investigate genotype-and-phenotype concordance of OPT-based genomic horn and poll predictions (Appendix 9.4) contained 16,733 animals of ten common breeds. However, EBVs for a subset of 5,909 animals of only two breeds (Brahman: n = 3240, and Hereford: n = 2669) were investigated for polledness effects on other traits, as other breeds were limited by fewer number of animals (Table 5). Brahman and Hereford samples with phenotype-genotype discrepancy (n = 1,110) were excluded as possible phenotyping and data recording errors (section 4.1.7).

Table 4. Phenotypic traits available through BREEDPLAN database for selected breeds. Details on how EBVs are computed and what are the preferred values are provided in Supplementary file.

Head-status	Weight (Kg)	Fertility/Calving	Carcase	Other
• Horned • Scurred • Polled	• Birth weight • 200 days growth • 400 days weight • 600 days weight • Mature cow weight • Milk yield	• Scrotal size • Days to calving Gestation length • Calving ease direct • Calving ease daughter	• Carcase weight • Eye muscle area • Fat (Rib and Rump) depth • Retail beef yield • Intramuscular fat • Shear force	• Net feed intake • Docility • Flight time

Table 5. BREEDPLAN phenotypes and head status data on ten breeds. Full list of OPT tested breeds is given in Appendix 9.4.

Breeds	BREEDPLAN sourced animals		OPT tested animals	
	Head-status	EBVs	Head-status	EBVs
Brahman	535,005	295,160	3,617	3,240
Brangus	86,227	47,223	23	-
Charolais	43,926	39,853	518	-
Droughtmaster	264,967	112,224	608	-
Hereford	835,402	668,386	3,402	2,669
Limousin	129,760	133,290	24	-
Santa Gertrudis	358,470	150,447	114	-
Shorthorn	173,777	140,934	67	-
Simmental	125,478	122,907	2	-
Wagyu	114,518	115,557	8,358	-
Total	**2,132,525**	**1,825,981**	**16,733**	**5,909**

3.3.2 Analyses of EBVs to determine genetic effects of polledness on different traits

Currently, EBVs can only be compared within each breed. Moreover, each breed has registered animals from a range of years since early 1950s. Therefore, a few considerations were applied to the acquired BREEDPLAN data to minimize confounding factors and biases such as;

- All analyses were performed within each of the eight breeds. Given that Brangus and Wagyu have very limited number of horned and polled animals respectively (Fig. 1), they were excluded from EBVs analyses.
- Animals born in birth-years from 2000 to 2018 were used (EBVs data were obtained during January to March 2020).
- All analyses were performed as pair-wise comparisons between two cohorts, polled and horned. Phenotypically scurred animals were excluded as head-status misclassification between scurred and horned animals (12.2%; Table 8) was deemed to be a major bias in the comparisons.

- All analyses were performed considering the levels of EBV accuracies as well as the number of animals with EBVs above a particular accuracy threshold.

The filtered and structured dataset analyses were performed by using the R program (R Core Team 2019) to compute the following statistics for each trait of a breed and were probed for the poll-*vs*-horn cohorts.;

- Summary statistics of Mean and Standard Deviation (SD) were computed for the two cohorts (horned and polled).
- Descriptive statistics for pairwise comparisons between the means were performed by the t-tests with pooled SD and p-values were obtained by using the *t.test* function in R-package: `stats`.
- Effect sizes (ES) on each trait due to polledness were computed using the Cohen's *d* (Cohen 1977*)* using this formula;

$$d = \frac{\text{Mean of polled} - \text{Mean of horned}}{Pooled\ SD}$$

where;

$$Pooled\ SD = \sqrt{\frac{(\text{SD of polled})^2 + (\text{SD of horned})^2}{2}}$$

- Point Biserial Correlations (PBC) – a correlation measure *r* of the strength of association between continuous-level variable (i.e., trait-wise EBVs) and a binary variable of head-status (poll, horn) – were computed with *biserial.cor* function in R-package: ltm (Rizopoulos 2006).

Results of these analyses were interpreted such as;

- P-values were deemed significant above the adjusted α for multiple-comparison by using the Bonferroni correction at $\alpha = 0.05/N$, and *N* is total number of recorded traits with EBVs (i.e., the number of comparisons) in each breed.
- ES (*d*) and PBC (*r*) values provide direction for interpretation that which of the two cohorts are better in a comparison to communicate the practical significance of results. Our analyses were performed such that positive values indicate polled-as-favourable while negative values indicate horned-as-favourable, except four traits are interpreted as opposite (because a lower or negative value of their EBVs is considered preferable) which are; Days to Calving, Gestation Length, Shear Force and Net Feed Intake Finishing (NFIF). Moreover, two other traits (Rib Fat, Rump Fat) can be considered in either direction depending on the breeding objectives (Supplementary file about why traits are preferred for high or low EBVs).

Results of these analyses are provided in several graphs and tables. Considering the extensive information available, a summary comparison of all traits is provided by averaging the ES (*d*) and PBC (*r*) of all traits within each breed, after adjusting each traits' importance in the cattle industry, to compare the overall genetic merit of naturally polled and horned animals.

4 Results

With the help of previously available genomic data, phenotypic records from BREEDPLAN and the validation population, the first milestone of the project has been successfully achieved by developing the new SNP-based optimized poll testing (OPT), which have already been implemented for commercial testing. Briefly, the steps during the development of OPT are described.

4.1 Development of the SNP-based optimized poll testing (Objective 1)

Ten SNPs known for strong LD with associated insertion-deletions (Pc and Pf) were investigated to for the cause of their failure and the degree of informativeness for predicting of the putative genotypes for head-status in Australian cattle.

4.1.1 Excluding the monomorphic SNPs

Overall, two SNPs (rs798116945 and rs800767839) were homozygous across most of the European and Zebu breeds and their cross-bred populations. Hence, these two SNPs were declared non-informative and superfluous for the Poll gene testing assays across all the genotyped samples (Table 2, Fig. 2). The lack of true predictiveness in animals led us to exclude those two SNPs from further investigations.

4.1.2 Disruptive and missing SNPs in the current poll testing assays

During the investigation of previously available SNP genotypes on 16,863 animals, it was noted that one genetic marker, denoted as SNP rs800947704, was consistently failing across various genotyping platforms in several breeds (Table 6). Overall, this SNP failed 4% of tests, but was most problematic in *Bos indicus* (Brahman) and crossbred (Droughtmaster) cattle where fail rates were more than 10%.

Table 6. Frequency of missing target SNP rs800947704 in different Australian breeds.

Breed	Number genotyped	Missing SNP (n)	Missing SNP (%)
Brahman	2,603	375	14.4 %
Wagyu	9,047	186	2.10 %
Angus	1,594	49	3.10 %
Droughtmaster	289	31	10.7 %
Hereford	2,820	10	0.40 %
Santa Gertrudis	81	5	6.20 %
Composite	35	2	5.70 %
Limousin	24	2	8.30 %
Total	**16,863**	**664**	**4.00 %**

4.1.3 Excluding the disruptive SNPs

This project also identified two SNPs, rs799187101 and rs800947704, as the major cause of "No Result" in the Zebu cattle when running the SPTv1 assay. Neither SNP is in complete LD in the haplotypes predicting the presence of Pc or Pf.

The highly disruptive SNP (rs800947704) was investigated for the high rates of fails for genotyping calls and HWE. High quality targeted genome sequences containing SNP region were completed for 55 samples out of 60 selected Brahman samples. Alignment of 55 sequenced fragments (665 bp) showed that the region was enriched with several neutral SNPs within a few bps of the rs800947704 (Figure 5). Genomic sequences generated in this project have been deposited in NCBI's GenBank (accession numbers: MN473394 to MN473448). Genotype call failure at the target SNP (C>T, rs800947704) in the genotyping assays can be explained by a neighbouring SNP rs381418143 (g.2378742A>G) located 3 bp upstream to rs800947704 in Brahman cattle. This SNP, uncommon in taurine breeds, caused probe hybridisation issues. The current probes are designed based upon the Taurine reference genome to recognise allele A only of rs381418143. Therefore, all samples carrying the allele A at rs381418143 (n=22) were correctly genotyped at the target SNP rs800947704 for all alleles (CC, TT or CT=Y). However, DNA samples (n=16) with G at rs381418143 resulted in incorrect genotype or failure to generate a genotype depending upon whether animals tested were heterozygous GA or homozygous AA for rs800947704. SNPs rs799187101 and rs799920960 were also deemed unreliable and excluded.

↓Samples	DNA Sequence results around rs381418143 (-3) & rs800947704 (0)	Genotypic assay
SNPs →	*******-3**0******	rs800947704
1	TGTGGCC**G**TTCGGGGTG	--
2	TGTGGCC**R**TTCGGGGTG	CC
3	TGTGGCC**G**TTCGGGGTG	--
4	TGTGGCC**G**TTCGGGGTG	--
5	TGTGGCC**A**TTYGGGGTG	CT
6	TGTGGCC**R**TTYGGGGTG	TT
7	TGTGGCC**R**TTCGGGGTG	CC
8	TGTGGCC**G**TTCGGGGTG	--
9	TGTGGCC**A**TTTGGGGTG	TT
10	TGTGGCC**R**TTYGGGGTG	TT

Fig. 5. Results of targeted DNA sequences of Brahman samples. Ten of the 55 Brahman samples are presented, highlighting the variation observed surrounding SNP rs800947704, marked as position 0 [heterozygous C/T=Y]. Upstream variant SNP rs381418143 [heterozygous A/G=R] was identified 3 base pairs upstream of rs800947704 (marked as -3) and affecting accurate probe binding. Two types of errors were observed in the genotyping assays: 1: genotype failed (--) when -3=G (yellow highlights), 2: wrong genotypes when -3=R and 0=Y (purple highlights).

4.1.4 Selection of informative SNPs for OPT

The remaining 5 SNPs (rs801127025, rs383143898, rs799403053, rs210350155 and rs797088784) were retained to assess accurate predictions of Pc and Pf alleles, with rs383143898 (P202ID) as the sole predictor for Pc, while the other 4 SNPs are associated with Pf (Fig.2, Table 2). Collectively, these 5 SNP markers constitute the optimized poll testing (OPT) assay, which predicts 5 possible genotypes. The OPT assay predicts the genetic make of an animal for the number of copies of a horn (H) and poll (Pc, Pf) allele.

The results are provided as;

- **HH** Two copies of horn allele
- **HPc** One copy of horn + one copy of Celtic poll
- **HPf** One copy of horn + one copy of Friesian poll
- **PcPf** One copy of Celtic poll + one copy of Friesian poll
- **PcPc** Two copies of Celtic allele
- **PfPf** Two copies of Friesian allele

No Result (**NR**) will most likely be due to a failure to amplify one or more markers during the genotyping process, and hence resolved with resampling. A novel haplotype could also be possible, although no evidence of such haplotypes was observed in this study.

4.1.5 Validation of OPT predictions with known genotypes

OPT was validated by re-evaluating 18,637 samples previously tested successfully by SPTv1 (Table 7). Importantly, results confirmed that the predictions remain unchanged (100%) using OPT relative to the original SPTv1 prediction (Table 7), ensuring no revision of prior test results was needed. Of the samples previously unable to be predicted based upon the SPTv1 translations (No Result, n=1,999), 1,990 (99.6%) were effectively classified into one of the 5 genotypic predictions (Table 7, Appendix 9.1). Thus, out of the total genotyped samples (n=20,636) the success rate for OPT predictions was 99.96% as compared to 90.31% for the SPTv1 assays.

Subsequent investigation of the remaining unpredicted (No Result) 9 samples found that 8 of them were due to genotyping failure for rs383143898 (Pc) and one failed for multiple markers associated with Pf. Both genotyping error rates were within the expected < 0.01% range (Wu et al. 2019) and can be resolved by resampling. Hence, the newly developed OPT markers can effectively predict poll status in multiple breeds of cattle including European, Zebu and their composite animals (Fig. 6, Appendix 9.1).

4.1.6 Resolving the "No results" issues with OPT

As previously noted, both MSAT and SNP-based SPTv1 poll testing assays, were consistently failing to provide a prediction in a subset of samples, denoted as "Not Determined" or "No Result" (Fig. 3). While re-testing 20,636 genotyped samples with OPT, it was possible to resolve 1990 out of 1999 "No result" samples (Table 7). This is demonstrated in Fig. 6, where the Not Determined/No Results (highlighted in red) are resolved or dramatically reduced across all breeds. Only nine samples remained as "No result" because of marker amplification failure. Hence using OPT, the issue of "No Results" is completely resolved, except where samples fail, likely due to sample quality issues.

Table 7: Comparison of SPTv1 and OPT on 20,636 Australian breed samples. List of breed-wise No results is in Appendix 9.1.

Poll results	Comparison of known genotypes			Comparison including No results		
	SPTv1	OPT	Change	SPTv1	OPT	Change
HH	11,113	11,113	0	11,113	12,785	1,672
HPc	3,420	3,420	0	3,420	3,696	276
HPf	135	135	0	135	147	12
PcPc	3,641	3,641	0	3,641	3,651	10
PcPf	300	300	0	300	311	11
PfPf	28	28	0	28	37	9
No Results	-	-	-	1,999	9	-1,990
Total	**18,637**	**18,637**	**0**	**20,636**	**20,636**	**0**

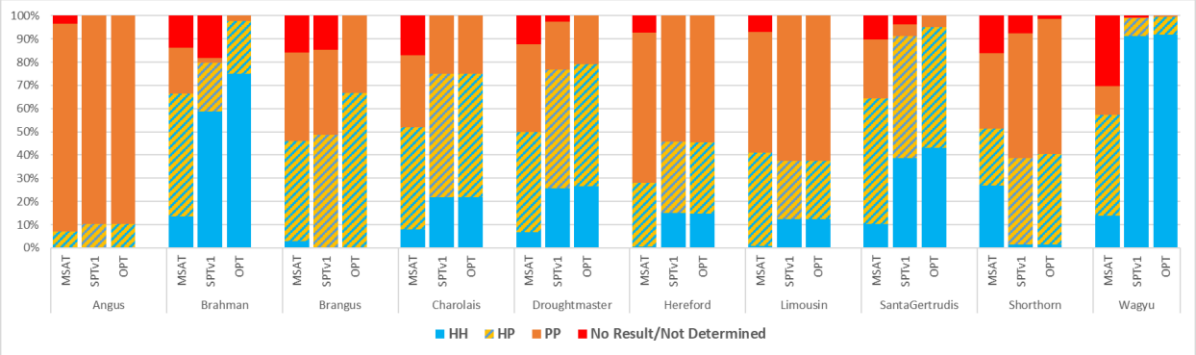


Fig. 6. Comparison of poll testing by microsatellites (MSAT) and SNP poll test – v1 (SPTv1) and optimized poll test (OPT) in various Australian breeds.

Based on this OPT data from 20,636 animals it was also possible to get an accurate estimate of frequencies for the poll-associated markers within a large subset of the Australian herd (Fig. 7), including at a breed level. The Horn allele (H) remains dominant at 71% of all alleles identified, and is most common in Brahman, Wagyu and Santa Gertrudis herds. The Celtic poll variant (Pc) was identified in 28% of test cases, while the Friesian mutation (Pf) is uncommon (~1%). Pc is the dominant poll allele in all breeds except for Shorthorn where Pf was very common and representing 48% of all identified alleles.

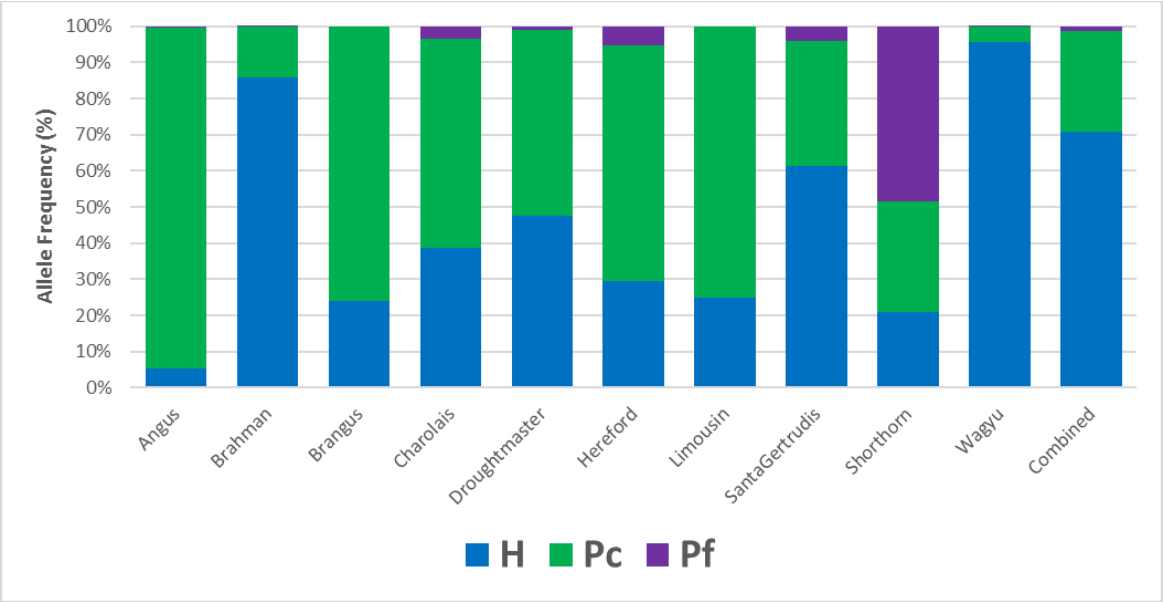


Fig. 7. Allele frequencies of Horn (H) Celtic Poll (Pc) and Friesian Poll (Pf) in various Australian breeds.

4.1.7 Phenotypic concordance of OPT-based results

An important question to consider is whether the predicted poll tested genotypes match with the actual phenotypes of the animals. For this purpose, the OPT-predicted genotypes and BREEDPLAN phenotypes (Horn, Scur or Poll) of 18, 417 animals were compared (Table 8). Current expectations are that the three head-status categories are expected to be due to genotypes status as below;

- **Horn** HH
- **Scur** HPc or HPf
- **Poll** HPc, HPf, PcPc, PcPf or PfPf

Out of 18,417 tested animals, 97.5% (n=17,962) phenotypes were found with the expected genotypes. However, 2.5% (n=455) did not align to expected genotypes based upon phenotypes available.

Specific observations of phenotypic concordance for each head-status are as following;

- Horned animals were 97.4% concordant with 2.4% genotyped as HPc.
- Scurred animals were 87.1% concordant with 12.2% genotyped as HH.
- Polled animals were 99.0% concordant with only 1.0% genotyped as HH.

It is important to note that the phenotypic records for head-status provided through BREEDPLAN can be inaccurate for several reasons, including recording errors, making wrong observations at a very early age before physical appearance of horns or scurs, disbudding and healing leading to a polled status or when scur and horns are not easily differentiated. While not the only source of discordance, it is proposed that these recording errors explain much of the discordances between head-status and OPT genotypes.

Table 8. Number of Head-phenotypes for various OPT-based genotypes and their phenotypic concordance (%). Breed-wise list is provided in Appendix 9.4.

OPT genotypes	BREEDPLAN Phenotypes and concordance (%)			
	Horn	Scur	Poll	Total
HH	11,528 (97.4%)	84 (12.2%)	60 (1%)	11,672
HPc	286 (2.4%)	579 (84.2%)	1852 (31.4%)	2,717
HPf	2 (0.016%)	20 (2.9%)	114 (1.9%)	136
PcPc	16 (0.1%)	4 (0.6%)	3531 (59.9%)	3,551
PcPf	1 (0.008%)	1 (0.1%)	311 (5.3%)	313
PfPf	1 (0.008%)		27 (0.5%)	28
Total	**11,834**	**688**	**5,895**	**18,417**

4.1.8 Accuracy of OPT results in validation population

For the validation population of 393 animals of varying ages which were carefully phenotyped by project staff and successfully OPT tested, all 393 animals were found concordant for phenotype and genotype (Tables 9 and 10). These results suggest that OPT has been very accurate. Note that a single animal was initially classified as horned with HPc genotypes (Fig. 4A). Upon reassessing with consultation with experienced industry leaders, the animal was reclassified as scurred, given the head-shape and small size of appendages even though attached to the skull. This is instructive as to the difficulties at times to distinguish horns from scurs, and inherent errors introduced. This animal was deemed worthy of further investigated by whole genome sequencing. Moreover, results of validation population and the BREEDPLAN concordance data suggest that some ambiguities remain to be understood about the mechanisms involving the heterozygote genotypes (HPc and HPf) which can develop as scurs or poll.

Table 9. Head-status phenotypes and OPT genotypes of the validation population.

Breeds	**OPT based genotypes**						
Head-status	**HH**	**HPc**	**HPf**	**PcPc**	**PcPf**	**PfPf**	**Total**
Brahman	**15**	**57**	**1**	**7**	**2**		**82**
Horn	15						15
Poll		28		7	2		37
Scur		29	1				30
Brangus	**3**	**4**		**11**			**18**
Horn	3						3
Poll		1		11			12
Scur		3					3
Crossbred	**24**	**12**					**36**
Horn	24						24
Poll		7					7
Scur		5					5
Droughtmaster	**30**	**87**	**3**	**40**	**2**		**162**
Horn	30						30
Poll		34	3	40	2		79
Scur		53					53
Hereford	**9**	**30**	**19**	**3**	**2**	**3**	**66**
Horn	9						9
Poll		15	10	3	2	3	33
Scur		15	9				24
Santa Gertrudis	**7**	**22**					**29**
Horn	7						7
Poll		9					9
Scur		13					13
Total	**88**	**212**	**23**	**61**	**6**	**3**	**393**

Table 10. Phenotypic concordance (%) with OPT Genotypes of validation population. Percentage of a phenotype (observed by project sources) detected with different OPT genotypes.

OPT results	**Phenotypic concordance (%) with OPT Genotypes**		
	Horn	**Scur**	**Poll**
HH	100.0%		
HPc		92.2%	53.1%
HPf		7.80%	7.30%
PcPc			34.5%
PcPf			3.40%
PfPf			1.70%
Total	**100%**	**100%**	**100%**

4.1.9 Implementation of the OPT within the genomic SNPchip assays and their success rate in commercial testing services

Through Dr Lyons' relationship with a major cattle genotyping lab in Australia and internationally, we have ensured that learnings from this study in relation to the most predictive assay (OPT) are now being applied routinely whenever poll testing is issued through that laboratory for any animals with an Indicus influence. Results have been published (Randhawa et al, 2019) that will allow other service providers to likewise adjust their assays. All 5 SNP are available on most commercially available SNP chips.

The commercial implementation of the OPT test has been very successful, with the overall rate of poll genotype determination at 99.42%. Of 75,031 samples tested across breeds and across a range of genomic and standalone testing options, only 434 animals were returned as No Result. Further investigation of these confirmed that absent genotypes for key markers is the predominant cause, usually associated with generally poor call rates or other QC issues indicating a sampling issue.

Pc and Pf mutations are approximately 200 kb apart and expected to be in trans-arrangements (residing on separate chromosomes) given their independent evolutions. This study has found no occurrence of cis-gametes (i.e. same chromosome carrying both Pc and Pf) in PcPf, PfPf and PcPc animals (n=3,978). However, there is non-negligible recombination probability (~0.2%) of a PcPf cis-gamete coexisting with H allele on the second chromosome which may cause growth of scurs as an unexpected phenotype. Further investigation is required to assess that possibility.

4.2 Investigation of scurs genetics (Objective 2)

Preliminary investigations have found that scurs development is influenced by some non-genetic factors such as sex of the animals. Therefore, to understand the scurs genetics in this project, the phenotypic concordance data were investigated for different aspects causing scurs development in cattle.

4.2.1 Scur expression based on sex (Male vs Female)

Of the total 18, 417 animals, 11,453 animals were identified with a known sex status (female, male and steer). For this analysis, steers were grouped with other males (Table 11). Any data where phenotype and genotype were discordant were excluded from this analysis as it was considered highly likely due to phenotyping error. The following results were observed related to scurs for animals identified as heterozygous (HPc or HPf);

- 622 females (HP) showed 87% (n=542) to be polled and 13% (n=80) to be scurred.
- 1725 males (HP) showed 70% (1206) to be polled and 30% (n=519) to be scurred.
- The differences in the rate of scurs development between females and male were statistically significant (Fisher exact test, $p < 0.00001$).

4.2.2 Scur expression based on polled genotypes (HPc vs HPf)

The poll alleles (Pc and Pf) appear to influence scurs development (Table 11). Results observed include:

- HPc animals (n = 2,213) developed scurs in 26% (n=579) of cases
- HPf animals (n = 134) developed scurs in 15% (n=20) of cases

These differences in scur development between the Celtic and Friesian alleles were statistically significant (Fisher exact test, $p = 0.003$). These observations support results of a very recent publication (Gehrke *et al.* 2020) who noted that in their dataset the Pf allele supressed the development of scurs more efficiently than the Pc allele.

When considering sex and allele effects combined:

- Allelic effects of scur development within male cohorts were also found significant (Fisher exact test, $p = 0.0058$).
- Allelic effects of scur development within female cohorts were found non-significant (Fisher exact test, $p = 0.159$).

This indicates that female cattle are most likely to be polled with either of the HPc or HPf genotypes and suggests that sex is a stronger influence on phenotype that allelic effects. However, more animals are required to confirm this is the case.

Table 11. Distribution of head-status phenotypes in Female and Male cattle. Red text highlights results considered highly likely to be due to phenotyping error. These animals were excluded from the female versus male comparison.

Sex of animal	**OPT based genotypes**						
Head-status	**HH**	**HPc**	**HPf**	**PcPc**	**PcPf**	**PfPf**	**Total**
Female	**3,359**	**635**	**30**	**506**	**42**	**6**	**4,578**
Horn	3,332	43		2		1	3,378
Poll	16	513	29	504	42	5	1,109
Scur	11	79	1				91
Male	**3,076**	**1,864**	**106**	**1,550**	**257**	**22**	**6,875**
Horn	2,959	243	2	14	1		3,219
Poll	44	1,121	85	1532	255	22	3,059
Scur	73	500	19	4	1		597
Total	**6,435**	**2,499**	**136**	**2,056**	**299**	**28**	**11,453**

4.2.3 Regions of interest related to scurs found in the previous studies

For the investigation of the scurs for gene mapping, a set of all previously known regions of interest in cattle were compiled (Table 12). To date, seven genomic regions localized on six chromosomes, designated *Bos taurus* autosome (BTA). BTA 4, 5, 12, 16, 18 and 19 have been proposed by different studies. Interestingly, none of these regions have been validated in any independent population. In addition, four other regions on BTA 2, 9, 10 and 23 have been mentioned as suggestive in various research articles although genomic positions have not been provided. This project investigated the high-density 770K SNP genotypes to detect putative regions in the bovine genome.

4.2.4 Candidate genome regions for scurs by GWAS and CSS analyses

A total of 657,543 SNPs, which were mapped to 29 *Bos taurus* autosomes (BTA) and passed the quality control (minor allele frequency ≥ 0.5 and genotyping call rate ≥ 0.95) were used to find scurs regions through both GWAS and CSS approaches.

The control analyses successfully detected the poll locus on BTA1 by GWAS and CSS. However, the GWAS scores for breed-wise analyses did not reach the genome-wide significance threshold, indicating that GWAS was sensitive to lower sample size (n=8) in both cohorts (Polled vs Horned) of the control analyses.

The discovery analyses by GWAS failed to identify any significant regions for scurs development, although several suggestive regions were detected. However, the CSS analyses (Fig. 8) identified

multiple regions in the combined as well as in the breed-wise datasets. These include four CSS peaks that have been identified as regions of interest in previous research (Table 12).

Neither GWAS or CSS identified regions or genes that were strongly associated with the scurs phenotype. In attempting to minimise the risks of inaccurate phenotypes, these analyses were restricted to a relatively small number of animals. With more samples, it may be possible to increase the power to identify true associations. However, to do so would require a large investment of resources beyond the scope of this project to ensure accurate and reproducible phenotypic data sets.

These investigations give support to the findings of Gehrke *et al.* (2020) who suggest that the development of scurs is a polygenic trait affected by several genes and/or other factors (e.g. non-coding RNA), as well as possible non-genetic effects. There does not appear to be a simple mono-genetic inheritance model as initially proposed by White and Ibsen (1936).

Table 12. Regions of interest for scurs development in cattle. Genomic positions are based on ARS-UCD1.2 bovine genome assembly. Regions marked with # also identified in current study.

BTA	Position (Mb)*	Genes	Discovery breeds	Reference
4	27 – 28	*TWIST1* (Type 2 scurs)	French Charolais	Capitan *et al.* (2011)
5#	44– 45		Holstein Friesian	Gehrke *et al.* (2020)
12	7.5 – 8.5		Holstein Friesian	Gehrke *et al.* (2020)
16#	40 – 41	*SUCO*	Holstein Friesian	Gehrke *et al.* (2020)
18#	46 – 47.5	*ARHGAP33*	Holstein Friesian	Gehrke *et al.* (2020)
19	26 – 29	*CTDNEP1, SHBG, SOX15, FGF11, DHRS7C*	Canadian Beef cattle	Asai *et al.* (2004); Ketel and Asai-Coakwell (2020)
19#	48 – 49		Simmental	Tetens *et al.* (2015)
19	-	Failed to validate chromosome 19 region	French Charolais	Capitan *et al.* (2009)

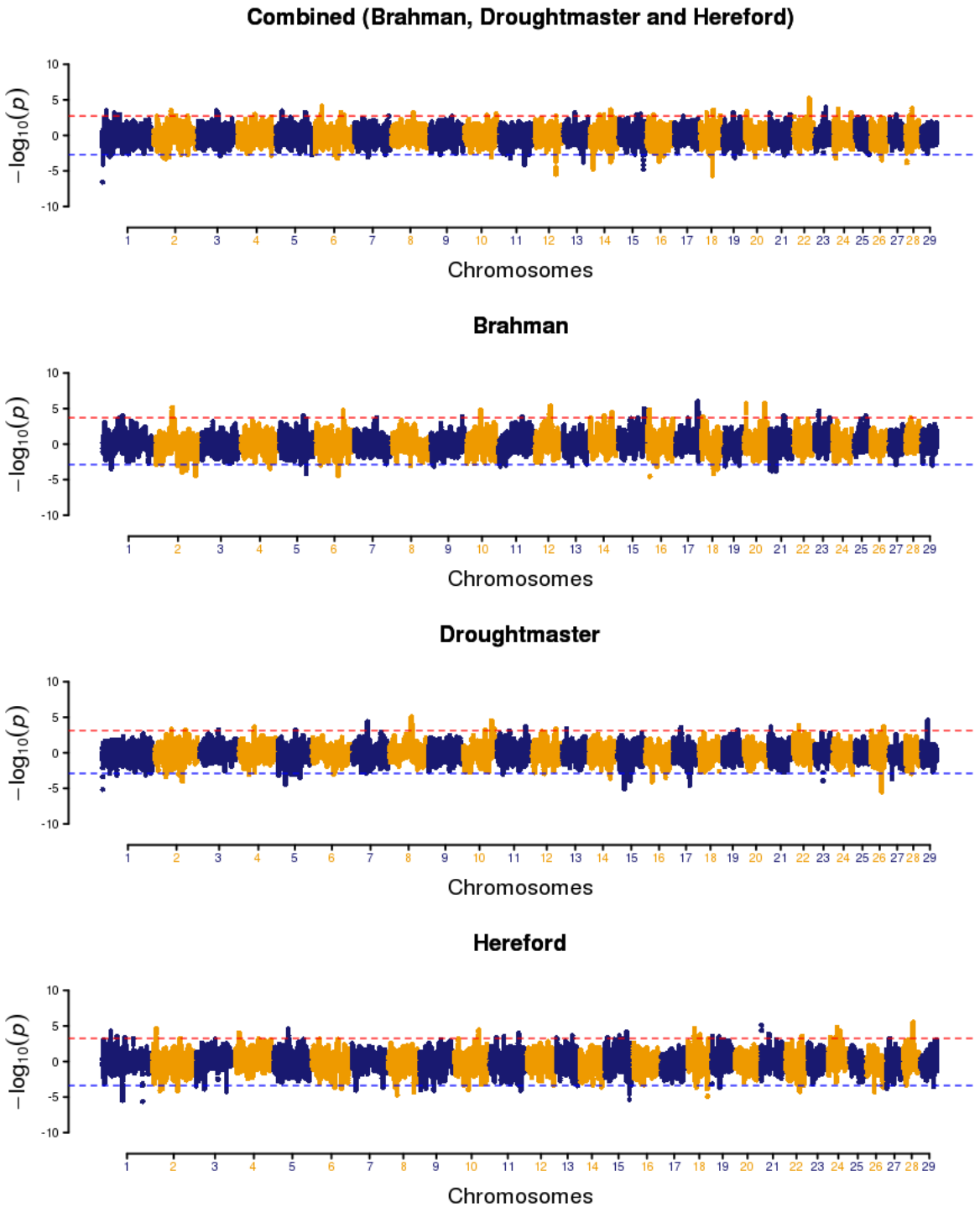


Fig. 8. Manhattan plots of CSS results in combined and breed-wise data using 657,543 SNPs. Peaks above-red and below-blue dashed lines capture candidate regions for polledness and scurs, respectively at the significance thresholds of 0.1%.

4.2.5 Whole-genome sequencing for scurs investigation

Sequencing of the obtained samples have been completed and data analysis is in progress by the collaborators at the CSIRO.

4.3 Polledness effects on production and reproduction (Objective 3)

The past two decades have seen changing preferences of the commercial beef producers and feedlots in Australia for polled cattle (Fig. 1) due to increased awareness of animal welfare, consumer choices and costs and risks associated with physical dehorning. Hence, the genetic merit (EBVs) for the recorded traits in the eight breeds studied have experienced significant changes within the last two decades. Comparing the EBVs distributions of polled and horned cohorts born between 2000 and 2018 will provide insights about the practical outcomes and statistical significance (t-tests) of producing polled cattle by computing the effect sizes (Cohen's d) and correlations (point biserial r) with polledness for a particular trait.

4.3.1 Four traits provide baseline comparisons

The BREEDPLAN database provides access to all commercial cattle breeds in Australia with a range of traits including production, reproduction, carcase, behaviour etc. However, all breeds are not recorded uniformly for each measurable trait. Hence, only a few traits are common in the eight selected breeds. Four traits were chosen that are representatives of major trait classes of production (Birth Weight and Mature Cow Weight) and reproduction (Scrotal Size and Gestation Length). A comprehensive comparison of these traits across the eight breeds at various thresholds of accuracy (%) of BREEDPLAN EBVs assessed whether the traits were favourable in polled versus horned animals and their effect size (Fig. 10). Effect size is a quantitative measure of the magnitude of the effect. The larger the effect size the stronger the relationship between two variables.

Important observations include:

- **Birth Weight EBVs** were generally significantly higher in the horned cohorts (Bonferroni corrected $p = 0.05$) in five breeds (Brahman, Charolais, Hereford, Limousin and Simmental). , Effect size ranged from 0.2-0.9, and relatively small correlations (0.05-0.4) The other three breeds (Droughtmaster, Santa Gertrudis and Shorthorn) had significantly higher EBVs for the polled cohort but with very small effect sizes ($d < 0.15$) and correlations ($r < 0.1$).
- **Mature Cow Weight EBVs**, interestingly, showed the opposite trend to that of the Birth Weight in most breeds. For instance, polled cohorts had significantly higher mature cow weight EBVs (d = 0.1-0.8, r = 0.05-0.4) in five breeds (Droughtmaster, Santa Gertrudis, Hereford, Limousin and Shorthorn) compared to the horned cohort. In the other three breeds (Brahman, Charolais and Simmental) the trait remained positively associated with horned status. However, the effect sizes ($-d$ = 0.04-0.2) and correlations ($-r < 0.01$) became very small. This showed that polled animals in several breeds which started with a lower Birth Weight were able to compensate with a higher growth rate at later stages of life (see section 4.3.4).
- **Scrotal Size EBVs** were consistently observed to be significantly higher (d = 0.15-0.9, r = 0.1-0.4) in polled animals in all breeds except Simmental (d = -0.2, r = -0.09).
- **Gestation Length EBVs** were consistently negatively associated with polledness. This is desirable, with lower gestation lengths being selected for to increase fertility in herds. Therefore, polled cohorts have shown better performance than the horned beef cattle cohorts.

Overall, the trends of four traits have shown that polledness has no detrimental effects on production and reproduction traits, and indeed, for most traits are statistically superior (Fig. 9). However, the effect size in all cases is relatively small such that there would be no significant genetic gains for evaluated traits through selection of polled animals over horned animals.

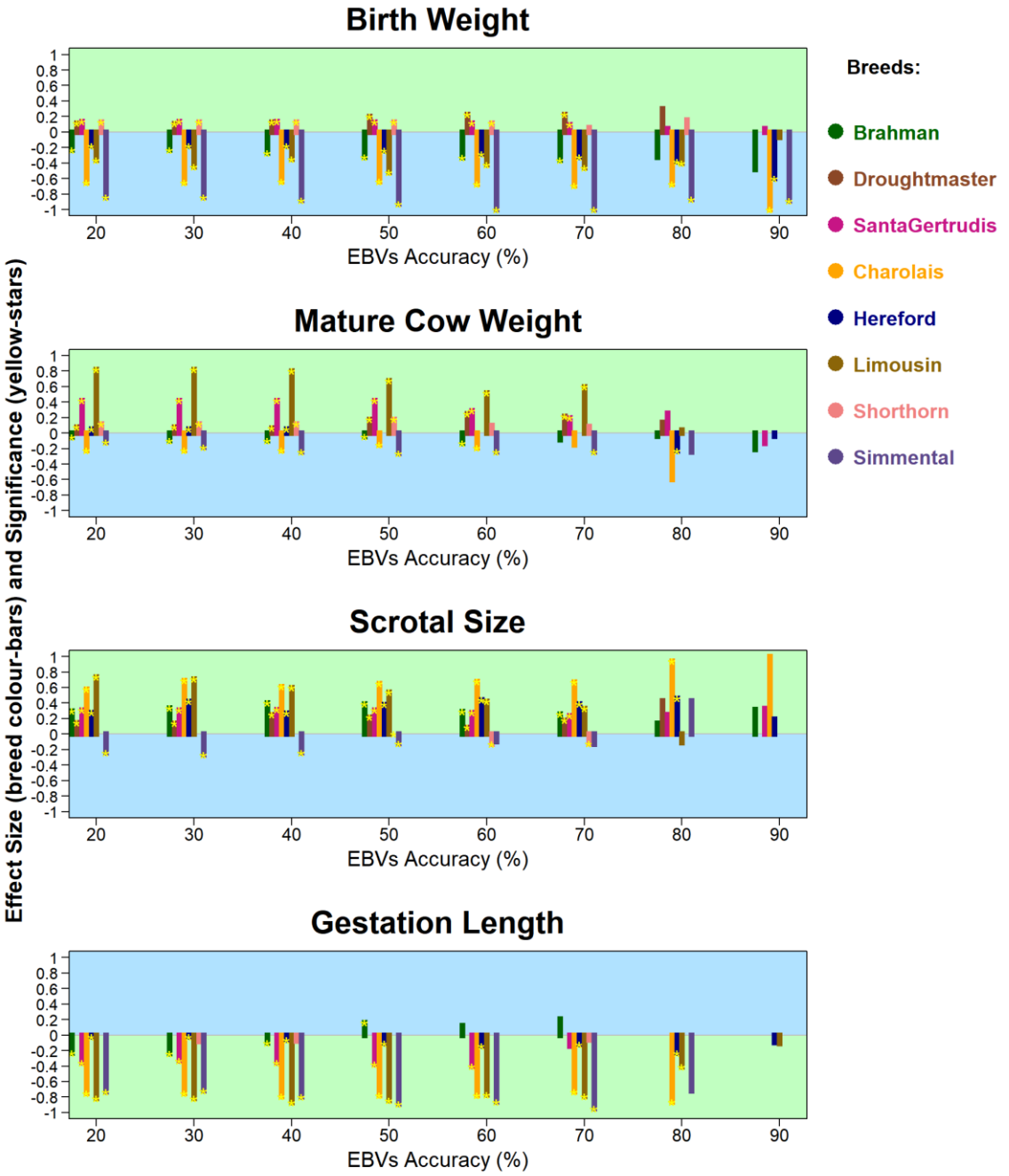


Fig. 9. Statistical significance and effect size (Y axis) measured for four traits in polled and horned cohorts of eight breeds using BREEDPLAN EBVs above different levels of accuracy. Where the traits are favourable in the polled cohort, the lines will be in the Green (polled) zone, and conversely when favourable in the horned cohort, the lines are in the Blue zone. Yellow stars indicate statistically significant differences between cohorts.

4.3.2 All traits compared at EBVs accuracy thresholds of 50% and 70%

As can be seen in Fig. 9, EBV statistics (*d, r*) were consistent at accuracy thresholds of 20 to 60%. However, as the accuracy levels increased up to 70% and above, several traits started to decrease the magnitude of those statistics (*d, r*) and most trait-wise EBVs can be found non-significant. It indicated that accuracy thresholds are important considerations for the comparison of EBVs between polled and horned cohorts. While using EBVs with higher accuracies is intuitively a better comparison, smaller sample sizes are an inevitable consequence (Appendix 9.5). Therefore, further analyses were performed at two different levels of EBVs' accuracy, these being 50% (Appendix 9.6) and 70% (Appendix 9.7), respectively.

The breed-wise results of published EBVs for a range of traits covering fertility, calving ease, milking ability, growth and carcase merit suggested that polled and horned genetic effects varied for these traits and the trends were not consistent across the eight breeds (Appendix 9.6; Supplementary Tables S2-S9). Moreover, it was observed that as the EBVs accuracy were increased to 70% many of the significance differences were not sustained Appendix 9.7; Supplementary Tables S10-S17). Additional analyses by using the OPT genotyped animals demonstrated that increasing the accuracy for head-status classifications can further reduce the number of traits with significance differences between the polled and horned cohorts (Fig. 10; Appendix 9.8). Furthermore, OPT-based analyses reinforced that the direction of effect sizes were not consistent across breeds, indicating that these variations may not be linked to head-status. Overall, the effect size (*d*) and correlation (*r*) were consistently in the small range such that statistically significant difference will have very small practical significances.

4.3.3 Average effects of polledness on all traits

Selection decision by using EBVs are vital to achieve a balance between the different groups of traits and to place emphasis on those traits that are important to the herd, markets and environment. Considering all traits equally important for each breed, the trait-wise statistics were averaged within each breed (Table 13). The results suggest that breeding for polledness was favourable (positive mean values) in all breeds except Droughtmaster. It is also noted that higher levels of EBVs accuracy (70%) results in reduced variation (lower effect size and correlations) among the polled and horned cohorts. All in all, this project found that polledness is not a negatively associated characteristics with any of recorded phenotypes in the beef cattle industry.

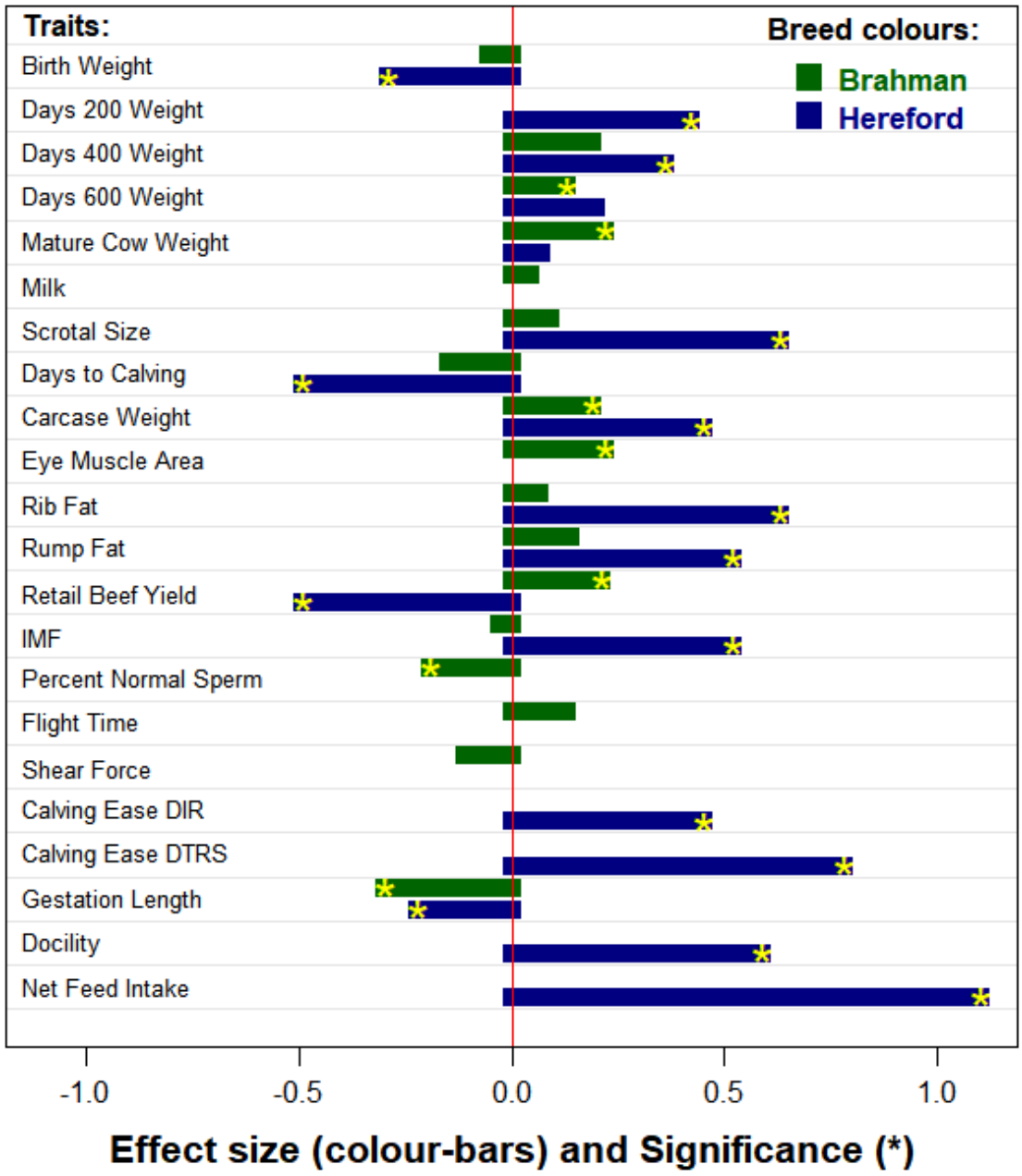


Fig. 10. Statistical significance and effect size (colour-bars) for trait-wise comparisons between polled and horned cohorts in Brahman and Hereford, using only animals with phenotypes (Horned, Polled) consistent with OPT predictions (see descriptive and inferential statistics in Appendix 9.8). Yellow stars indicate statistically significant differences between cohorts.

Table 13. Standardized Mean Effect size and Point Biserial Correlation (PBC) of polledness on other traits.

Breeds	EBVs accuracy ≥ 50%			EBVs accuracy ≥ 70%		
	Traits (N)[A]	**Effect Size** (Mean ±SD)	**PBC** (Mean ±SD)	**Traits** (N) [A]	**Effect Size** (Mean ±SD)	**PBC** (Mean ±SD)
Brahman	18	0.057±0.26	0.013±0.07	17	0.037±0.31	0.010±0.09
Droughtmaster	13	-0.012±0.19	-0.001±0.07	12	-0.023±0.20	-0.013±0.08
Santa Gertrudis	17	0.200±0.21	0.087±0.09	17	0.130±0.17	0.062±0.08
Charolais	16	0.200±0.48	0.090±0.21	16	0.260±0.70	0.100±0.26
Hereford	19	0.130±0.28	0.049±0.11	19	0.100±0.44	0.032±0.18
Limousin	17	0.540±0.44	0.240±0.19	17	0.300±0.47	0.130±0.21
Shorthorn	17	0.019±0.13	0.003±0.03	14	0.039±0.21	0.005±0.03
Simmental	18	0.200±0.71	0.077±0.29	17	0.110±0.69	0.032±0.28

[A] Number of traits with recorded EBVs (above respective accuracy threshold) in each breed. Note that four traits (Days to Calving, Shear Force, Gestation Length and NFIF), which are considered favourable at lower (negative) EBV values, were standardized (switched values for +/-) to account for their favourability to compute the breed-wise averages of Effect Size (Cohen's d) and PBC of all traits. Hence, positive values show an overall favourable effect of polledness within each breed. We also note that two other traits (Rib Fat, Rump Fat) can be considered favourable depending on the breeding objectives of a herd/breed. Therefore, Appendix 9.10 shows values after removing these two traits, which showed high correlation (Pearson's) with the values in this table.

4.3.4 Polled animals can potentially catch up the growth rate

One of the earlier observations was that polledness may be negatively associated with early in life traits (Birth weight, yearling weight etc.), but compensate in later life at higher levels of growth rate. This observation was further investigated by using the six growth-related traits (weights of cohorts at birth, 200 days, 400 days, 600 days, mature cow and carcase) in eight breeds using EBV accuracy thresholds at 50% (Fig. 11), 70% (Fig 12) and 75% (Appendix 9.9). Results were consistent with the previous observations. The trends showed that polled animals have genetic merit to perform equal to or better than horned animals in later stages of life traits across all breeds. This further reinforces the conclusion that polledness is not a detrimental characteristic.

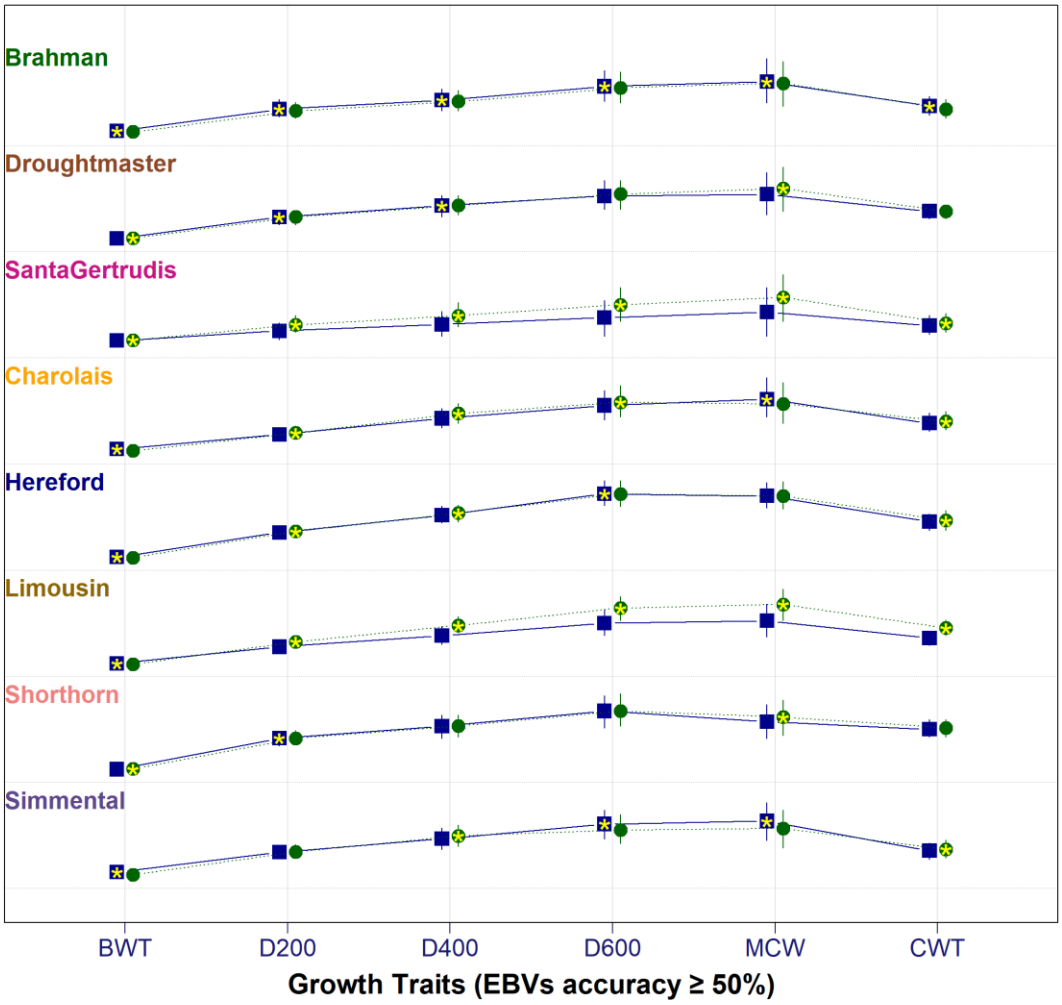


Fig. 11. Comparison of BREEDPLAN EBVs (Accuracy ≥ 50%) of growth (weight) traits at different life-stages in eight breeds. Blue squares and green circles represent the means of the horned and polled cohorts, while vertical lines in each square and circle show the range of 1st and 3rd quartiles. A steric (*) in a square or circle represents a statistically significant, higher value.
Trait abbreviations are **BWT**: Birth weight; **D200**: Days 200 weight; **D400**: Days 400 weight; **D600**: Days 600 weight; **MCW**: Mature cow weight; **CWT**: Carcase weight.

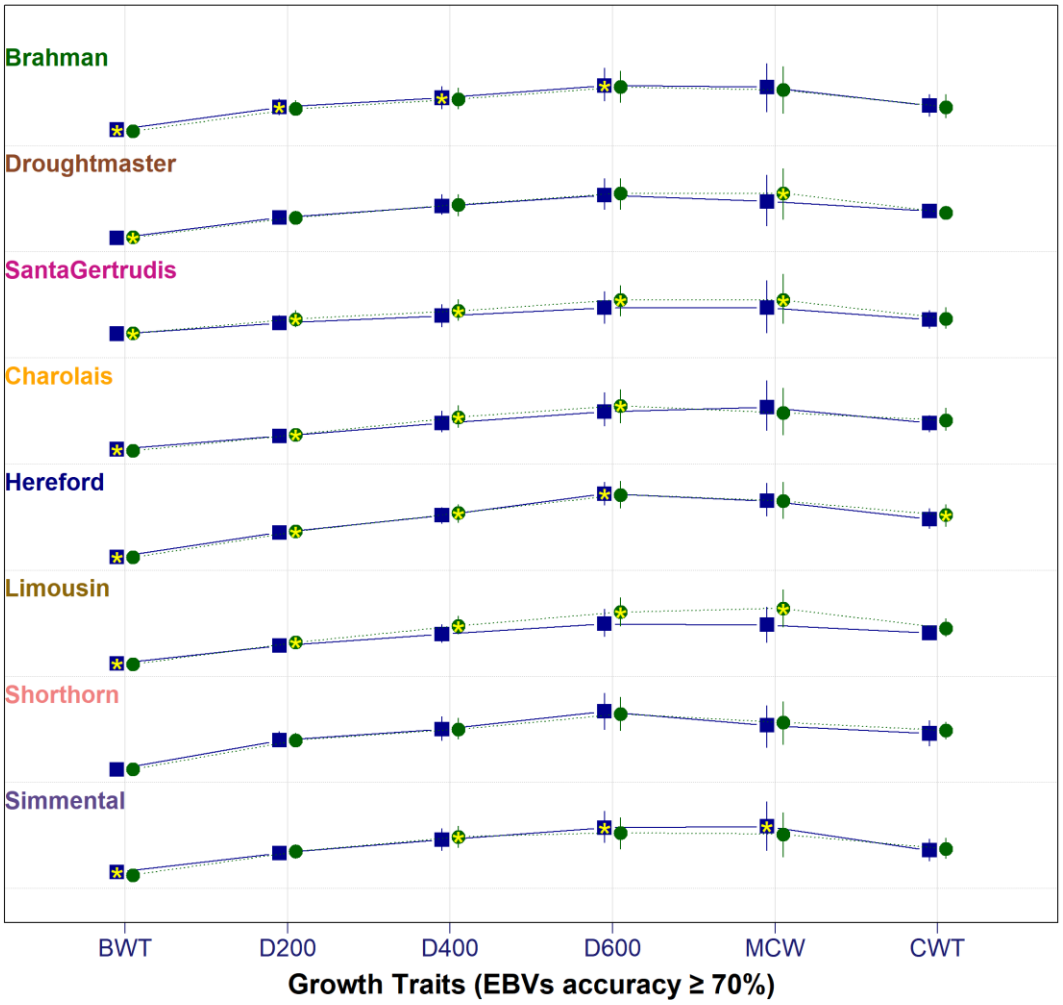


Fig. 12. Comparison of BREEDPLAN EBVs (Accuracy ≥ 70%) of growth (weight) traits at different life-stages in eight breeds. Blue squares and green circles represent the means of the horned and polled cohorts, while vertical lines in each square and circle show the range of 1st and 3rd quartiles. A steric (*) in a square or circle represents a statistically significant, higher value.
Trait abbreviations are **BWT**: Birth weight; **D200**: Days 200 weight; **D400**: Days 400 weight; **D600**: Days 600 weight; **MCW**: Mature cow weight; **CWT**: Carcase weight.

5 Discussion

5.1 Optimized polled testing assay

Economic sustainability and animal welfare have driven recent progress in modern livestock production systems especially in efforts to minimize or eliminate undesirable traits such as the presence of horns in cattle. Genetic dehorning is being progressively adopted (Fig. 1) as the non-invasive approach to breed hornless cattle through genetic selection (Carlson *et al.* 2016; Mueller *et al.* 2019). Genotype-phenotype relationships of horn growth are however complex (Medugorac *et al.* 2012) limiting the informativeness of poll gene testing assays for early detection for polled alleles in some breeds. This project identified limitations of the commercial poll gene testing assays, especially in tropical cattle common in Northern Australia and throughout Asia. This project successfully delivered the optimized poll testing (OPT) assay – based on a robust set of 5 SNPs – which can effectively eliminate the ambiguous and undetermined results that previously limited the effectiveness of both SNP-based and MSAT-based poll predictions.

Accuracy of horn phenotypes is a major challenge for scurs research, as evident in this and other recent studies (Connors *et al.* 2018). Future efforts should focus on providing industry with better guidelines on how and when to record head status, and at differentiating horn from scurs. Without these accurate phenotypes, and at sufficiently high numbers per category, genomic studies are difficult and likely to generate false leads. This was evident in the current study. Phenotypic records on a subset of 18,417 animals were obtained from the BREEDPLAN database to test concordance between phenotypes and genotypes. OPT-based genotypes have shown high concordance with known head-status of horned as 97.5% HH and polled as 99% PP (PcPc, PcPf or PfPf) animals. However, the potential for scur phenotypes in heterozygous animals (HPc and HPf) significantly compromises the informativeness of these genotypes. It is very unlikely that HH animals can be either scurred (84 out of 11,672; 0.72%) or polled (60 out of 11,672; 0.51%), and these results likely reflect errors due to imprecise phenotyping at a very early age, improper dehorning resulting in partial regrowth of horns mimicking scurs and data recording errors.

The inaccuracy of phenotype recording was evident in the BREEDPLAN data set. In the validation data, phenotype-genotype concordance was found as expected with HH and PP, but heterozygotes HP corresponded to 54.5% scurred and 45.5% polled animals. Some rare cases were observed (Fig. 4), such as an HPc Brahman animal with horn-like fixed scurs. A farmer also reported three Droughtmaster PcPc bulls, initially recorded as polled, that grew small scurs after two years of age. These cases are being investigated through whole-genome sequencing to find out whether they carry a unique genetic variant, although the data is suggesting that other factors, such as epigenetics or other unidentified biological and/or environmental triggers, can lead to unexpected rare combinations of phenotypes and poll gene test. On the other hand, the vast majority of PcPf animals (n=313) were found to be 99.4% polled (Table 8).

Overall, OPT was shown to significantly increase the prediction rates from 90.3% to 99.42% in 75,031 commercially tested samples. Remaining No Result samples were due to incomplete genotypes. The number (0.04%) was below the expected failure rate as observed in previous research (Wu *et al.* 2019).

Economic benefits to industry directly as a result of increased accuracy and reduced costs, and based upon assumptions below, are conservatively estimated at more than $400,000 per annum. SNP-based testing can be, and is, routinely bundled into value-added genomic packages including genomic data (50,000 SNP is common), parentage, breed-relevant defects and traits, generally available through the relevant breed societies. If extracted and translated from this genomic data retrospectively, one commercial service provider costs this service at $5 per sample. That compares

to an average cost of $25 per sample tested using available stand-alone MSAT or SNP poll tests. Therefore, the price differential between OPT and available stand-alone poll tests available is estimated at $20 per sample. It is conservatively predicted that more than 20,000 poll tests will be requested in any 12-month period, based on recent genomic testing. Based upon these figures, industry will save upward of $400,000 per annum with greater than 99% poll test success rates across all breeds using OPT. This excludes benefits of polled herds including reduced horn management costs and risks, increased animal and worker welfare, and less downgrading of horn-damaged meat and hide, which would also be significant but are not quantified in this report.

5.2 Understanding scurs

The genetics underpinning scurs development remains largely unknown, and global research efforts have to date proven unsuccessful in identifying diagnostic markers predictive of scurs. Previous research has shown that variation in rate of development, size and shape of horns, scurs and skull is affected by non-genetic factors such as sex, nutrition, age and photoperiod (Randhawa *et al.* 2019). While OPT is an excellent tool for prediction of whether horns will form or not, scurs development cannot at this time be predicted, though it appears most commonly in heterozygous (HPc, HPf) animals. This project investigated the non-genetic and genetic factors in targeted populations by using high-density SNP genotypes, whole genome sequencing and accurate phenotyping.

In this study the scur phenotype is significantly influenced by an animal's sex. Scurs were much more common in heterozygous male than females in the data analysed, an observation consistent with other studies (Long and Gregory 1978; Asai *et al.* 2004; Gehrke *et al.* 2020; Ketel and Asai-Coakwell 2020). All recent studies, including the present research, points to complex inheritance patterns involving many genetic factors and/or non-genetic effects.

Several extensive efforts to map the scur gene have identified genomic regions potentially involved through various bone development and hormone regulation (Table 12), though disappointingly none have proven reproducible across studies and none of the proposed regions were coincided to a single or a few common genes. This project found some genes – *SUCO* (BTA16) and *ARHGAP33* (BTA18) – and non-genic regions (BTA5, BTA19) that overlapped with previous research. Several other genomic regions were identified by CSS suggesting that many genes interacting through complex polygenic networks may control scurs development, rather than one or two genes of major effect. GWAS did not identify any regions associated with scurs in the project data. Further investigation is in progress to compare the proposed genomic regions across various studies to find any genomic structural variants related to scurs development.

Although scurs are substantially less common than horns (Fig. 1), and are less damaging and easily manageable, they will continue to appear as the beef breeds transition from horned to polled cattle. However, with increased frequencies of polled alleles (Pc and Pf), and as more homozygous polled breeding stock become available, the incidence of scurs is expected to decrease. While efforts to date to identify diagnostic scurs-associated markers have been unsuccessful, further anticipated reduction of horns in herds in years ahead may allow re-investigation with more accurate phenotypes as horns become less common.

Interestingly, genetic heterogeneity (Pc or Pf) at the poll-locus appears to influence the rate of scurs development in heterozygous animals (Gehrke *et al.* 2020). This project has confirmed this recent observation (Gehrke *et al.* 2020) that scurs were more likely in HPc animals than HPf animals. Efforts to increase the frequency of the rarer Pf allele within and across herds, most common in Shorthorn cattle, could lead to further reduction of the scur phenotype.

5.3 Polledness effects on production and reproduction

Previous research on the impact of polledness on production and fertility traits of different breeds and cross-bred cattle have generally shown no significant difference for production traits, including live weight, growth rate, carcass weight and quality, dystocia, fertility and mortality rates (Frisch *et al.* 1980; Stookey and Goonewardene 1996; Kommisrud and Steine 1997; Goonewardene *et al.* 1999). Results in the current project, based on the estimated breeding values (EBVs), generally support previous findings. Most of the significant differences were found to regress at higher thresholds of EBVs' accuracy (>70%) and when the contrasting cohorts (polled versus horned) were based on OPT-genotypes to account for the head-phenotype misclassifications. Even when some sustained trends showed significant differences, these were of little practical impact as had very small effect sizes. Moreover, several EBVs trends were not consistent towards one or other phenotype (polled or horned) across all breeds and may indicate head-status is not a direct cause for differences observed.

Interestingly, reproduction (scrotal size, gestation length, days to calving, calving ease and sperm quality) and behaviour (flight time and docility) traits have consistently shown better genetic merit for the polled cohorts. Another promising observation in favour of poll breeding was that the polled-cohorts that had lower EBVs at early age (birth weight and subsequent growth to weaning) showed above average EBVs for later growth and became as good as or better than their horned contemporaries, which was consistent with previous research (Frisch *et al.* 1980). The data analysed herein conclusively shows that poll status across all breeds is not negatively associated with EBVs for measured trait. An increased prevalence of the polled condition within herd or industry-wide would not be expected to have a measurable negative influence on production, carcase, fertility and behaviour traits.

The greater risk is in loss of diversity or inbreeding effects in breeds with few polled animals (e.g., Brahman, Santa Gertrudis and Wagyu), and this will need to be managed carefully until there are larger numbers of polled breeding bulls and cows. Inbreeding depression, as well as the potential increase of detrimental alleles and associated disease, are risks. Therefore, care must be taken to avoid high levels of relatedness and inbreeding. Fortunately, breeding high genetic merit polled animals has become a realistic perspective with the help of new tools being used for genomic selection and precision breeding. Gene editing technologies (e.g. CRISPR) may also have a role to play in introducing more genetic diversity into these herds, depending upon federal regulations.

5.4 Success against Objectives

Deliverables against the main research objectives are outlined below:

1) **A more cost effective, SNP-based test based on haplotypes, which can be incorporated into multi-purpose SNP panels**.
    - Markers specific to Indicus breeds were identified that rendered some SNP as non-informative or inconsistent in the earlier version of SNP poll test (SPTv1) SNP). These have been removed in a new optimised poll test (OPT).
    - Results generated with OPT did not change genotypes for any of the 18,637 samples previously tested successfully with SPTv1.
    - OPT resolved 99.5% of previous "No result" samples.
    - Correlations between predicted genotypes and observed phenotypes is very strong.
    - Markers are available on most commonly used genomics SNP chips, and the new assay has been implemented in a major service provider lab as both stand-alone and bundled (multi-purpose) offerings.

- Significant savings are possible through development of multi-purpose SNP panels (bundles), including poll.
- Outcomes have been published and is available for any lab that may choose to implement this improved poll test.

2) Elimination or accurate prediction of carriers of the scur gene.

- Scurs remains an elusive trait to understand.
- Study is hindered by a lack of consistency in recording, largely due to the variability of Scurs phenotype, including some being difficult to distinguish from horns, as well as emergence at different ages in different animals.
- No genes, genomic regions or other markers were able to be identified that could be used as a diagnostic test to predict scurs.
- Scurs are most likely to occur in heterozygous (HP) animals
- Scurs is sex-linked, with rate of Scurs at 30 and 13% in HP males versus HP females respectively.
- Scurs were more likely in HPc animals than HPf animals, suggesting the Pf allele (rare all breeds except Shorthorn) is superior for preventing scurs.

3) **Determining whether there are negative effects of the Poll gene on reproductive phenotypes**.

- Across multiple breeds, and across all available recorded BREEDPLAN traits (EBVs; maximum of 22), no negative effects were observed in polled versus horned animals. This included reproduction-associated traits e.g. Scrotal Circumference and Gestation Length.
- For most traits, polled animals were considered favourable to horned animals.
- Effect sizes were low, meaning the benefits of selecting for or against poll for these traits is very low.
- All indication is that polledness has no detrimental effects on production and reproduction traits and should not be considered a risk in increased selection of polledness across herds.

.

6 Conclusions and Recommendations

Poll testing has changed dramatically in recent years including a transition away from microsatellites to SNP based predictions. The microsatellite test employed from 2014-2018 was a valuable tool and allowed many producers to employ selective breeding to successfully increase polledness in their herds. Similarly, the earliest iteration of the SNP Poll test worked very well for most taurine breeds but was found lacking in Brahman and other indicus-influenced composite breeds, with large numbers of unassignable animals (No results). Given the early SNP assays were based upon a taurine reference genome, variation in Brahman cattle seemed a likely reason for these problems.

This was largely the catalyst for this research project which looked to identify a more universally predictive test for those breeds not well served by the tests available at that time. The current project has succeeded in identifying mutations that were responsible for inaccuracies or the inability to determine poll status in all common breeds, particularly Brahman and other indicus breeds. Importantly, prior SNP-based poll results were not compromised or changed by the new test, leading to a seamless transition to the new improved assay.

Anecdotally, some in the industry have indicated a resistance to use poll genetics for fears of negative outcomes for other production or fertility traits. Consistent with several previous studies and based upon an extensive dataset of animals and 22 traits, the project concludes there is no negative association of polledness with any of the measured traits in all breeds assessed. While the project team do not advocate breeding for polledness solely without consideration for balancing of traits or target market specification, there is no quantifiable reason to breed horned animals beyond personal preference or lack of suitable polled genetics to achieve breeding objectives. With the later point in mind, it is recommended that Brahman, Santa Gertrudis and Wagyu breeders carefully monitor inbreeding in their herds, especially where there are currently limited polled animals to select.

Scurs remains a problem, and this study was unsuccessful in identifying markers that would be diagnostic of the phenotype. Data presented does confirm that both genetic and non-genetic determinants affect an animal's likelihood of developing scurs. Heterozygous (HP) males are more likely to produce scurs than females, and interestingly the most common form of polled allele (Pc), when present with a horn allele (HPc) is more likely to result in scurs than when a Friesian polled allele in present (HPf). The later finding is suggestive of genetic heterogeneity within the polled region on chromosome 1 and may serve as a beacon for future studies of targets within the region. However, to date no studies have identified mechanisms or triggers to explain how and when scurs develop. Despite this, increased frequencies of Pc and especially Pf across and within herds will be expected to have the double benefit of less horn and scur phenotypes.

Transitioning away from the previously used MSAT and SNP poll tests to the SNP assay developed though this project, especially when used in conjunction with genomic tests, offers significant savings for the industry, estimated at greater than $400,000 per annum in testing costs alone. This does not include additional value derived from removing or reducing costs associated with dehorning/disbudding, lost productivity or reduced meat and hide quality, nor societal value through improved animal welfare.

7 Key Messages

A key message that needs to be better explained is that while the Optimised Poll test is a significant improvement on forerunner tests, there is no perfect poll test that will deliver the correct result 100% of the time. Why?

- No current markers are confirmed as the cause of polledness.
- These markers are however closely associated, and data presented shows they correctly predict the phenotype most of the time.
- In recent years the scientific community has, and will continue to, make significant gains in knowledge on this trait.

Scurs remains poorly understood. However,

- Data shows it is predominantly observed in heterozygous animals (HP).
- More likely to occur in males than females.
- Genetic factors such poll allele present (Pc vs Pf) affect likelihood of developing scurs.
- All evidence to date is that many genes, not 1 or 2, are involved.
- Non-genetic factors such as nutrition and environment may also play a role
- Age of Scurs development is variable and not understood.
- By selecting for polled animals (ideally using PP bulls) over a few generations, scurs and horns will both decline.

Breeding with polled animals will not compromise fertility and production traits. However,

- Selection on polledness alone is not recommended.
- Management of inbreeding is required.

7.1 Project contributions and acknowledgements

This project was completed with contributions from Dr. Imtiaz Randhawa (Postdoctoral research fellow), Dr. Russell Lyons (Principal Investigator), Prof. Michael McGowan, Prof. Ben Hayes, Dr. Laercio Porto-Neto and the late Dr. Brian Burns.

We would like to acknowledge the contributions to the Project of the following organisations, groups and individuals;

- MLA funding
- Collaborating organisations and staff at CSIRO and QAFFI
- Staff at the AGL/Neogen DNA lab, with special mention to Dr Karen Schutt.
- BREEDPLAN
- Breed Societies and their members
- Dr Michael Flynn (Valera Vale) and Dr Matthew Kelly (AACo) for samples and practical insights.

7.2 Scientific Output

The following scientific output were achieved for the project:

- ✓ Randhawa, IAS, Burns, BM, McGowan, MR, Porto-Neto, LR, Hayes, BJ, Ferretti, R, Schutt, KM, Lyons, RE (2020) Optimized Genetic Testing for Polledness in Multiple Breeds of Cattle. *G3: Genes|Genomes|Genetics* **10**, 539-544.
- ✓ Randhawa, IAS, McGowan, MR, Porto-Neto, LR, Hayes, BJ, Lyons, RE (2019) DNA Testing and Genetic Evaluation for Poll Breeding in Tropically Adapted Beef Cattle. *Proceedings* **36**, 98.
- ✓ Randhawa, IAS, Lyons, RE, Hayes, BJ, Porto-Neto, LR, McGowan, MR (2019) Factors affecting development of horns and scurs in domesticated ruminants. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* **23**, 484-487.
- ✓ Randhawa, IAS, McGowan, MR, Porto-Neto, LR, Hayes, BJ, Schutt, KM, Lyons, RE (2019) Poll testing efficiency, accuracy and trends in Australian cattle. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* **23**, 342-345.
- ✓ Randhawa, IAS, Porto-Neto, LR, McGowan, MR, Hayes, BJ, Schutt, KM, Lyons, RE (2019) Production and polledness: Genetic correlations between target traits in beef cattle. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* **23**, 350-353.
- ✓ Randhawa, IAS, Hayes, BJ, Porto-Neto, LR, Schutt, KM, McGowan, MR, Burns, BM, Lyons, RE (2019) 'New diagnostic tools helping poll breeding for sustainable beef production, Northern Beef Research Update Conference.' Brisbane, Australia, 19-22 August 2019.
- ✓ Randhawa, IAS, McGowan, MR, Porto-Neto, LR, Hayes, BJ, Lyons, RE (2019) 'Poll diagnostics, scur genetics and production concurrence in naturally hornless cattle, International Society for Animal Genetics (ISAG) Conference.' Lleida, Spain, 7-12 July 2019.
- ✓ Randhawa, IAS, Porto-Neto, LR, Hayes, BJ, McGowan, MR, Burns, BM, Lyons, RE (2019) 'Improving poll gene testing in Australian cattle, Farm animal welfare and gene technologies (RSPCA Australian Animal Welfare Seminar).' Canberra, Australia, 22 February 2019.

In addition, two more scientific articles are in preparation to be published in peer-reviewed journals.

8 Bibliography

Ajmone-Marsan, P, Garcia, JF, Lenstra, JA (2010) On the origin of cattle: How aurochs became cattle and colonized the world. *Evolutionary Anthropology* **19**, 148-157.

Allais-Bonnet, A, Grohs, C, Medugorac, I, Krebs, S, Djari, A, Graf, A, Fritz, S, Seichter, D, Baur, A, Russ, I, Bouet, S, Rothammer, S, Wahlberg, P, Esquerré, D, Hoze, C, Boussaha, M, Weiss, B, Thépot, D, Fouilloux, M-N, Rossignol, M-N, van Marle-Köster, E, Hreiðarsdóttir, GE, Barbey, S, Dozias, D, Cobo, E, Reversé, P, Catros, O, Marchand, J-L, Soulas, P, Roy, P, Marquant-Leguienne, B, Le Bourhis, D, Clément, L, Salas-Cortes, L, Venot, E, Pannetier, M, Phocas, F, Klopp, C, Rocha, D, Fouchet, M, Journaux, L, Bernard-Capel, C, Ponsart, C, Eggen, A, Blum, H, Gallard, Y, Boichard, D, Pailhoux, E, Capitan, A (2013) Novel insights into the bovine polled phenotype and horn ontogenesis in *Bovidae*. *PloS One* **8**, e63512.

Asai, M, Berryere, TG, Schmutz, SM (2004) The scurs locus in cattle maps to bovine chromosome 19. *Animal Genetics* **35**, 34-39.

Aulchenko, YS, Ripke, S, Isaacs, A, van Duijn, CM (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294-1296.

Bunter, KL, Johnston, DJ, Wolcott, ML, Fordyce, G (2013) Factors associated with calf mortality in tropically adapted beef breeds managed in extensive Australian production systems. *Animal Production Science* **54**, 25-36.

Capitan, A, Grohs, C, Gautier, M, Eggen, A (2009) The scurs inheritance: new insights from the French Charolais breed. *BMC Genetics* **10**, 33.

Capitan, A, Grohs, C, Weiss, B, Rossignol, M-N, Reversé, P, Eggen, A (2011) A Newly Described Bovine Type 2 Scurs Syndrome Segregates with a Frame-Shift Mutation in TWIST1. *PloS One* **6**, e22242.

Carlson, DF, Lancto, CA, Zang, B, Kim, E-S, Walton, M, Oldeschulte, D, Seabury, C, Sonstegard, TS, Fahrenkrug, SC (2016) Production of hornless dairy cattle from genome-edited cell lines. *Nature Biotechnology* **34**, 479.

Cohen, J (1977) 'Statistical power analysis for the behavioral sciences. .' (Routledge: New York)

Connors, NK, Tier, B, Johnston, DJ (2018) 'Current status of Australia's diagnostic poll haplotype test, World Congress on Genetics Applied to Livestock Production.'

Davis, EB, Brakora, KA, Lee, AH (2011) Evolution of ruminant headgear: a review. *Proceedings. Biological sciences* **278**, 2857-2865.

Dove, WF (1935) The physiology of horn growth: A study of the morphogenesis, the interaction of tissues, and the evolutionary processes of a mendelian recessive character by means of transplantation of tissues. *Journal of Experimental Zoology* **69**, 347-405.

Frisch, JE, Nishimura, H, Cousins, KJ, Turner, HG (1980) The inheritance and effect on production of polledness in four crossbred lines of beef cattle. *Animal Science* **31**, 119-126.

Gehrke, LJ, Capitan, A, Scheper, C, König, S, Upadhyay, M, Heidrich, K, Russ, I, Seichter, D, Tetens, J, Medugorac, I, Thaller, G (2020) Are scurs in heterozygous polled (Pp) cattle a complex quantitative trait? *Genetics Selection Evolution* **52**, 6.

Georges, M, Drinkwater, R, King, T, Mishra, A, Moore, SS, Nielsen, D, Sargeant, LS, Sorensen, A, Steele, MR, Zhao, X, Womack, JE, Hetzel, J (1993) Microsatellite mapping of a gene affecting horn development in Bos taurus. *Nature Genetics* **4**, 206-210.

Goonewardene, LA, Price, MA, Liu, MF, Berg, RT, Erichsen, CM (1999) A study of growth and carcass traits in dehorned and polled composite bulls. *Canadian Journal of Animal Science* **79**, 383-385.

Grobler, R, Visser, C, Capitan, A, van Marle-Köster, E (2018) Validation of the POLLED Celtic variant in South African Bonsmara and Drakensberger beef cattle breeds. *Livestock Science* **217**, 136-139.

Grossman, SR, Shylakhter, I, Karlsson, EK, Byrne, EH, Morales, S, Frieden, G, Hostetter, E, Angelino, E, Garber, M, Zuk, O, Lander, ES, Schaffner, SF, Sabeti, PC (2010) A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science* **327**, 883-886.

Harlizius, B, Tammen, I, Eichler, K, Eggen, A, Hetzel, D (1997) New markers on bovine Chromosome 1 are closely linked to the polled gene in Simmental and Pinzgauer cattle. *Mammalian Genome* **8**, 255-257.

Henderson, CR (1975) Use of Relationships Among Sires to Increase Accuracy of Sire Evaluation. *Journal of Dairy Science* **58**, 1731-1738.

Henshall, JM, Piper, E, McDonald, CA (2011) 'Effects of ascertainment bias when testing for polled in Australian beef cattle, Association for the Advancement of Animal Breeding and Genetics.' Perth, WA, Australia.

Herskin, MS, Nielsen, BH (2018) Welfare Effects of the Use of a Combination of Local Anesthesia and NSAID for Disbudding Analgesia in Dairy Calves-Reviewed Across Different Welfare Concerns. *Frontiers in Veterinary Science* **5**, 117.

Ketel, C, Asai-Coakwell, M (2020) Heterozygosity of the Celtic Polled locus in Canadian scurred beef cattle. *Canadian Journal of Animal Science*

Knierim, U, Irrgang, N, Roth, BA (2015) To be or not to be horned—Consequences in cattle. *Livestock Science* **179**, 29-37.

Kommisrud, E, Steine, T (1997) Semen Collection, Semen Production and Fertility Rates in Polled vs Horned Bulls in Norway. *Reproduction in Domestic Animals* **32**, 221-223.

Long, CR, Gregory, KE (1978) Inheritance of the horned, scurred, and polled condition in cattle. *Journal of Heredity* **69**, 395-400.

Lu, H, Giordano, F, Ning, Z (2016) Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics* **14**, 265-279.

Lundrigan, B (1996) Morphology of Horns and Fighting Behavior in the Family Bovidae. *Journal of Mammalogy* **77**, 462-475.

Mariasegaram, M, Harrison, BE, Bolton, JA, Tier, B, Henshall, JM, Barendse, W, Prayaga, KC (2012) Fine-mapping the *POLL* locus in Brahman cattle yields the diagnostic marker *CSAFG29*. *Animal Genetics* **43**, 683-688.

Mariasegaram, M, Reverter, A, Barris, W, Lehnert, SA, Dalrymple, B, Prayaga, K (2010) Transcription profiling provides insights into gene pathways involved in horn and scurs development in cattle. *BMC Genomics* **11**, 370.

Medugorac, I, Graf, A, Grohs, C, Rothammer, S, Zagdsuren, Y, Gladyr, E, Zinovieva, N, Barbieri, J, Seichter, D, Russ, I, Eggen, A, Hellenthal, G, Brem, G, Blum, H, Krebs, S, Capitan, A (2017) Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. *Nature Genetics* **49**, 470.

Medugorac, I, Seichter, D, Graf, A, Russ, I, Blum, H, Göpel, KH, Rothammer, S, Förster, M, Krebs, S (2012) Bovine polledness – an autosomal dominant trait with allelic heterogeneity. *PloS One* **7**, e39477.

Mueller, ML, Cole, JB, Sonstegard, TS, Van Eenennaam, AL (2019) Comparison of gene editing versus conventional breeding to introgress the *POLLED* allele into the US dairy cattle population. *Journal of Dairy Science* **102**, 4215-4226.

Neely, CD, Thomson, DU, Kerr, CA, Reinhardt, CD (2014) Effects of three dehorning techniques on behavior and wound healing in feedlot cattle. *Journal of Animal Science* **92**, 2225-9.

OMIA (2019) 'Online Mendelian Inheritance in Animals. Sydney School of Veterinary Science.' Available at http://omia.org/

Piper, EK, Tier, B, Henshall, JM (2014) A haplotype diagnostic for polled in Australian beef cattle. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production* **10**, 1-3.

Prayaga, K (2007) Genetic options to replace dehorning in beef cattle—a review. *Australian Journal of Agricultural Research* **58**, 1–8.

R Core Team (2019) 'R: A language and environment for statistical computing.' (R Foundation for Statistical Computing: Vienna, Austria)

Randhawa, IAS, Khatkar, MS, Thomson, PC, Raadsma, HW (2014) Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genetics* **15**, 34.

Randhawa, IAS, Khatkar, MS, Thomson, PC, Raadsma, HW (2016) A Meta-Assembly of Selection Signatures in Cattle. *PloS One* **11**, e0153013.

Randhawa, IAS, Lyons, RE, Hayes, BJ, Porto-Neto, LR, McGowan, MR (2019) Factors affecting development of horns and scurs in domesticated ruminants. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* **23**, 484-487.

Rizopoulos, D (2006) ltm: An R Package for Latent Variable Modeling and Item Response Analysis. *2006* **17**, 25.

Rothammer, S, Capitan, A, Mullaart, E, Seichter, D, Russ, I, Medugorac, I (2014) The 80-kb DNA duplication on BTA1 is the only remaining candidate mutation for the polled phenotype of Friesian origin. *Genetics Selection Evolution* **46**, 44.

Sabeti, PC, Varilly, P, Fry, B, Lohmueller, J, Hostetter, E, Cotsapas, C, Xie, X, Byrne, EH, McCarroll, SA, Gaudet, R, Schaffner, SF, Lander, ES, Consortium, TIH (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-918.

Schafberg, R, Swalve, HH (2015) The history of breeding for polled cattle. *Livestock Science* **179**, 54-70.

Scheper, C, Wensch-Dorendorf, M, Yin, T, Dressel, H, Swalve, H, König, S (2016) Evaluation of breeding strategies for polledness in dairy cattle using a newly developed simulation framework for quantitative and Mendelian traits. *Genetics Selection Evolution* **48**, 50.

Spurlock, DM, Stock, ML, Coetzee, JF (2014) The impact of 3 strategies for incorporating polled genetics into a dairy cattle breeding program on the overall herd genetic merit. *Journal of Dairy Science* **97**, 5265-5274.

Stafford, KJ, Mellor, DJ (2005) Dehorning and disbudding distress and its alleviation in calves. *The Veterinary Journal* **169**, 337-349.

Stookey, JM, Goonewardene, LA (1996) A comparison of production traits and welfare implications between horned and polled beef bulls. *Canadian Journal of Animal Science* **76**, 1-5.

Tetens, J, Wiedemar, N, Menoud, A, Thaller, G, Drögemüller, C (2015) Association mapping of the scurs locus in polled Simmental cattle – evidence for genetic heterogeneity. *Animal Genetics* **46**, 224-225.

Utsunomiya, YT, Torrecilha, RBP, Milanesi, M, Paulan, SdC, Utsunomiya, ATH, Garcia, JF (2019) Hornless Nellore cattle (Bos indicus) carrying a novel 110 kbp duplication variant of the polled locus. *Animal Genetics* **50**, 187-188.

Weir, BS, Cockerham, CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

White, WT, Ibsen, HL (1936) Horn inheritance in Galloway-Holstein cattle crosses. *Journal of Genetics* **32**, 33-49.

Wiedemar, N, Drögemüller, C (2015) A 1.8-kb insertion in the 3'-UTR of RXFP2 is associated with polledness in sheep. *Animal Genetics* **46**, 457-461.

Wiedemar, N, Tetens, J, Jagannathan, V, Menoud, A, Neuenschwander, S, Bruggmann, R, Thaller, G, Drögemüller, C (2014) Independent polled mutations leading to complex gene expression differences in cattle. *PloS One* **9**, e93435.

Wiener, DJ, Wiedemar, N, Welle, MM, Drögemüller, C (2015) Novel Features of the Prenatal Horn Bud Development in Cattle (Bos taurus). *PloS One* **10**, e0127691.

Windig, JJ, Hoving-Bolink, RA, Veerkamp, RF (2015) Breeding for polledness in Holstein cattle. *Livestock Science* **179**, 96-101.

Wu, XL, Xu, J, Li, H, Ferretti, R, He, J, Qiu, J, Xiao, Q, Simpson, B, Michell, T, Kachman, SD, Tait, RG, Jr., Bauck, S (2019) Evaluation of genotyping concordance for commercial bovine SNP arrays using quality-assurance samples. *Animal Genetics* **50**, 367-371.

Zeder, MA (2008) Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 11597-11604.

9 Appendices

9.1 Breed-wise samples tested with SPTv1

Appendix 9.1. Comparison of SNP-based current poll testing (SPTv1) and the optimized poll testing (OPT) assays on 20,636 samples for the frequency of "No results" in various cattle breeds.

Breed	Number tested	No results with SPTv1	No results with OPT
Brahman	2,819	492	0
Wagyu [A]	9,050	60	0
Droughtmaster	558	9	0
Shorthorn	67	5	1
Santa Gertrudis	136	5	0
Brangus	72	6	0
Hereford	3341	1	0
Angus	1,602	0	0
Charolais	900	0	0
Limousin	207	0	0
Holstein-Friesian	*31*	*0*	*0*
Braford	*14*	*0*	*0*
Murray Grey	*8*	*0*	*0*
Boran	*7*	*0*	*0*
Poll Hereford	*7*	*0*	*0*
Red Angus	*4*	*0*	*0*
Red Wagyu	*2*	*0*	*0*
Simmental	*2*	*0*	*0*
Gelbvieh	*1*	*0*	*0*
Composite	*35*	*1*	*0*
Angus x Brahman	*33*	*7*	*0*
Angus x Simmental	*2*	*0*	*0*
Wagyu Cross	*2*	*0*	*0*
Ultrablack x WAGX	*2*	*0*	*0*
Cross Breed	*2*	*0*	*0*
Brahman Cross	*1*	*0*	*0*
Unknown breeds [B]	*1,731*	*1,413*	*8*
Total	**20,636**	**1,999**	**9**

[A] Wagyu includes F1-F4 classes as well as Fullblood population in Australia.

[B] Data provided to the testing laboratory does not always identify the breed of the sample. These samples are primarily acquired for the "No Results" output of SPTv1 and suspected to be of Brahman and Brahman-infused cross-bred cattle.

9.2 Genomic positions of microsatellites

Appendix 9.2. List of 14 microsatellites mapped within the Poll locus on chromosome 1.

Microsatellites	ARS-UCD2.1[A]	UMD3.1 [A]	Btau4.0
CSAFG26	2341080	1617119	1441734
CSAFG27	2368949	1646009	1470624
CSAFG28	2394182	1670889	1495504
CSAFG29	2406659	1683376	1507991
RP42-351B8_MS1	2410219	1686934	1511549
RP42-351B8_MS2	2432556	1709169	1533784
CSAFG30	2435911	1712696	1537311
CSAFG31	2457207	1733999	1558614
CSAFG33	2512705	1790169	1614784
CSAFG34	2575232	1852525	1677140
CSAFG35	2588208	1865512	1690127
CSAFG37	2704710	1984932	1809547
CSAFG38	2735847	2016159	1840774
CSAFG22	3014463	2294700	2119315

[A] Positions on ARS-UCD1.2 and UMD3.1 are approximated based on the nearest rs-SNP to each MSAT.

9.3 Call rate of ten SNPs in different breeds

Appendix 9.3. Call rate of 10 SNPs in different cattle breeds and cross-bred populations.

Breeds	Number tested	rs801127025	rs799187101	rs800947704	rs798116945	rs383143898	rs799403053	rs210350155	rs797088784	rs800767839	rs799920960
Angus	1594	99.89	100	96.96	100	99.96	99.94	99.65	99.50	100	-
Angus X Brahman	33	100	100	90.91	-	100	100	100	100	-	100
Boran	7	100	100	28.57	100	100	100	100	100	100	-
Braford	14	100	100	100	100	100	100	100	100	100	-
Brahman	2604	98.27	99.92	85.71	100	99.97	99.73	98.78	99.77	99.60	100
Brangus	9	100	100	100	100	100	100	100	100	100	-
Charolais	370	99.19	99.73	98.92	100	100	99.80	99.19	98.65	100	-
Composite	35	97.14	100	95.71	100	99.05	100	98.57	100	97.14	-
Droughtmaster	289	96.93	99.83	89.62	100	99.88	98.70	95.85	99.65	100	-
Hereford	2820	99.88	99.98	99.66	99.96	99.81	99.95	99.59	98.09	100	-
Holstein-Friesian	31	100	100	100	100	100	100	100	100	100	-
Limousin	24	100	100	91.67	100	100	100	100	100	100	-
Murray Grey	8	100	100	87.50	100	100	100	100	100	100	-
Poll Hereford	7	100	100	100	100	100	100	100	100	100	-
Red Angus	4	100	100	75.00	100	100	100	100	100	100	-
Santa Gertrudis	81	97.88	100	93.83	100	100	99.07	96.91	100	100	-
Shorthorn	67	100	100	100	100	100	100	93.28	64.18	100	-
Unknown	1631	100	100	98.41	100	99.52	100	100	100	100	100
Wagyu	9047	99.45	100	98.03	99.99	99.99	99.79	99.27	99.93	99.99	-

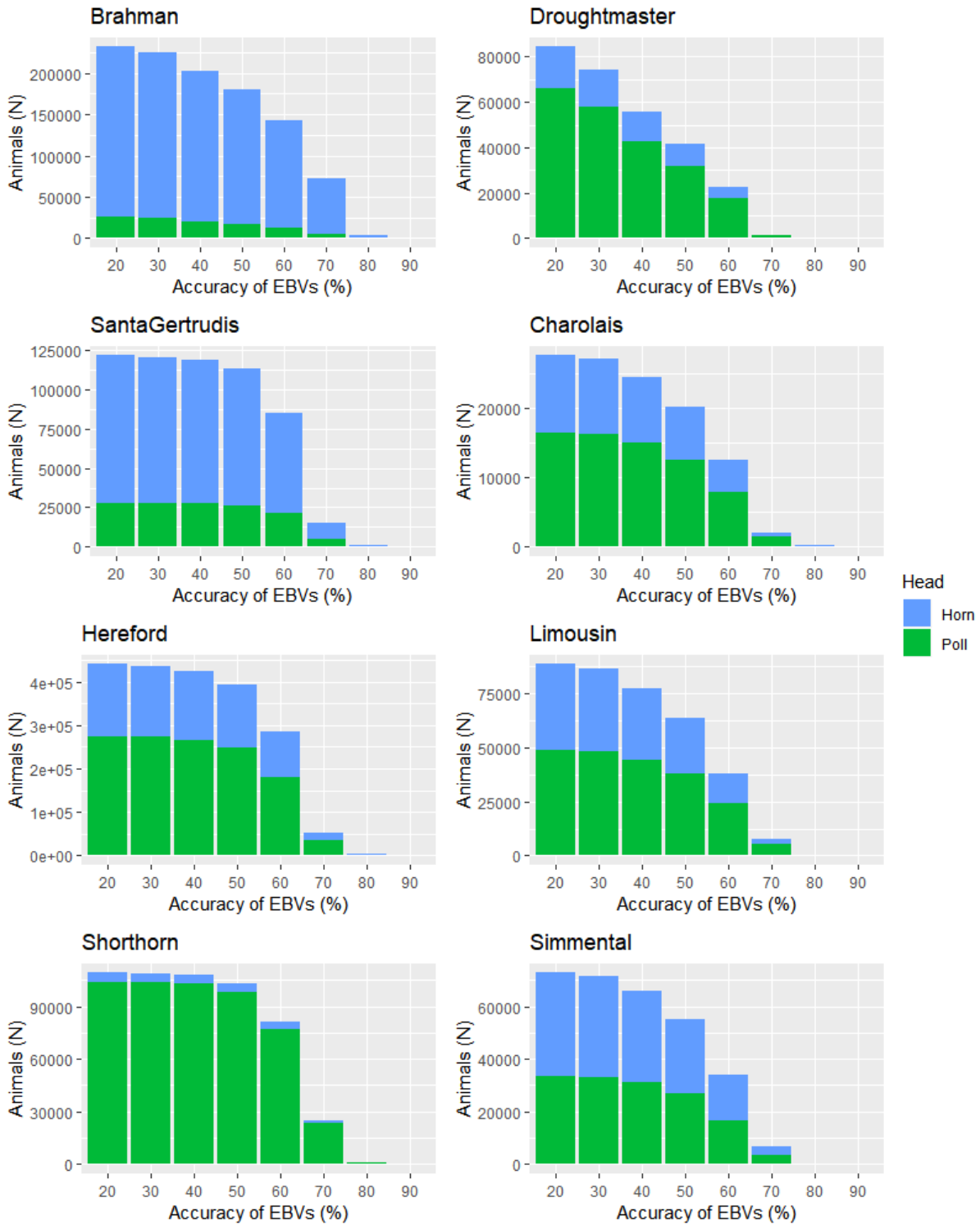
9.4 Phenotypic concordance with OPT genotypes in different breeds

Appendix 9.4. Phenotypic concordance between OPT genotypes and BREEDPLAN phenotypes.

Breeds	Head-status	OPT predicted genotypes							
		HH	HPc	HPf	NR	PcPc	PcPf	PfPf	Total
Angus		**1**	**166**			**1,426**	**12**		**1,605**
	Poll		166			1,426	12		1,604
	Unknown	1							1
Boran		**7**							**7**
	Horn	7							7
Braford		**1**	**8**	**1**		**4**			**14**
	Unknown	1	8	1		4			14
Brahman		**2,701**	**850**	**1**		**84**	**2**		**3,638**
	Horn	2,613	231			9			2,853
	Poll	17	431			74	2		524
	Scur	65	174	1					240
	Unknown	6	14			1			21
Brangus		**3**	**5**			**15**			**23**
	Horn	3							3
	Poll		2			15			17
	Scur		3						3
Charolais		**97**	**288**	**19**		**217**	**25**	**1**	**647**
	Horn	57	5			1			63
	Poll	7	212	15		194	17	1	446
	Scur	4	5						9
	Unknown	29	66	4		22	8		129
Composite		**48**	**26**			**1**			**75**
	Unknown	48	26			1			75
Crossbred		**27**	**44**			**2**			**73**
	Horn	24							24
	Poll		7						7
	Scur		5						5
	Unknown	3	32			2			37
Droughtmaster		**128**	**343**	**9**		**153**	**5**		**638**
	Horn	104	28			3			135
	Poll	10	151	9		146	5		321
	Scur	8	140			4			152
	Unknown	6	24						30

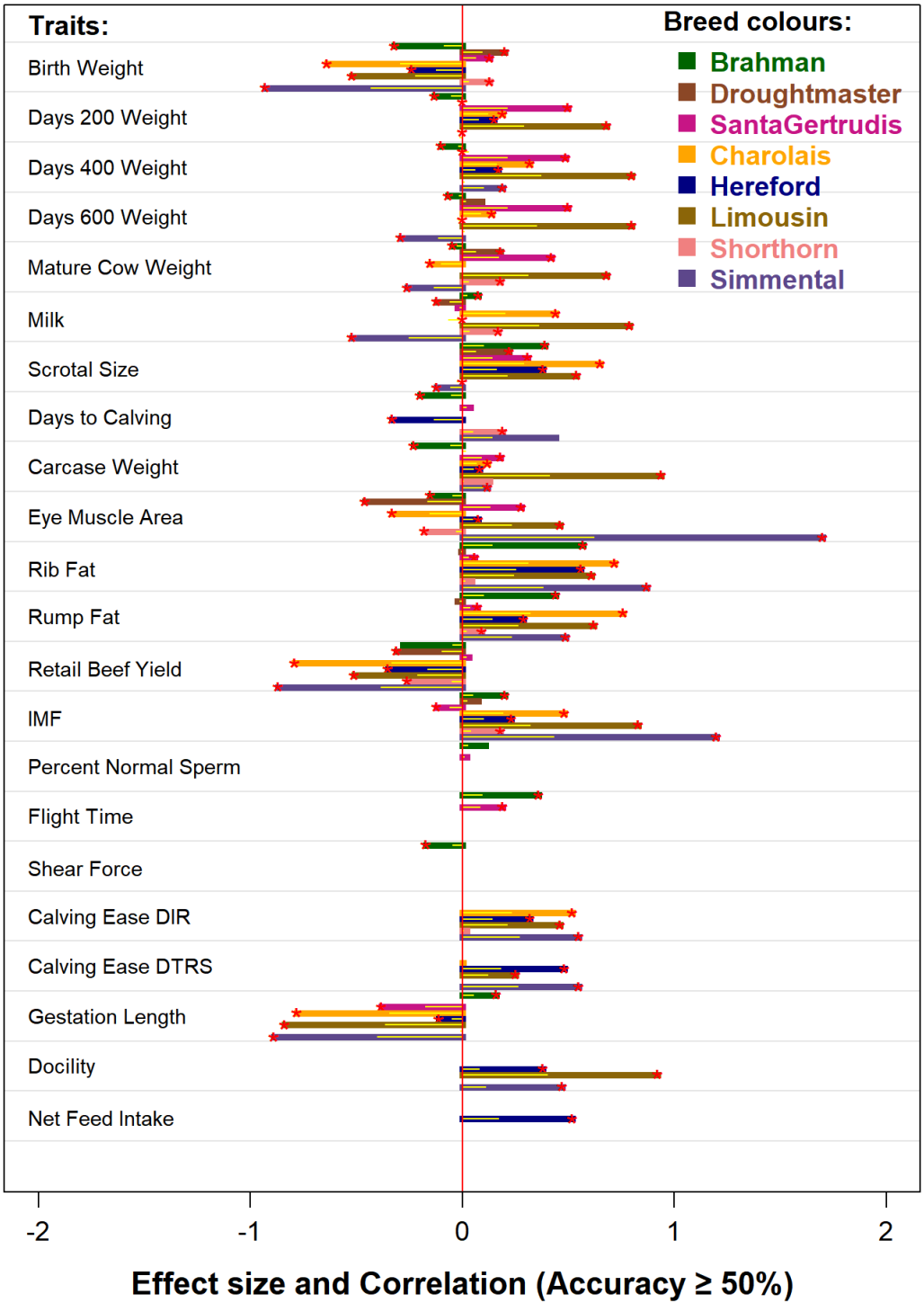
Breeds	Head-status	OPT predicted genotypes							
		HH	HPc	HPf	NR	PcPc	PcPf	PfPf	Total
Gelbvieh						**1**			**1**
	Poll					1			1
Hereford		**448**	**1,025**	**91**		**1,596**	**252**	**13**	**3,425**
	Horn	417	12	1		2	1		433
	Poll	23	756	70		1591	250	13	2,703
	Scur	7	239	19			1		266
	Unknown	1	18	1		3			23
Holstein-Friesian		**31**							**31**
	Horn	31							31
Limousin		**3**	**6**			**15**			**24**
	Horn	3							3
	Poll		5			15			20
	Scur		1						1
Murray Grey						**8**			**8**
	Poll					8			8
Santa Gertrudis		**41**	**67**	**7**		**10**	**1**	**1**	**127**
	Horn	32	9	1		1		1	44
	Poll	2	41	5		9	1		58
	Scur		12						12
	Unknown	7	5	1					13
Shorthorn		**1**	**11**	**15**	**1**	**3**	**24**	**13**	**68**
	Poll	1	10	15	1	3	24	13	67
	Unknown		1						1
Simmental						**2**			**2**
	Unknown					2			2
Unknown		**706**	**174**	**12**	**8**	**30**	**6**	**7**	**943**
	Unknown	706	174	12	8	30	6	7	943
Wagyu		**8,237**	**692**	**6**		**49**			**8,984**
	Horn	8,237	1						8,238
	Poll		71			49			120
	Unknown		620	6					626
Total		**12,480**	**3,705**	**161**	**9**	**3,616**	**327**	**35**	**20,333**

9.5 Sample sizes and accuracy of EBVs



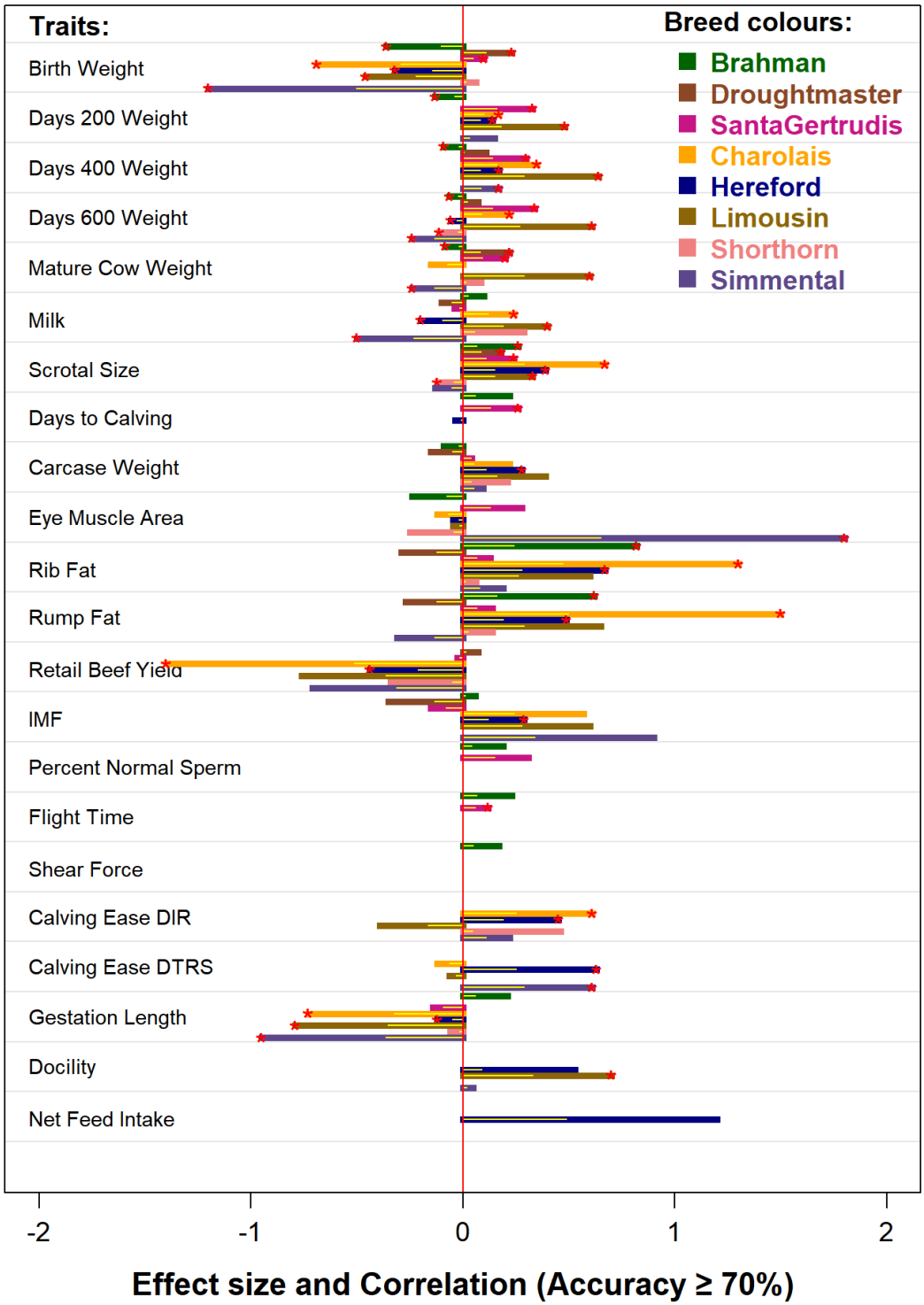
Appendix 9.5. Sample sizes (number of animals) with recorded EBVs above different levels of accuracy (%) of EBVs. Note that it includes any animal with one or more EBVs and the charts represent maximum animals in each breed at each threshold. Some traits were recorded on limited animals and hence have sample sizes much smaller even at lower accuracies.

9.6 Effect size and correlation of EBVs at 50% threshold



Appendix 9.6. Effect sizes and correlations of polled cohorts for 8 breeds using BREEDPLAN EBVs (accuracy ≥ 50%). Effect sizes (Cohen's *d*) are shown as breed-coloured bars and correlations (Point Biserial r) are shown with yellow lines within each bar. Statistically significant results are shown with a red-star.

9.7 Effect size and correlation of EBVs at 70% threshold



Appendix 9.7. Effect sizes and correlations of polled cohorts of 8 breeds using BREEDPLAN EBVs (accuracy ≥ 70%). Effect sizes (Cohen's *d*) are shown as breed-coloured bars and correlations (Point Biserial r) are shown with yellow lines within each bar. Statistically significant results are shown with a red-star.

9.8 Descriptive statistics of EBVs concordant with OPT genotypes

Appendix 9.8-A. Descriptive statistics between the poll and horn cohorts of Brahman and Hereford using animals (n=5,909) found concordant for OPT-genotypes and BREEDPLAN phenotypes.

Traits (EBV units) [A]	**Brahman**				**Hereford**			
	d	*PBC*	*t-test*	*p-value*	*d*	*PBC*	*t-test*	*p-value*
Birth Weight (kg)	0.053	0.009	0.51	0.609	-0.29*	-0.110	-4.60	4.51E-06
Days 200 Weight (kg)	0.00	0.033	1.70	0.087	0.42*	0.150	6.10	2.67E-09
Days 400 Weight (kg)	0.19	0.049	2.50	0.011	0.36*	0.140	5.30	1.53E-07
Days 600 Weight (kg)	0.13*	0.071	3.60	0.0004	0.20	0.074	2.80	0.00512
Mature Cow Weight (kg)	0.22*	0.084	4.00	5.8E-05	0.07	0.031	1.30	0.199
Milk (kg)	0.04	0.017	0.83	0.407	0.00	0.024	1.00	0.305
Scrotal Size (cm)	0.05	0.015	0.90	0.368	0.63*	0.196	11.00	7.09E-24
Days to Calving (days)	-0.10	-0.033	-1.80	0.069	-0.49*	-0.180	-8.40	7.86E-16
Carcase Weight (kg)	0.19	0.086	3.70	0.00023	0.45*	0.170	6.80	4.17E-11
Eye Muscle Area (sqcm)	0.22*	0.085	4.40	1.1E-05	0.00	-0.012	-0.55	0.582
Rib Fat (mm)	0.03	0.014	0.69	0.491	0.63*	0.221	10.00	1.21E-22
Rump Fat (mm)	0.11	0.043	2.20	0.025	0.52*	0.175	8.60	1.06E-16
Retail Beef Yield (%)	0.21*	0.079	3.70	0.00025	-0.49*	-0.195	-8.40	6.53E-16
IMF (%)	-0.05	-0.019	-1.10	0.281	0.52*	0.168	9.20	1.55E-18
Percent Normal Sperm (%)	-0.21*	-0.067	-3.60	0.0004				
Flight Time (secs)	0.10	0.039	2.10	0.037				
Shear Force (kg)	-0.05	-0.019	-1.00	0.318				
Calving Ease DIR (%)					0.45*	0.152	7.20	2.32E-12
Calving Ease DTRS (%)					0.78*	0.242	12.00	6.02E-27
Gestation Length (days)	-0.25*	-0.11	-4	9.0E-05	-0.22*	-0.075	-3.80	0.000178
Docility					0.59*	0.162	4.50	2.80E-05
NFIF (kg/day)					1.10*	0.317	11.00	6.42E-20

d = Cohen's d (effect size) ***PBC*** = Point Biserial Correlation (***r***)

* Significant at $p < 0.05$ (Bonferroni corrected)

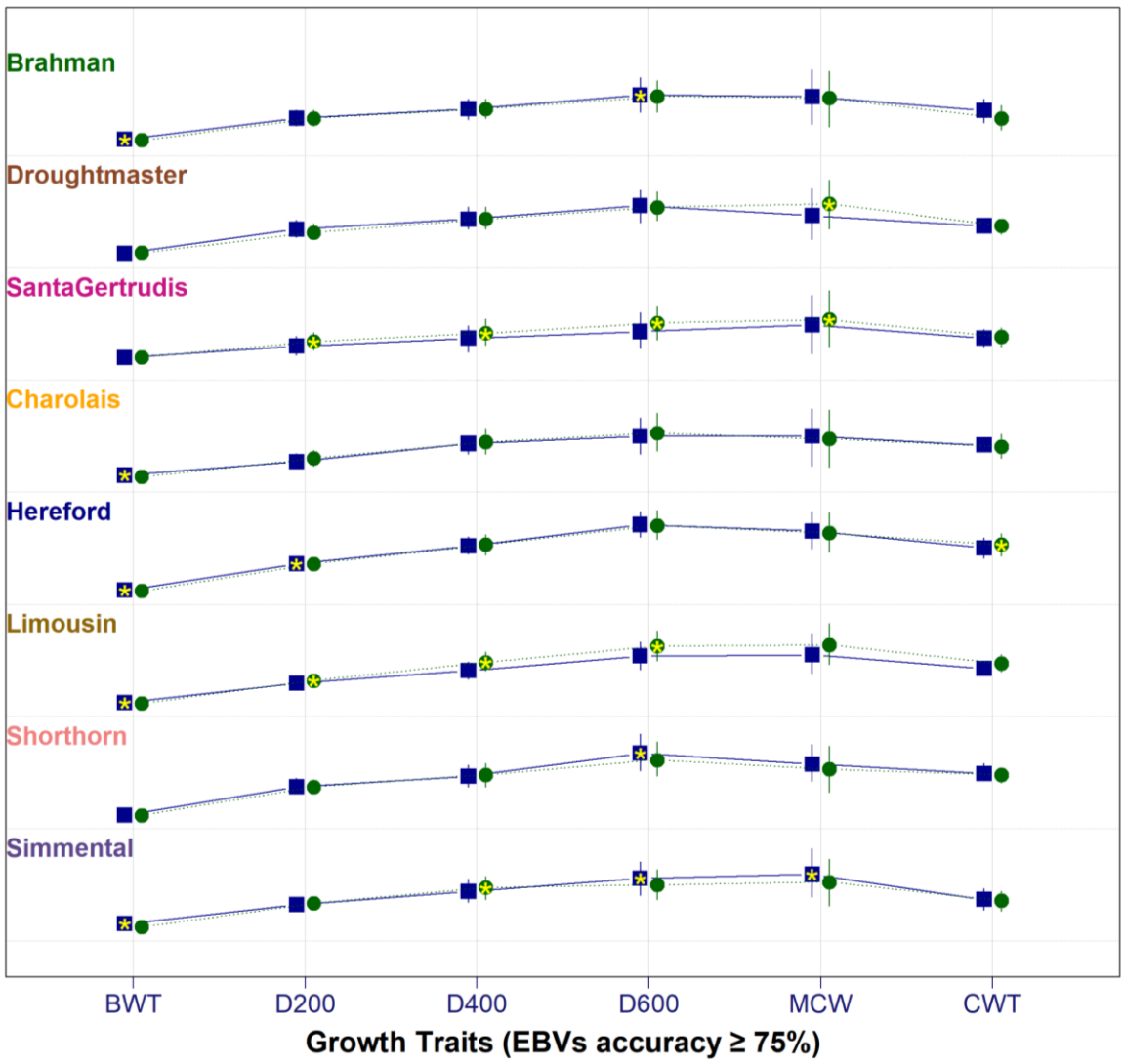
[A] Sample sizes and summary statistics are given in Table A28.

Appendix 9.8-B. Number of samples and summary statistics of poll and horn cohorts of Brahman and Hereford breeds. These animals have been found concordant for OPT-genotype and BREEDPLAN phenotype.

Traits (EBV units)	Brahman				Hereford			
	Poll cohort		*Horn cohort*		*Poll cohort*		*Horn cohort*	
	N [A]	Mean±SD	N	Mean±SD	N	Mean±SD	N	Mean±SD
Birth Weight (kg)	502	2.50±1.9	2271	2.60±1.9	2039	4.40±2.2	333	5.10±2.6
Days 200 Weight (kg)	503	19.00±7.5	2275	19.00±7	2039	34.00±8	333	30.00±11
Days 400 Weight (kg)	504	28.00±11	2288	26.00±10	2039	56.00±14	333	50.00±19
Days 600 Weight (kg)	505	39.00±16	2309	37.00±14	2039	78.00±20	333	73.00±29
Mature Cow Weight (kg)	502	45.00±25	2278	40.00±21	2037	68.00±25	333	66.00±31
Milk (kg)	475	-1.50±2.6	2177	-1.60±2.2	2029	15.00±5	318	15.00±6
Scrotal Size (cm)	489	0.90±1	2180	0.80±1.2	2033	2.10±1	333	1.50±0.9
Days to Calving (days)	453	-3.20±6.4	2120	-2.20±7	1969	-2.60±1.8	293	-1.70±1.9
Carcase Weight (kg)	502	24.00±12	2268	22.00±8.3	2035	52.00±15	333	44.00±20
Eye Muscle Area (sqcm)	473	2.80±1.9	2159	2.40±1.8	2023	3.60±1.8	306	3.60±2.0
Rib Fat (mm)	471	-0.22±0.95	2154	-0.28±0.88	2029	0.64±0.9	331	0.05±1.0
Rump Fat (mm)	493	-0.31±1.3	2215	-0.49±1.2	2031	0.80±1.4	332	0.08±1.4
Retail Beef Yield (%)	385	0.63±0.49	1980	0.53±0.44	2024	0.71±0.88	312	1.20±1.1
IMF (%)	432	-0.08±0.19	2087	-0.07±0.22	2028	0.60±0.78	329	0.22±0.7
Percent Normal Sperm (%)	232	1.40±2.8	1230	2.10±4.5				
Flight Time (secs)	471	0.02±0.1	2150	0.002±0.1				
Shear Force (kg)	448	-0.0008 ±0.26	2119	0.028±0.26				
Calving Ease DIR (%)					2019	1.30±5.9	302	-1.40±6.1
Calving Ease DTRS (%)					1973	1.60±4.0	273	-1.50±4.0
Gestation Length (days)	355	-0.39±1.6	1392	0.02±1.1	2035	-0.73±2.2	329	-0.24±2.2
Docility					597	3.4±10.0	59	-2.30±9.3
NFIF (kg/day)					829	-0.003±0.2	102	-0.22±0.2

[A] Number of animals which have EBVs for a specific trait within each cohort based on the head-status.

9.9 Trends of weight traits' EBVs at 75% accuracy



Appendix 9.9. Comparison of BREEDPLAN EBVs (accuracy ≥ 75%) of growth traits at different life-stages in eight breeds. Trait abbreviations are **BWT**: Birth weight; **D200**: Days 200 weight; **D400**: Days 400 weight; **D600**: Days 600 weight; **MCW**: Mature cow weight; **CWT**: Carcase weight.

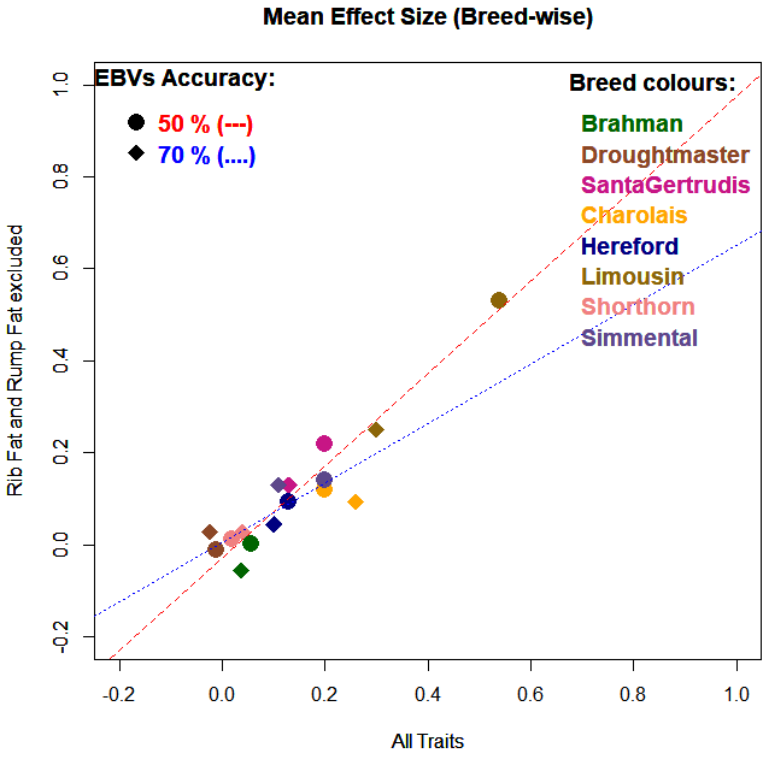
9.10 Breed-wise standardized effect sizes and correlations with polledness

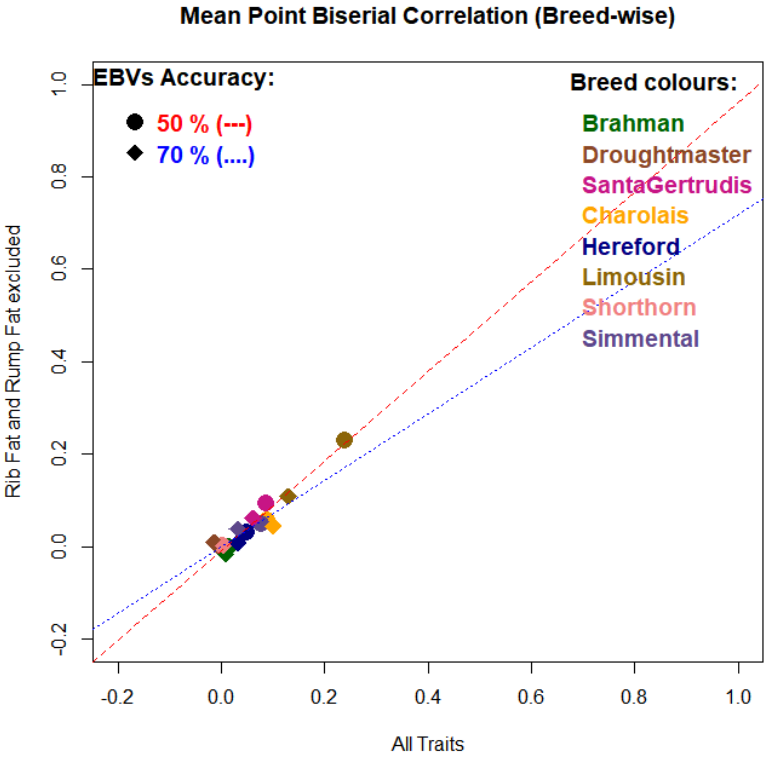
Appendix 9.10-A. Standardized Mean Effect size and Point Biserial Correlation of polledness on other traits by excluding Rib and Rump fat traits.

Breeds	EBVs accuracy ≥ 50%			EBVs accuracy ≥ 70%		
	Traits (N)[A]	**Effect Size** (Mean ±SD)	**PBC** (Mean ±SD)	**Traits** (N) [A]	**Effect Size** (Mean ±SD)	**PBC** (Mean ±SD)
Brahman	16	0.001±0.22	6.3e-5±0.06	15	-0.055±0.18	-0.016±0.05
Droughtmaster	11	-0.011±0.21	0.001±0.08	10	0.028±0.18	0.009±0.07
Santa Gertrudis	15	0.220±0.21	0.094±0.09	15	0.130±0.18	0.062±0.09
Charolais	14	0.120±0.47	0.058±0.21	14	0.093±0.58	0.045±0.23
Hereford	17	0.094±0.27	0.032±0.11	17	0.045±0.43	0.009±0.17
Limousin	15	0.530±0.47	0.230±0.20	15	0.250±0.48	0.110±0.22
Shorthorn	15	0.012±0.14	0.002±0.03	12	0.029±0.22	0.003±0.03
Simmental	16	0.140±0.73	0.049±0.30	15	0.130±0.73	0.039±0.29

PBC: Point Biserial Correlation

[A] Number of traits with recorded EBVs (above respective accuracy threshold) in each breeds.





Appendix 9.10-B. Comparison of overall ES and PBC of polledness (breed-wise) by including and excluding two traits (Rib and Rump fat).