

Current status of Australia's diagnostic poll haplotype test

N.K. Connors¹, B. Tier¹ & D.J. Johnston¹

¹Animal Genetics and Breeding Unit, University of New England, New South Wales, Australia*
natalie.connors@une.edu.au (Corresponding Author)

Summary

Australia's diagnostic poll haplotype test requires reliable and varied phenotype submissions to estimate polled probabilities of haplotypes, and is dependent on unbiased sampling of the population. This paper provides a review of the effectiveness of the haplotype poll test and shows clear potential for significant ascertainment bias to be affecting the accuracy of the test, resulting from industry submission of mostly unknown or polled phenotypes, with little control over phenotype scoring accuracy. A new project targeting the supplementation with horned animals is underway to address the resulting phenotype proportions, with the aim of greatly increasing the accuracy of the test.

Keywords: poll, horn, haplotype, beef.

Introduction

Currently, bruising injury from horns is estimated to cost the Australian meat industry \$30 million per year (CSIRO, 2014). Removing the horns, 'dehorning', is now commonly accepted management practice in cattle husbandry (Medugorac *et al.*, 2012). Dehorning requires the removal of the horn 'bud' from the animal at an early age (less than 6 months) and requires special tools/knives and skilled farm-hands to ensure a successful removal.

Although commercially necessary and common practice, dehorning is a painful procedure regardless of the method used, and as such is likely to be subject to renewed animal welfare legislation in the future (Capitan *et al.*, 2009). Mustering practices within Northern Australia often lead to calves being dehorned at up to 10 months of age (Bortolussi *et al.*, 2005); late dehorning can lead to larger wounds, longer wound healing time, and secondary infections, translating into short term weight loss and increased mortality rates (Winks *et al.*, 1977). Studies into calf mortality within Australian production systems found that the incidence of calf deaths after 3 months of age was associated with the process of dehorning (Bunter *et al.*, 2013). While dehorning can provide an economic improvement at the point of slaughter, these improvements are likely negated with short-term loss from weight loss and mortality, along with increased labour costs for the process. Dehorning requires repeating at each generation and as such is generally perceived to be "treating the symptom, and not the cause" (Capitan, *et al.*, 2009).

The long-term alternative to dehorning of horned cattle is to breed naturally hornless cattle, called polled. Cattle horns form as a free-floating bud, usually evident at birth, which later fuses to the skull to form as a fixed bony extension. However, horn development and morphology demonstrates significant polymorphism within the species (Medugorac, *et al.*, 2012). Scurs appear similar to horns, though often smaller and only loosely attached, they are often not fixed to the skull and are able to move (Asai *et al.*, 2004), while polled cattle are naturally hornless (Seichter *et al.*, 2012).

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England

Poll genetics

At least three genes, poll, scur, and African horn, have been associated with the presence/absence of horns, with phenotypes dependent on epistatic interactions, allele dominance, and sex-influenced expression. Polledness is determined by an autosomal dominant locus characterized by two alleles, P (polled) and p (horned) (White & Ibsen, 2008) and has been mapped to bovine chromosome one (Georges *et al.*, 1993). The scurs locus is also characterised by two alleles, Sc (scurs) and sc (no scurs), and is sex-linked; males only require one scur allele to form scurs but females require two scur alleles to form scurs (White & Ibsen, 2008). The African horn gene is theorised to have a sex-linked epistatic effect on the polled locus, such that males require only one allele to be horned (Georges, *et al.*, 1993), but this theory remains unproven. Because horns grow in the same position as scurs, a horned genotype (pp) would mask the presence of scurs (Asai, *et al.*, 2004). Furthermore, it is suggested that homozygosity at the polled locus masks growth of scurs, unless the animals is also homozygous for scurs (Long & Gregory, 1978).

Diagnostic analyses for polledness have been developed by a number of parties, targeting various locations of the poll locus. An insertion-deletion (InDel) event was identified as a possible causal variant in European *Bos Taurus* cattle; a 202 base pair (bp) sequence is duplicated and replaces a sequence of 10 bp (Medugorac, *et al.*, 2012). As all animals with at least one InDel allele were polled and originate from geographically Celtic areas, this mutation has been designated as Celtic Polled, P_C. Furthermore, a 260 bp haplotype block consisting of five mutations, three SNPs flanked by two InDels, was identified in Friesian cattle breeds, known as Friesian Polled, P_F (Medugorac, *et al.*, 2012). A commercial SNP test targeting the P_C InDel was developed specifically for *Bos Taurus* cattle, and patented (Ingenuity/GeneSeek), but was not proven for *Bos Indicus*.

Ten microsatellite markers have been mapped within the polled locus, between 1,495,504 bp and 2,119,315 (Mariasegaram *et al.*, 2012). The Beef Cooperative Research Centre (Beef CRC) in Australia released a single marker DNA test for polled which could be applied to *Bos Indicus* cattle in 2010, pre-dating the P_C SNP test. The test was based on a 303 bp allele at the microsatellite marker CSAFG29, which was strongly associated with the polled phenotype in Brahman cattle (Mariasegaram, *et al.*, 2012; Prayaga *et al.*, 2009), and formed the basis for Australia's commercial polled test from 2010-2013 (Piper *et al.*, 2014).

Pre-commercialisation testing of the single marker CSAFG29 poll test on various breeds in Australia demonstrated its limitations to assign poll genotype accurately in some breeds. Limousin and Angus breeds obtained genotype assignment accuracies as low as 39% (Henshall *et al.*, 2014). Another allele within marker CSAFG29, 305 bp in length, was discovered to be associated with both polled and horned (Henshall *et al.* 2011); one that was associated as horned at the poll locus (prevalent in French Limousin) and one that was associated as polled at the poll locus (prevalent in Angus). The test was modified to include the two sources of the 305 bp allele at marker CSAFG29, in addition to expansion to ten microsatellite markers forming haplotypes. This newly improved haplotype test was shown to have improved accuracy for the Limousin/Angus breeds (Henshall, *et al.*, 2014), and was released commercially in 2014 as Australia's diagnostic polled haplotype test (Henshall *et al.*, 2011; Henshall, *et al.*, 2014; Piper, *et al.*, 2014).

Australia's diagnostic haplotype poll test

Test data structure during development

The structure of the test reference population has changed over time. Development of the poll test began with single marker genotypes and phenotypes from unrelated industry and research animals, consisting of 68 Brahmans and 20 Herefords (Prayaga, *et al.*, 2009), in addition to phenotypes and genotypes of 91 Angus and 52 Limousin available from previous research (Mariasegaram, *et al.*, 2012). An additional 142 industry-provided Limousin cattle were added to the analysis in 2011, of which all were polled but one scurred animal (Henshall, *et al.*, 2011). Once development of the test moved toward a ten marker haplotype basis in 2014, a reference population of 1,759 animals of various breeds and phenotypes had been established (Piper, *et al.*, 2014). At that time, phenotypes came from approximately 20 various breeds of both *Bos Indicus* and *Bos Taurus* cattle, and consisted of approximately 29% horned, 5% scurred, 12% unknown, and 53% polled. Some phenotypes were progeny-tested (pt), whereby selective matings showed consistent progeny phenotypes; for example, horned progeny of polled parents suggest that parents are not homozygous PP and likely to be heterozygous Pp. Of those polled phenotypes in the reference population, 4% were progeny tested polled, ptPP and ptPp, and or the horned phenotypes <1% were progeny tested horned, ptpP.

This initial population generated 448 haplotypes, of which 200 haplotypes were only seen once, and the first 100 common haplotypes accounted for 80% of observations (Henshall, *et al.*, 2014). Data validation of submitted phenotypes and genotypes was finalised in 2014, at which point data submission to the test was provided entirely by industry. The test has been offered through The University of Queensland Animal Genetics Laboratory (UQAGL), where genotypes are analysed at Animal Genetics and Breeding Unit (AGBU), and results sent back to UQAGL.

Test methods and improvements

Since 2014 animal samples submitted for testing may include a phenotype scored by the breeder (i.e. Horned, Polled, Scurred) which varied in its reliability and accuracy. No progeny tested phenotypes (e.g. ptPP, ptPp, ptpP) have been submitted in this time, and if no phenotype is provided, it is scored as 'Unknown'. The test uses diploid genotypes from ten microsatellite markers to estimate haplotype pairs for each sample, using the haplo.em function of the haplo.stats R package (Sinnwell *et al.*, 2007). Haplotypes are assigned as either horned or polled based on the phenotypes provided with the associated sample, using an MCMC sampler which applies the Metropolis-Hastings algorithm (Hastings, 1970). Assumptions of phenotype based on genotype are handled by a penetrance function to weight the estimations appropriately given there are possibilities of phenotyping, genotyping, and mislabelling errors (Piper, *et al.*, 2014). Each haplotype is provided a polled probability based on the following criteria of specific phenotype and genotype observations: (i) observed in polled animals with homozygous haplotypes; (ii) observed within progeny tested animals (ptPP, ptPp, ptpP); (iii) observed in horned animals; or (iv) observed in polled or scurred animals, where the other haplotype is horned; if the haplotypes are not observed in any of these situations, then they cannot be assigned as horned or polled (Henshall, *et al.*, 2014; Piper, *et al.*, 2014). Animals with genotype estimations lower than 90% probability do not

receive a result.

In 2016 improvements were made to clarify some uncertainty in the polled probabilities of haplotypes, and determine probable genotypes for animals that were previously omitted. Briefly, haplotype estimation was performed 100 times, instead of once, enabling variable haplotype estimations to be captured within each batch run (Connors & Tier, 2016). Animals that received haplotype probabilities less than 0.99 were re-estimated; varied and/or multiple haplotype probabilities were accumulated over the 100 chains and the mean polled probabilities reported, enabling previously omitted animals to receive probable genotypes (Connors *et al.*, 2016). The MCMC sampler run-time was shortened and repeated 100 times in parallel, enabling approximately 80% reduction in computation time.

Current status and data structure

At the time of writing, the poll haplotype test has been applied to more than 25,000 animals. This total includes more than 50 different breeds, though over 75% of all animals are from only six breeds: Brahman, Santa Gertrudis, Hereford, Charolais, Droughtmaster, and Limousin (as shown in Figure 1), while over 40 breeds each consist of less than 1% of the total (grouped as ‘Other’ in Figure 1).

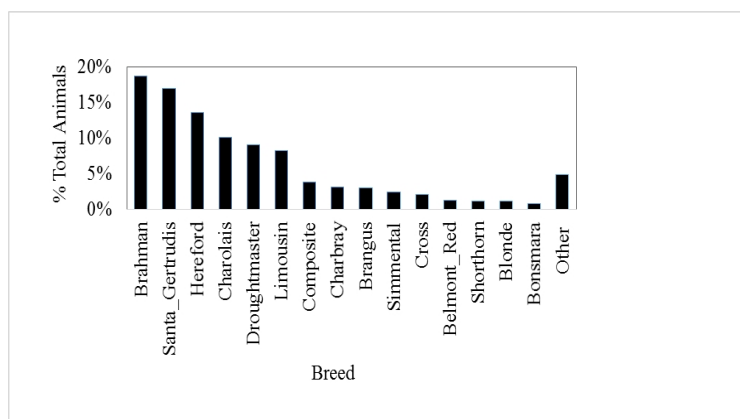


Figure 1. Proportion of breeds submitted to poll haplotype test.

Haplotype numbers have increased significantly since the release to industry, with more than 2400 haplotypes now observed. This is likely due to a number of reasons; (i) increasing haplotype estimations to 100 times per batch and inclusion of rare haplotypes, (ii) increase in the number of breeds and crosses sampled within the test, (iii) increased sampling of varied sub-populations within breeds, (iv) possible recombination within the marker locations over time. Despite the increase in haplotype variability, the first 250 common haplotypes account for approximately 90% of assignments within the test; the remaining 2150 haplotypes, while only accounting for 15% of haplotype assignments, represents a level of uncertainty in the test, providing a much more accurate assignment of poll probability.

This uncertainty is better represented in Figure 3, showing the number of times a haplotype is seen, and its polled probability. As expected, the more a haplotype is observed, the more certain its polled probability becomes, and it is assigned close to either one (polled) or zero (horned). If observed only a few times, that uncertainty of its polled probability is a reflection of both the number of observations and its association with varied or unknown phenotypes. Initial investigations show that haplotypes that are seen many times (>100) but

still show some uncertainty (probability <0.9 , >0.1) are a result of contradicting phenotypes (e.g. horned animals submitted as polled phenotypes)(data not shown).

It must be noted, that haplotype poll probability is entirely reliant on the phenotypes associated with them; this means that should a haplotype only ever be submitted with polled phenotypes, it is going to receive a poll probability close to one. That haplotype may in fact exist within horned populations, but without horned phenotypic data submitted to the test, the poll probability cannot reflect that existence.

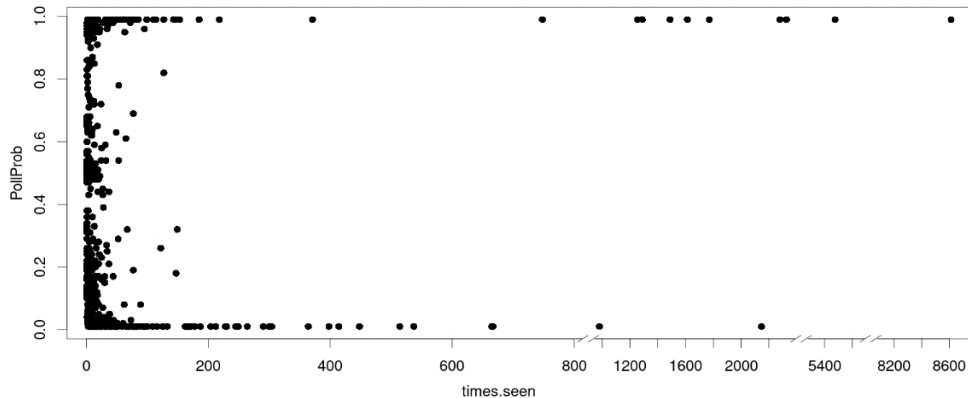


Figure 3. Estimated haplotype probability (polled) as a function of number of observations.

Analysis of sampling trends

Prior to commercialisation, approximately 1,750 animals, from 20 breeds, were submitted to the test, with phenotypes of relatively high accuracy, including progeny-tested data (Piper, *et al.*, 2014). Since its use in industry, more than 23,000 additional animals have been included in the test, from over 50 breeds, with little or no quality control over the phenotypes submitted. Breeders are unlikely to spend money on submission of horned animals to determine their poll genetics (which is obviously not homozygous polled), nor are they likely to submit scurred animals for similar reasons. More commonly, breeders submit polled or unknown phenotyped animals, seeking confirmation of homozygous polled status for breeding or marketing purposes. Without economic incentives or subsidies, it seems unlikely breeders will voluntarily bear the extra cost required for submission of horned phenotypes.

In addition, the quality of the phenotyping from industry is also unreliable (Henshall, *et al.*, 2011). Phenotyping can be problematic due to the similarity between scurs and horns. In some cases, animals phenotyped with scurs, if given time to develop, may actually become horned, though often dehorning or animal slaughter makes it impossible to know which. Animals submitted as polled, may actually have been dehorned as calves, thus should have horned phenotypes. There are known instances of scurred animals submitted to the test with polled phenotypes. Whether deliberate, to influence the result of the test, or a simple error in submission remains unknown, though regardless of the reason, haplotype probabilities are affected.

It was previously suggested that there was potential to obtain some ascertainment bias in the testing of animals (Henshall, *et al.*, 2011). A plot showing phenotypes submitted over time, since its release to industry in 2014 to early 2017, is shown in Figure 4, and demonstrates a growing number of unknown phenotype submissions, with a steady increase in polled phenotypes but little increase in informative horned phenotypes.

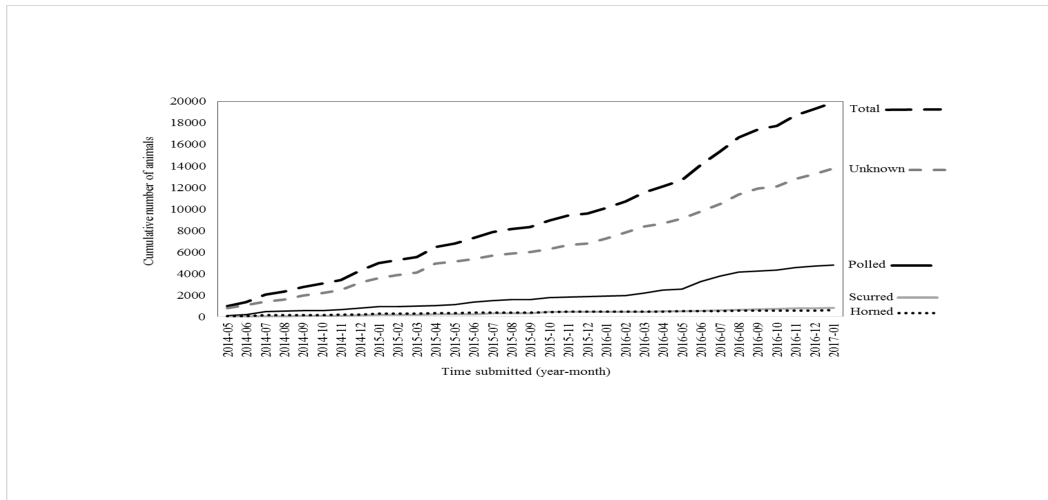


Figure 4. Cumulative number of animals and phenotypes submitted over time.

At the time of writing (2017), the vast majority of samples submitted to the test are phenotyped as unknown (>60%), with over a quarter phenotyped as polled (27%) and horned and scurred phenotypes account for approximately 5 percent each (Figure 5). In comparison to the data used for validation of the test in 2014 (Figure 5) before release to industry, the current number of informative phenotypes, particularly horned animals, is alarmingly low. Whether a submission strategy was suggested to breeders is unknown, though a clear lacking in horned phenotypes is evident.

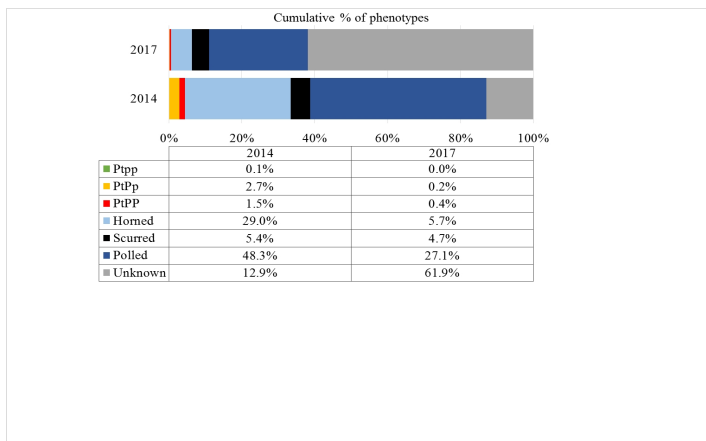


Figure 5. Comparison of phenotype proportions from validated research data (2014) and industry data (2017). Where: PtPP = progeny tested homozygous polled, PtPp = progeny tested heterozygous polled, PtpP = progeny tested homozygous horned.

Figures 4 and 5 demonstrate an obvious ascertainment bias in the test, for which the effects are only now becoming apparent. Anecdotal evidence has revealed animals with inaccurate genotype estimations, frequently based on inaccurate phenotype submissions; for example homozygous polled (PP) bulls with horned progeny. In addition, contradicting phenotypes for some haplotypes are increasing their uncertainty and moving their poll probabilities toward the middle (i.e. closer to 0.5). This is increasing the number of animals

that are not assigned a genotype estimation (probability drops below 90%). Table 1 shows commonly tested breeds with large percentage of animals that are not assigned a genotype estimation; also included is the number of horned animals in the breed's test population. This shows a trend where animals that have poor genotype assignment rates are usually from breeds for which we have very small numbers of horned animals sampled. There are some exceptions however, such as Brangus, with reasonable assignment rates and few horned animals; this is more so related to the lack of horned animals within the breed, and fewer numbers of haplotypes.

The machine-learning nature of the test, that is the dependency of haplotype poll probabilities on associated phenotypes, allows a vulnerability to ascertainment bias. Studies have demonstrated a quantifiable impact of mislabelled phenotypes on accuracy of genotype assignments (Biffani *et al.*, 2017), though being able to quantify the impact relies on knowledge of what phenotypes have been mislabelled. In this case, due to a lack of validation of data submitted, there is no way of knowing which samples are inaccurate.

Table 1. Observations for commonly tested breeds; sorted in descending order of non-assignment.

Breed¹	No. of haplotypes seen	No. of horned animals	No. of total animals	% Not Assigned
Bonsmara	64	1	182	62.64
Charbray	139	73	515	49.32
Charolais	148	89	1836	41.99
Simmental	97	70	446	35.87
Brahman	410	371	3878	28.37
Droughtmaster	193	54	1103	25.93
Santa Gertrudis	197	113	2682	23.08
Brangus	139	14	694	19.45
Limousin	120	62	1661	11.80
Hereford	87	57	2873	11.45

¹ *Columns for each breed: number of haplotypes seen, number of horned animals, total number of animals, percentage of animals with genotypes not assigned (less than 90% probability).*

Future improvements

Recently funding has been acquired to supplement the test population with validated horned phenotypes, to clarify uncertainty of haplotype probabilities and improve accuracy of genotype estimations and assignment rates. During the test development, it was suggested that ongoing refinement would be required, as it is dependent on the continual goodwill of users, and as such appropriate breeder engagement and potential subsidy of horned phenotypes would be required (Henshall, *et al.*, 2014). Supplementation with horned phenotypes will address this request and address the sampling bias that has occurred, though ongoing engagement with industry must be enhanced, along with encouragement and subsidies for submission of horned phenotypes in addition to polled.

Recently, the patent expired for the commercially available SNP test for Celtic polled mutation, P_C, and as such the test is likely to be available at an affordable price in Australia. This test will provide further indication of polled genetics, particularly for *Bos Taurus* breeds, though its application in *Bos Indicus* is still under investigation. In the process of applying the SNP test to *Bos Indicus* cattle, the haplotype test can be used as a research validation tool, but

will require continued funding. In addition, the SNP test can be retrospectively applied to samples showing some probability uncertainty within the haplotype test, to clarify its genotype. Each of these processes are currently ongoing.

Further to the identification of poll genetics, the question of scur and African horn genetics will also have to be addressed in future. Scurred animals are also commonly dehorned and as such will be subject to animal welfare legislation. There is potential for the haplotype test to be developed further for investigation of scur genetics, though improvements in the quality control and validation of phenotypes will be required going forward.

Conclusions

Predicting the poll genetics of an animal requires accurate estimates of polled haplotype probabilities. The accuracy of these estimates is dependent on unbiased sampling of the population, and reliable and varied phenotypes. Samples submitted for the haplotype poll test are shown to have significant ascertainment bias, which has affected the accuracy of the test. A project targeting the supplementation of the test with horned animals is underway and will improve accuracies. Ongoing engagement with industry must be increased and economic incentives to entice submission of horned phenotypes should be employed, to prevent the same sampling biases from occurring again.

Acknowledgements

This work is supported by Meat and Livestock Australia, project B.BFG.0050. Test development (2011-2014) funded under MLA projects B.AHW.0144 and B.AWW.0222. We acknowledge the use of data from the Beef Cooperative Research Centre, and numerous breeders and breed societies.

List of References

- Asai, M., T.G. Berryere & S.M. Schmutz, 2004. The scurs locus in cattle maps to bovine chromosome 19. *Animal Genetics* 35 (1): 34-39.
- Biffani, S., H. Pausch, H. Schwarzenbacher & F. Biscarini, 2017. The effect of mislabeled phenotypic status on the identification of mutation-carriers from SNP genotypes in dairy cattle. *BMC research notes* 10 (1): 230.
- Bortolussi, G., J.G. McIvor, J.J. Hodgkinson, S.G. Coffey & C.R. Holmes, 2005. The northern Australian beef industry, a snapshot. 2. Breeding herd performance and management. *Australian Journal of Experimental Agriculture* 45 (9): 1075-1091.
- Bunter, K.L., D.J. Johnston, M.L. Wolcott & G. Fordyce, 2013. Factors associated with calf mortality in tropically adapted beef breeds managed in extensive Australian production systems. *Animal Production Science* 54 (1): 25-36.
- Capitan, A., C. Grohs, M. Gautier & A. Eggen, 2009. The scurs inheritance: new insights from the French Charolais breed. *BMC Genetics* 10 33.
- Connors, N.K. & B. Tier, 2016. Breeding polled cattle in Australia. In: *Breeding Focus 2016 - Improving Welfare*, editors (S. Hermes and S. Dominik), *Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW. p 59-66.
- Connors, N.K., B. Tier & D.J. Johnson, 2016. Increased accuracy of the Poll DNA marker test for Australian beef cattle. In *'Proceedings of the Northern Beef Research Update*

- Conference.' Rockhampton, QLD, Australia.
- CSIRO, 2014. Annual Report 2013-2014. p 71. Available at www.csiro.au/annualreport2014
- Georges, M., R. Drinkwater, T. King, A. Mishra, S.S. Moore, D. Nielsen, L.S. Sargeant, A. Sorensen, M.R. Steele, X. Zhao & et al., 1993. Microsatellite mapping of a gene affecting horn development in *Bos taurus*. *Nature Genetics*, 4 (2): 206-10.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57 (1): 97-109.
- Henshall, J.M., E.K. Piper & A. McDonald, 2011. Effects of ascertainment bias when testing for polled in Australian beef cattle. In 'Proceedings of the Association for the Advancement of Animal Breeding and Genetics.' Perth, WA, Australia.
- Henshall, J.M., E.K. Piper & B. Tier, 2014. Haplotype Tests for Diagnosis of QTL and Genes. In 'Proceedings of the 10th World Congress on Genetics Applied to Livestock Production.' Vancouver, Canada.
- Long, C.R. & K.E. Gregory, 1978. Inheritance of the horned, scurred, and polled condition in cattle. *Journal of Heredity* 69 (6): 395-400.
- Mariasegaram, M., B.E. Harrison, J.A. Bolton, B. Tier, J.M. Henshall, W. Barendse & K.C. Prayaga, 2012. Fine-mapping the POLL locus in Brahman cattle yields the diagnostic marker CSAFG29. *Animal Genetics* 43 (6): 683-8.
- Medugorac, I., D. Seichter, A. Graf, I. Russ, H. Blum, K.H. Göpel, S. Rothhammer, M. Förster & S. Krebs, 2012. Bovine Polledness - An Autosomal Dominant Trait with Allelic Heterogeneity. *PLoS One*, 7 (6): e39477.
- Piper, E.K., B. Tier & J.M. Henshall, 2014. A Haplotype Diagnostic for Polled in Australian Beef Cattle. In 'Proceedings of the 10th World Congress on Genetics Applied to Livestock Production.' Vancouver, Canada.
- Prayaga, K.C., M. Mariasegaram, B. Harrison, B. Tier, J.M. Henshall & W. Barendse, 2009. Genetic markers for polled condition in cattle: the current status and the future plans. In 'Proceedings of the Association for the Advancement of Animal Breeding and Genetics.' Adelaide, SA, Australia.
- Seichter, D., I. Russ, S. Rothhammer, J. Eder, M. Förster & I. Medugorac, 2012. SNP-based association mapping of the polled gene in divergent cattle breeds. *Animal Genetics* 43 (5): 595-598.
- Sinnwell, J.P., D.J. Schaid & Z. Yu, 2007. haplo. stats: Statistical analysis of haplotypes with traits and covariates when linkage phase is ambiguous [Online]. URL <http://www.mayo.edu/research/labs/statistical-genetics-genetic-epidemiology/software>
- White, W.T. & H.L. Ibsen, 2008. Horn inheritance in Galloway-Holstein cattle crosses. *Journal of Genetics* 32 33-49.
- Winks, L., A. Holmes & P. O'Rourke, 1977. Effect of dehorning and tipping on liveweight gain of mature Brahman crossbred steers. *Australian Journal of Experimental Agriculture* 17 (84): 16-19.