

Selecting for Nut Characteristics in Macadamia Using a Genome-wide Association Study

Katie O'Connor, Ben Hayes, Craig Hardner, Mobashwer Alam, and Bruce Topp¹

Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St. Lucia 4072, Qld, Australia

Additional index words. GWAS, nut weight, kernel weight, kernel recovery, SNP markers, breeding

Abstract. Current macadamia breeding programs involve a lengthy and laborious two-stage selection process: evaluation of a large number of unreplicated seedling progeny, followed by replicated trials of clonally propagated elite seedlings. Yield component traits, such as nut-in-shell weight (NW), kernel weight (KW), and kernel recovery (KR) are commercially important, are more easily measured than yield, and have a higher heritability. A genome-wide association study (GWAS) combined with marker-assisted selection offers an opportunity to reduce the time of candidate evaluation. In this study, a total of 281 progeny from 32 families, and 18 of their 29 parents have been genotyped for 7126 single nucleotide polymorphism (SNP) markers. A GWAS was performed using ASReml with 4352 SNPs. We found five SNPs significantly associated with NW, nine with KW, and one with KR. Further, three of the top 10 markers for NW and KW were shared between the two traits. Future macadamia breeding could involve prescreening of individuals for desired traits using these significantly associated markers, with only predicted elite individuals continuing to the second stage of selection, thus potentially reducing the selection process by 7 years.

Macadamias (*Macadamia integrifolia*, *Macadamia tetraphylla*, and their hybrids) are grown commercially around the world for their high-quality nuts. The genus is from the ancient Proteaceae family and native to the east coast of Australia (Nock et al., 2016). Macadamia was first developed as an international commercial crop in Hawaii in the early 1920s (Hardner et al., 2009). Nuts consist of an edible kernel enclosed in a hard shell, which are grown inside a husk on the tree (Hardner et al., 2009). In 2017, nut-in-shell production in South Africa comprised ≈25% (46,490 mt) of the world's total crop, followed by Australia (43,000, 23%), Kenya (30,500, 16%) and the United States (Hawaii, 17,900, 9%) (Australian Macadamia Society, 2018). Breeding new macadamia varieties is usually focused on harnessing

higher nut yield per tree. However, selecting for yield is challenging due to the polygenic nature of the trait, low heritability ($H \approx 0.12$), and large genotype by environment interactions (Hardner et al., 2002).

Yield component traits such as NW, KW, and KR are also important selection criteria for new varieties (Hardner et al., 2009). NW between 6.5 and 7.5 g per nut is desirable, due to the ease of handling and cracking compared with smaller nuts (Hardner et al., 2009). Desirable KW is 2 to 3 g, which is a selection criteria that is described as “intermediate optimum” (Falconer, 1989); kernels <1.5 g and >3.5 g have issues with roasting and processing (Hardner et al., 2009). KR, the percentage of KW to NW, is very important in terms of production and processing costs, and as such is a major determinant of the per kilogram price farmers receive from processing factories for consignments. It is worth noting, however, that cultivars with high KR may have thin shells and be more susceptible to pests and diseases (Hardner et al., 2009).

GWAS examine genetic markers across the genome individually and test for a significant association with a particular trait (Khan and Korban, 2012). If genetic markers could be discovered that were associated with important traits, these markers explained a reasonable proportion of the genetic variation in these traits, and their chromosome location known, then the markers could be used to increase genetic gain. This can be achieved by combining GWAS with marker-assisted selection (MAS). MAS is a method whereby

candidate cultivars are selected indirectly based on genetic markers linked with desirable traits (Collard et al., 2005; Tester and Langridge 2010). Screening seedlings for markers significantly associated with important traits could predict these trait measurements years before they are actually expressed. As such, GWAS followed by MAS is a feasible way of accelerating selection cycles, by selecting candidate cultivars at an earlier stage, and thus improving genetic gain (Khan and Korban, 2012; Isik et al., 2015).

Long generation times of tree crops combined with large plant size generally means lengthy and laborious phenotyping to identify superior genotypes in tree crop breeding. Research into using GWAS for improving fruit and nut tree crop breeding is expanding, with significant associations found for fruit quality traits in Japanese pear (Iwata et al., 2013; Yamamoto et al., 2014), apple (Kumar et al., 2013), and peach (Cao et al., 2012). Recently, O'Connor et al. (2018) evaluated the potential to use genomics in macadamia breeding to improve varieties. The current study aimed to identify genetic markers associated with NW, KW, and KR in macadamia using GWAS for future use in MAS.

Materials and Methods

Plant material. The Australian macadamia breeding program's B1.2 population is the focus of this study. The entire population included 2000 seedlings from 141 families, which were planted across nine locations in southeast Queensland and northeast New South Wales, Australia, between 2001 and 2003 (Topp et al., 2016). This study involved 281 progeny from 32 families across four of these sites in Queensland, as well as 18 of their 29 parents ($n = 299$).

Phenotypic analysis. Historical data were used in the study, taken from trees in 2010, when the trees were 7 to 9 years old. A sample of 50 nuts per tree was taken and dried to 1% moisture content. An average value for NW was obtained. Nuts were cracked mechanically, with kernel and shell weighed separately to calculate average KW and also KR per tree. NW and KW were both log transformed [$\log_{10}(x + 1)$] due to the skewed nature of the data.

Genotyping and association analysis. Each tree was genotyped for 7126 SNP markers by Diversity Arrays Technology. Marker locations were unknown, as a complete reference genome is currently not available. SNPs were pruned for ≥50% call rate and ≥2.5% minor allele frequency across individuals, leaving 4352 markers for analysis. A genomic relationship matrix (GRM) was constructed using R to model the kinship of individuals.

Association analysis was performed for the three traits using ASReml (Gilmour et al., 2009) in R, using a mixed model:

$$y = Wb + Xg + Zu + e,$$

where y is a vector of phenotypes, W is a matrix allocating fixed effects to individuals,

Received for publication 12 June 2018. Accepted for publication 23 Aug. 2018.

This paper was presented as a part of the 2017 International Macadamia Research Symposium, 13–14 Sept. 2017, in Big Island, HI.

This research was funded by Hort Innovation Australia, using the macadamia research and development levy and contributions from the Australian Government. Hort Innovation is the grower-owned, not-for-profit research and development corporation for Australian horticulture. KO acknowledges the Australian Postgraduate Award and Charles Morphet Peglar scholarship for financial support. We thank anonymous reviewers for their suggestions and comments.

¹Corresponding author. E-mail: b.topp@uq.edu.au.

Table 1. Phenotypic minimum, maximum, mean, broad- and narrow-sense heritability for nut weight (NW), kernel weight (KW), and kernel recovery (KR).

	NW	KW	KR
Minimum	3.29 g	1.07 g	18.6%
Maximum	12.43 g	4.89 g	52.7%
Mean	6.21 g	2.28 g	36.9%
Broad-sense heritability (H^2)	0.59	0.53	0.62
Narrow-sense heritability (h^2)	0.56 ± 0.08	0.51 ± 0.09	0.62 ± 0.10

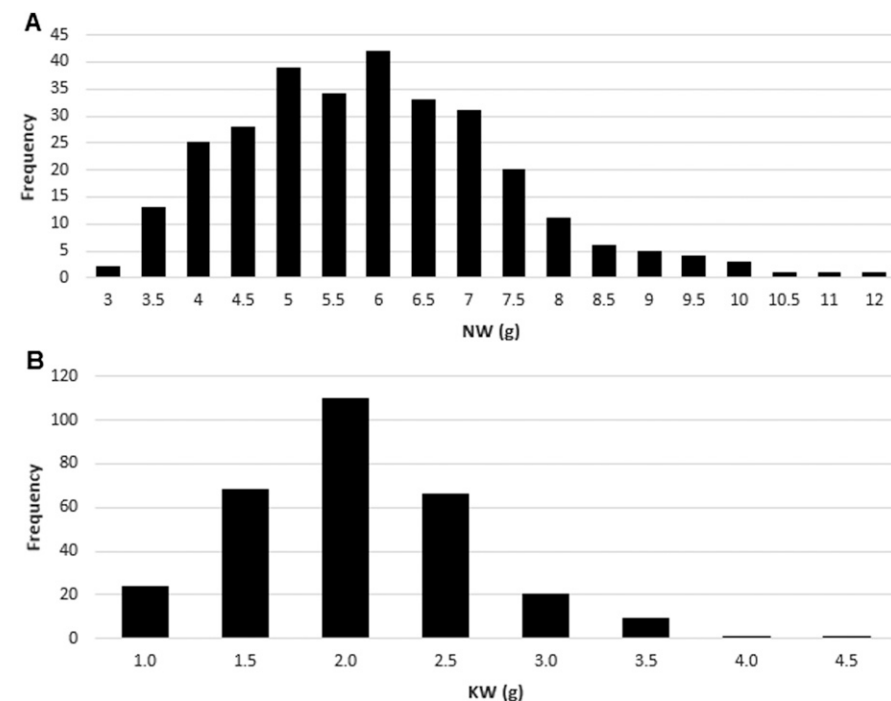


Fig. 1. Histogram of skewed phenotypes for (A) nut weight (NW) and (B) kernel weight (KW).

Table 2. Phenotypic correlations (above diagonal) and genetic correlations based on best linear unbiased predictions (below diagonal) between nut characteristics.

	NW	KW	KR
NW	—	0.85***	-0.16**
KW	0.80***	—	0.37***
KR	-0.27***	0.34***	—

** $P < 0.01$; *** $P < 0.001$.

b is a vector of fixed effects, X is a design matrix allocating records to the marker effect (modeled as 0, 1, or 2 for homozygous, heterozygous, and alternate homozygous genotypes, respectively), g is the fixed effect of the marker currently being fitted in the model, Z is a design matrix allocating records to individuals, and u is a vector of breeding values of the individuals, assumed random $\sim N(0, G\sigma_g^2)$, where G is the GRM among the individuals using the same 4352 SNPs, constructed following VanRaden (2008), σ_g^2 is the genetic variance captured by the SNP, and e is a vector of random error. This model is additive, in that two copies of the second allele will have double the effect of one copy.

We included SNP and site as fixed effects, and genotype as a random effect. Two models were tested for each trait, both including and excluding genotype by site as a random effect. Log likelihoods of the two models

were compared using a χ^2 test to determine if the models were statistically different. For all traits, there was no significant difference between models, so the genotype by site term was excluded. To further account for population structure, the first two principal coordinates (PCs) were calculated from the GRM and included as fixed effects in the model. PCs 1, 2, and 3 explained 32%, 19%, and 15% of the variance, respectively, but were not significant and were therefore not included in further analyses.

Best linear unbiased predictions (BLUPs; genetic values) were calculated for each genotype using the model above, but excluding SNP effects, using ASReml-R and the GRM (Gilmour et al., 2009). Phenotypic and genetic (based on BLUPs) correlations (Pearson's) were calculated among the three nut characteristics. Narrow-sense heritability was calculated using the GRM with the pin.r function in R (White 2013), and broad-sense heritability was calculated manually from variance components by including family in the model as a random effect.

Quantile-Quantile (Q-Q) plots were constructed to compare observed and expected significance of markers and ensure that population structure was accounted for in the analysis. SNP markers with a significance level of 1×10^{-4} or lower for each trait were fitted simultaneously in a multiple regression

model to determine if any were in linkage disequilibrium. In this case, SNPs that were no longer significant were therefore considered redundant, as it was assumed that these SNPs represented the same underlying quantitative trait loci (QTL). Significance of non-redundant SNPs was calculated.

Results and Discussion

NW ranged from 3.29 g to 12.43 g, with a mean of 6.21 g (Table 1; Fig. 1A). In comparison, weight of kernels varied between genotypes, from 1.07 g to 4.89 g, with an average KW of 2.28 g (Table 1; Fig. 1B). As a derivative of these two values, mean KR was 36.9%, and phenotypes ranged from 18.6% to 52.7% (Table 1).

Heritability ranged across traits from 0.51 for KW to 0.62 for KR (Table 1). Estimates of heritability for the three traits in this study were very similar to that of individual broad-sense heritability in a previous study of macadamia by Hardner et al. (2001), consisting of four replicates of 40 cultivars at three locations ($H^2 = 0.63$ for all three traits). The replication of cultivars may have allowed more accurate estimates and smaller error, and hence higher heritability in the study by Hardner et al. (2001) compared with the current study. Our estimates of heritability were higher than that of NW and KW in pecan ($h^2 = 0.35$ and 0.38 , respectively) (Thompson and Baker, 1993). In this study, estimates of narrow-sense and broad-sense heritability were similar for each trait (Table 1), indicating that the model could not detect any dominance. Alternatively, this could imply limited dominance exists for these traits.

NW and KW were highly correlated both phenotypically (0.85 , $P < 0.001$) and genetically (0.80 , $P < 0.001$), whereas KW and KR were moderately but significantly correlated ($r_p = 0.37$, $r_g = 0.34$, $P < 0.001$) (Table 2). KR decreased with larger NW ($r_p = -0.16$, $r_g = -0.27$) (Table 2). This may imply that nuts with thicker shells have smaller kernels and hence lower KR.

The Q-Q plot for NW (Fig. 2A) indicates that our model has accounted for population structure, because there were a similar number of observed and expected SNPs at low levels of significance [as suggested by Korte and Farlow (2013)]. However, genomic inflation has occurred, where the SNPs are rising above the one-to-one line as P values become more stringent, suggesting polygenic inheritance of the trait (Yang et al., 2011). Five SNPs were found to be significantly associated with NW after fitting 14 SNPs in a multiple regression model (Table 3), with the reduced number of significant markers in the multiple regression likely reflecting some linkage disequilibrium between the markers.

Similar to NW, some genomic inflation was apparent in the Q-Q plot for KW (Fig. 2B). Nine markers remained nonredundant after simultaneously fitting 19 markers significantly associated with KW (Fig. 2B; Table 3). Three of these 10 SNPs were also

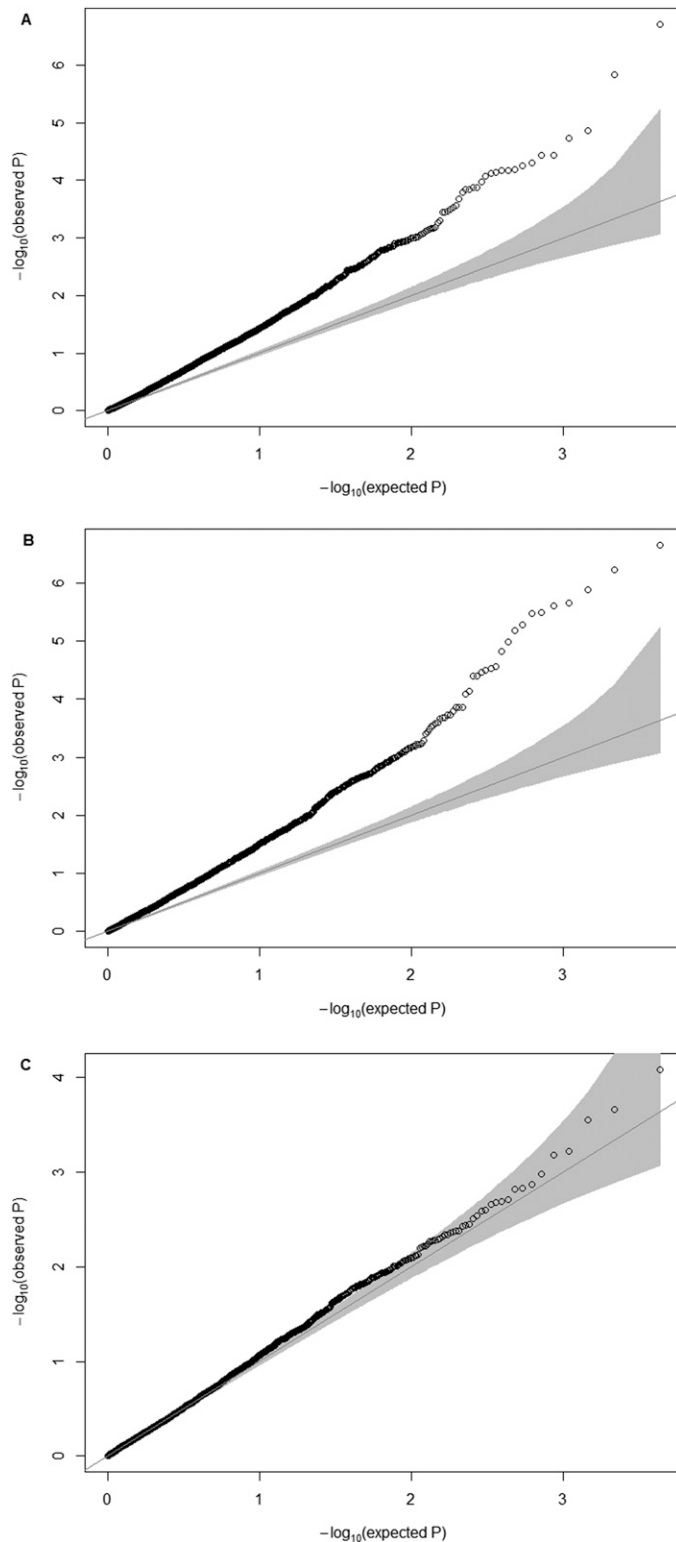


Fig. 2. Quantile-quantile plots of observed and expected $-\log_{10}(P)$ values for various nut characteristics: (A) nut weight, (B) kernel weight, and (C) kernel recovery. Each point represents one SNP marker.

among the most significant markers associated with NW (Table 3). One SNP was significantly associated with KR (Fig. 2C; Table 3). Because KR is a derivative of NW and KW, KR could be estimated if significant markers for NW and KW are used in MAS.

One limitation with our research was that there is currently no completely assembled macadamia genome, so our markers could

not be easily mapped; however, multiple regression results suggest that several significant markers are detecting the same QTL. Research by Nock et al. (2016) has sequenced $\approx 79\%$ of the macadamia genome. When a more complete reference genome becomes available, the location of significant markers for macadamia NW, KW, and KR can be determined. Then, MAS could be used to

increase genetic gain in macadamia breeding. Future progeny seedlings could be genotyped at target markers for each trait to determine the allelic state, and breeders predict NW and KW, and sequentially cull undesirable individuals before phenotypic expression of these traits.

In conclusion, we have identified markers linked to three commercially important nut traits in macadamia through a GWAS. This study is part of a larger project investigating genetic markers controlling target traits in macadamia. Future research will examine the economic benefits of using these markers to identify parents and progeny with desirable NW, KW, and high KR. GWAS combined with MAS for important nut characteristics would be an asset for the macadamia breeding program. The efficiency and accuracy of using genomic selection to improve nut yield also should be investigated by the Australia macadamia breeding program. Application of genomics may reduce the length of the selection cycle and increase genetic gain in macadamia breeding.

Literature Cited

- Australian Macadamia Society. 2018. Estimated World Macadamia Production. In: XXXVII International Nut and Dried Fruit Congress, Spain.
- Cao, K., L. Wang, G. Zhu, W. Fang, C. Chen, and J. Luo. 2012. Genetic diversity, linkage disequilibrium, and association mapping analyses of peach (*Prunus persica*) landraces in China. *Tree Genet. Genomes* 8:975–990.
- Collard, B.C.Y., M.Z.Z. Jahufer, J.B. Brouwer, and E.C.K. Pang. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* 142:169–196.
- Falconer, D.S. 1989. Introduction to quantitative genetics. Longman Scientific & Technical, Essex, England.
- Gilmour, A.R., B. Gogel, B. Cullis, R. Thompson, and D. Butler. (2009) ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.
- Hardner, C., C. Winks, R. Stephenson, and E. Gallagher. 2001. Genetic parameters for nut and kernel traits in macadamia. *Euphytica* 117:151–161.
- Hardner, C.M., C.W. Winks, R.A. Stephenson, E.G. Gallagher, and C.A. McConchie. 2002. Genetic parameters for yield in macadamia. *Euphytica* 125:255–264.
- Hardner, C.M., C. Peace, A.J. Lowe, J. Neal, P. Pisanu, M. Powell, A. Schmidt, C. Spain, and K. Williams. 2009. Genetic resources and domestication of *Macadamia*. *Hort. Rev.* 35:1–126.
- Isik, F., S. Kumar, P.J. Martínez-García, H. Iwata, and T. Yamamoto. 2015. Acceleration of forest and fruit tree domestication by genomic selection, p. 93–124. In: C. Plomion and A.-F. Adam-Blondon (eds.). *Advances in botanical research, land plants - trees*, vol 74. Elsevier, Oxford, UK.
- Iwata, H., T. Hayashi, S. Terakami, N. Takada, Y. Sawamura, and T. Yamamoto. 2013. Potential assessment of genome-wide association study and genomic selection in Japanese pear *Pyrus pyrifolia*. *Breed. Sci.* 63:125–140.
- Khan, M.A. and S.S. Korban. 2012. Association mapping in forest trees and fruit crops. *J. Expt. Bot.* 63:4045–4060.

Table 3. List of nonredundant significant single nucleotide polymorphisms (SNPs) and their significance values (*P*) for nut weight (NW), kernel weight (KW), and kernel recovery (KR). Bolded markers are also significant for another trait. See Supplemental Table 1 for marker sequences. SNP ID does not imply chromosome location.

NW		KW		KR	
SNP	<i>P</i>	SNP	<i>P</i>	SNP	<i>P</i>
SNP2607	1.31×10^{-8}	SNP0289	1.58×10^{-9}	SNP4603	8.17×10^{-5}
SNP4565	7.68×10^{-7}	SNP4668	6.94×10^{-9}		
SNP5959	9.06×10^{-7}	SNP3809	9.57×10^{-8}		
SNP6113	1.78×10^{-5}	SNP2607	7.96×10^{-8}		
SNP4829	3.65×10^{-5}	SNP2026	4.19×10^{-5}		
		SNP0803	6.00×10^{-5}		
		SNP4565	3.98×10^{-5}		
		SNP5959	4.85×10^{-5}		
		SNP1196	4.69×10^{-5}		

Korte, A. and A. Farlow. 2013. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* 9:29.

Kumar, S., D.J. Garrick, M.C. Bink, C. Whitworth, D. Chagné, and R.K. Volz. 2013. Novel genomic approaches unravel genetic architecture of complex traits in apple. *BMC Genomics* 14:393–406.

Nock, C.J., A. Baten, B.J. Barkla, A. Furtado, R.J. Henry, and G.J. King. 2016. Genome and transcriptome sequencing characterises the

gene space of *Macadamia integrifolia* (Proteaceae). *BMC Genomics* 17:937.

O'Connor, K., B. Hayes, and B. Topp. 2018. Prospects for increasing yield in macadamia using component traits and genomics. *Tree Genet. Genomes* 14:7.

Tester, M. and P. Langridge. 2010. Breeding technologies to increase crop production in a changing world. *Science* 327:818–822.

Thompson, T. and J. Baker. 1993. Heritability and phenotypic correlations of six pecan nut characteristics. *J. Amer. Soc. Hort. Sci.* 118:415–418.

Topp, B., C.M. Hardner, J. Neal, A. Kelly, D. Russell, C. McConchie, and P.J. O'Hare. 2016. Overview of the Australian macadamia industry breeding program. *Acta Hort.* 1127:45–50.

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423.

White, I. 2013. Pin function for asreml-R. 2017.

Yamamoto, T., S. Terakami, N. Takada, S. Nishio, N. Onoue, C. Nishitani, M. Kunihiisa, E. Inoue, H. Iwata, and T. Hayashi. 2014. Identification of QTLs controlling harvest time and fruit skin color in Japanese pear (*Pyrus pyrifolia* Nakai). *Breed. Sci.* 64:351–361.

Yang, J., M.N. Weedon, S. Purcell, G. Lettre, K. Estrada, C.J. Willer, A.V. Smith, E. Ingelsson, J.R. O'Connell, M. Mangino, R. Magi, P.A. Madden, A.C. Heath, D.R. Nyholt, N.G. Martin, G.W. Montgomery, T.M. Frayling, J.N. Hirschhorn, M.I. McCarthy, M.E. Goddard, and P.M. Visscher. 2011. Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* 19:807–812.

Supplemental Table 1. Single nucleotide polymorphisms (SNPs) significantly associated with nut traits, Allele ID, Allele sequences for reference allele, and SNP allele, and trimmed sequence for the reference allele.

SNP Code	Allele ID	Allele sequence reference	Allele sequence SNP	Trimmed sequence reference
SNP0289	19933185 F 0-29:A>G	TGCAGAGGTCATAGAGACCATCGGAGATCATG GATGCAAAAATCCATTACGAGGACGGAGAC TGTGAGTG	TGCAGAGGTCATAGAGACCATCGGAGATCGTGGATGC AAAAATCCATTACGAGGACGGAGACTGTGAGTG	TGCAGAGGTCATAGAGACCATCGGAGATCATGGATGCAA AATCCATTACGAGGACGGAGACTGTGAGTG
SNP0803	19935921 F 0-35:C>T	TGCAGTTGTAGTGAAGGCTCTTCTCTACCCCTG GTGCTGTTGCTAAATGGTATCATCAACAATA AGACCAT	TGCAGTTGTAGTGAAGGCTCTTCTCTACCCCTGGTGTG TTGCTAAATGGTATCATCAACAATAGACCAT	TGCAGTTGTAGTGAAGGCTCTTCTCTACCCCTGGTGTGTT GCTAAATGGTATCATCAACAATAGACCAT
SNP1196	19936059 F 0-24:C>A	TGCAGAAAATCGTTGCACAGACCCTCTCTC CTTCCATCTCTCCATCTCCAGCCGAAGA TCAATGATA	TGCAGAAAATCGTTGCACAGACCCTACTTCTCTCTCCAT CTTCTCCATCTCCAGCCGAAGATCAATGATA	TGCAGAAAATCGTTGCACAGACCCTCTCTCTCCATCT TCTCCATCTCCAGCCGAAGATCAATGATA
SNP2026	19931971 F 0-47:C>T	TGCAGAAATCTCGGAGCCGGAGGGAAGGGAA CCCACAGCAGCCTACTATCAAGGTTCA TGAAGAAGA	TGCAGAAATCTCGGAGCCGGAGGGAAGGAAACCCACA CGACGACCTATTAACAAGTTTCAATGAAGAAGA	TGCAGAAATCTCGGAGCCGGAGGGAAGGAAACCCACAC GACGACCTACTATCAAGGTTTCAATGAAGAAGA
SNP2607	19936700 F 0-14:G>A	TGCAGAGTGTGGAAGATGCGAGATCGGAAGA GCGGTTTCAGCAGGAATGCCGAGACCGATC TCGTATGCC	TGCAGAGTGTGGAAGATGCGAGATCGGAAGAGCGGTT CAGCAGGAATGCCGAGACCGATCTCGTATGCC	TGCAGAGTGTGGAAGATGCGG GCATAGAAAGTGAGGATGAACATCATGTGGA
SNP3809	19933411 F 0-16:T>G	TGCAGGAGCTACACCTCTTAGGCTACAACAT GCAAGTAGTTTATTTGGGGTCTGTTGG TTAGATG	TGCAGGAGCTACACCTCTTAGGCTACAACATGCAAA GTAGTTTATTTGGGGTCTGTTGGTTAGATG	TGCAGGAGCTACACCTCTTAGGCTACAACATGCAAAAGT AGTTTATTTGGGGTCTGTTGGTTAGATG
SNP4565	19935406 F 0-47:G>A	TGCAGCTTAGGATGGTCTAGAGGAGCATGAAG ATGATGTGCATAGAAAGTGAGGATGAACATC ATGTGGA	TGCAGCTTAGGATGGTCTAGAGGAGCATGAAGATGA TGTGCATAGAAATGAGGATGAACATCATGTGGA	TGCAGCTTAGGATGGTCTAGAGGAGCATGAAGATGATGT GCATAGAAAGTGAGGATGAACATCATGTGGA
SNP4603	19933117 F 0-20:G>C	TGCAGAAATCAGGCAGAGGGAGAAAAAAGCA GCAACAACAAGGTTACAAAAACCAAGA GAAATGG	TGCAGAAATCAGGCAGAGGGACAAAAAAGGCAGCA CAACAAGGTTAAAAAACCAGAGAAATGG	TGCAGAAATCAGGCAGAGGGAGAAAAAAGCAGCAACA ACAAAGGTTAAAAAACCAGAGAAATGG
SNP4668	19931867 F 0-8:T>C	TGCAGTCTTTGGTGGTCTGCGAGATCGGAAGA GCGGTTTCAGCAGGAATGCCGAGACCGATCT CGTATG	TGCAGTCTTTGGTGGTCTGCGAGATCGGAAGAGCGG TTCAGCAGGAATGCCGAGACCGATCTCGTATG	TGCAGTCTTTGGTGGTCTGCGG
SNP4829	19929082 F 0-56:A>G	TGCAGAAATGGTGTCTCCATGCCGTGTTTTCT GCCAAAGCTGATTAATGCTTACCATGAGATC AAGG	TGCAGAAATGGTGTCTCCATGCCGTGTTTTCTGCCT AAGCTGATTAATGCTTACCCTGAGATCAAGGC	TGCAGAAATGGTGTCTCCATGCCGTGTTTTCTGCCTAAG CTGATTAATGCTTACCATGAGATCAAGGC
SNP5959	19929268 F 0-26:C>G	TGCAGCATTAGTTTTGGTCTTTGTACTACGGT ATGCAATTAATCAGTTTTGCAACCATATIG ATGTT	TGCAGCATTAGTTTTGGTCTTTGTACTACGGTATGC AATTAATCAGTTTTGTCAACCATATTTGATGTT	TGCAGCATTAGTTTTGGTCTTTGTACTACGGTATGCAAT TAATCAGTTTTGTCAACCATATTTGATGTT
SNP6113	19930919 F 0-10:T>G	TGCAGCCTTCTCACTTTGTTGATGTTGGGT GCGGAAATCCTATCTCTCTACCACTTCG CGATCT	TGCAGCCTTCTCACTTTGTTGATGTTGGGTGCGGGA ATCCTATCTCTCTACCACTTCGCGATCT	TGCAGCCTTCTCACTTTGTTGATGTTGGGTGCGGGAAT CCTATCTCTCTACCACTTCGCGATCT