Final Report

Development of genomic multi-breed eating quality trait estimates using shared global data

Project code: L.GEN.2000

Prepared by: Drs Peter McGilchrist, Sam Clark, Hassan Aliloo, Ben Hayes, Mehrnush Forutan & Rod Polkinghorne

University of New England, University of Queensland & Birkenwood Pty Ltd

Date published: 2025

PUBLISHED BY
Meat & Livestock Australia Limited
PO Box 1961
NORTH SYDNEY NSW 2059

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

Abstract

This project aimed to determine if sharing global consumer data could predict eating quality of beef using genotypes. The main objective was to establish the accuracy of genomic multi-breed values for eating quality. Striploin samples were consumer-tested by 10 individuals per sample following the Meat Standards Australia (MSA) consumer testing protocols. The samples were derived from 3,273 USDA MARC carcasses, 1,331 ICBF carcasses, 1,298 Southern Multibreed and 564 Northern BIN cattle along with 3298 historic MSA samples. Each consumer evaluated the samples for tenderness, juiciness, flavour, and overall liking which were combined using the weightings of 0.3, 0.1, 0.3 and 0.3 to generate the meat quality 4 score (MQ4). After quality control, the dataset comprised 3,526 records from Australia (including historic data sets collected prior to the project), 3,197 records from the US, and 1,297 records from Ireland for analysis.

Genomic predictions for five meat eating quality traits were performed using single-trait analyses, conducted separately for each country and with the combined data from all countries. The analyses utilized 709k imputed markers and were conducted using a GBLUP and BayesR models. Prediction accuracies were evaluated using two reference sets: one comprising of only one countries data and the other incorporating Australian, USA and Irish data. Using the GBLUP model, heritability estimates for tenderness, juiciness, flavour, overall liking and MQ4 score ranged from 0.19 to 0.31 for Australia, 0.07 to 0.20 for the US, 0.09 to 0.17 for Ireland, and 0.14 to 0.22 for the combined dataset. Using the BayesR model, heritability estimates ranged from 0.23 to 0.36 for Australia, 0.07 to 0.18 for the US, 0.13 to 0.17 for Ireland, and 0.17 to 0.26 for the combined dataset. Prediction accuracies for all traits from the Australian-only reference were moderate (0.39 to 0.69), but they increased to high accuracies when international data was included in the reference population (0.55 to 0.89) depending on trait and analysis used. This highlights the value of incorporating international data into the reference population, creating a larger and more diverse dataset, and ultimately improving prediction accuracies. The multi-breed genomic prediction generated by this project could be utilised by the Australian beef industry to create an eating quality EBV. This genomic prediction could be updated with further data from Australia, USA and Ireland on a tri-annual basis.

Executive summary

Background

Historically, selection for meat quality in beef cattle has utilised indicator traits like intramuscular fat and shear force. A shift towards breeding for more consumer-driven priorities like tenderness, juiciness and flavour is necessary. Meat Standards Australia (MSA) ensures international consumer satisfaction of meat but lacks the ability for genetic evaluation of breeding stock. Generating a genomic breeding value for multiple breeds of beef cattle is necessary to drive increased selection for eating quality. This project integrated data from ICBF, USDA MARC, and MSA to enhance global prediction of eating quality traits. By combining diverse datasets, it aimed to develop genetic tools for commercial and seedstock cattle, enabling producers to enhance beef quality for both Australian and global markets.

Objectives

The main objectives of the project achieved were:

- Leverage international research and develop a global eating quality dataset – Create a multi-breed reference dataset with international partners (ICBF, USDA MARC, MSA) for genomic breeding value estimation.
- Enhance global linkages – Connect Australian cattle data with international populations to improve genetic imports and ranking.
- Establish multi-breed genomic values and assess genomic prediction accuracy – Evaluate how well genomic data predicts eating quality, compare country-specific versus global predictions along with their suitability to integrate into Breedplan for widespread adoption and long-term herd improvement.

Methodology

Striploin samples from 3,273 USDA MARC, 1,331 ICBF, 1,298 Southern Multibreed, 564 Northern BIN cattle were consumer-tested by 10 individuals per sample using MSA protocols. This was collated with historic data from 3298 cattle from MSA. Consumers rated tenderness, juiciness, flavour, and overall liking for all samples which were combined into the meat quality 4 score (MQ4). The final dataset for analysis included 3,526 Australian, 3,197 US, and 1,297 Irish records. Genomic predictions for all meat quality traits were conducted separately for each country and combined using 709k imputed markers with GBLUP and BayesR models. Prediction accuracies were assessed using country-specific and international reference sets in a 5-fold cross-validation.

Results/key findings

Using the GBLUP model, heritability estimates for tenderness, juiciness, flavour, overall liking and MQ4 ranged from 0.19 to 0.31 for Australia, 0.07 to 0.20 for the US, 0.09 to 0.17 for Ireland, and 0.14 to 0.22 for the combined dataset. Using the BayesR model, heritability estimates ranged from 0.23 to 0.36 for Australia, 0.07 to 0.18 for the US, 0.13 to 0.17 for Ireland, and 0.17 to 0.26 for the combined dataset. Prediction accuracies for all traits from the Australian-only reference were moderate at 0.62 to 0.69 (GBLUP analysis) and 0.39 to 0.58 (BayesR analysis). These accuracies increased to high accuracies of 0.83 to 0.89 (GBLUP analysis) and 0.55 to 0.64 (BayesR analysis) when international data was included in the reference population.

Benefits to industry

The heritabilities of all 5 eating quality traits (Tenderness, Juiciness, Flavour, Overall liking and MQ4) are moderately heritable traits and there is significant variation between animals in all eating quality traits, hence there is merit in selection for improvements in eating quality across the international beef herd. The multi-breed genomic prediction generated by this project is considered accurate and therefore could be utilised by the Australian beef industry to create an eating quality EBV. The accuracies for the GBLUP analysis from the AUS+USA+IRE reference population showed notable improvements compared to the AUS-only reference, increasing by 0.21 for tenderness (from 0.62 to 0.83), by 0.20 for juiciness (from 0.69 to 0.89), by 0.25 for flavour (from 0.64 to 0.89), by 0.23 for overall liking (from 0.64 to 0.87), and by 0.23 for MQ4 (from 0.63 to 0.86). This indicates that combining international data gives much greater accuracy of prediction for all 5 eating quality traits. This benefit was not only limited to Australia but was seen across the 3 populations, strongly demonstrating the benefit of combining forces to improve the genomic prediction accuracy of hard to measure, expensive traits.

Future research and recommendations

- Genomic predictions for eating quality are demonstrating high enough heritability and accuracy to explore commercial multi-breed products for delivery through BREEDPLAN or other genetic analyses.
- Careful consideration is required into the design of future reference populations to ensure they are developed to predict the future target animals of interest, whilst also utilising data from animals consumer tested for other projects.
- The multibreed prediction has low predictive accuracy for within breeds. Therefore, the application of the multibreed algorithm developed in this project is not suitable for use within purebred populations. The international partners (USDA and ICBF) will have access to the prediction equation developed via commercial licences.
- Permitted use of the prediction equation will be for the purpose of generating genomic breeding value products on the breeds and breed types involved in the project. Its use will be restricted to providing a service within country.
- Modelling is required to understand the numbers of eating quality records that need to be collected by each country annually, biannually or tri-annually to maintain constant improvement of the genomic prediction for eating quality.
- Understand the appetite from USDA and ICBF for further collaboration, further collection of sensory phenotypes. There is opportunity to establish a collaboration based on routine data collection and routine updates to the equation using the numbers modelled from the point above. This would be arranged as an update to the licence prepared for the above or under a separate licence.
- Combining data from multiple countries improved prediction accuracy. This should be investigated for other key hard to measure and expensive to record traits.
- Combining data is very beneficial for all countries that contributed. The expense to generate enough eating quality data individually is cost prohibitive to the development of a breeding value.
- Generate an MSA database that is not excel based to allow phenotypes and genotypes to be stored.

Table of contents

1. Background

Historically the value of breeding cattle has been principally based on improved growth, milk and fertility traits and correlated measures of meat quality and quantity. Recent years have seen a shift from producer-driven to consumer-driven beef production, with significant emphasis placed on consumer satisfaction and interest in eating quality indicators, with incentives implemented by beef brands for improved compliance and product quality. The Meat Standards Australia (MSA) grading system underpins international consumer satisfaction. However, MSA development has been poorly linked to beef cattle genetic evaluation due to scarcity of carcass data that is used in the evaluation. This is more evident for consumer satisfaction as there is a lack of a direct eating quality trait estimate.

A number of challenges exist to create a closer link between consumer satisfaction and prediction of genetic merit and selection of breeding animals. The first challenge is that consumer preference is hard to measure and expensive to record. Whilst in the past this has been seen as a major limitation, the use of DNA based genomic prediction methods has improved the possibility for using trait records from animals outside of breeding populations. Using such data allows for the creation of a larger reference dataset (or reference population) for genomic prediction. To increase the size of the reference population used for genomic prediction, the number of animals used for analysis in this research project (L.GEN.2000) will combine data from different breeding organisations, the Irish Cattle Breeding Federation (ICBF), the United States Department of Agriculture Meat Animal Research Centre (USDA MARC) and Meat Standards Australia (MSA). This sharing of data is significant and may help to underpin global prediction of consumer satisfaction traits in beef cattle.

This provides an opportunity for the development of analyses that incorporate eating quality and production traits from both commercial and seedstock animals of varying breed content. To develop these tools, large and varied data are required and must include extensive phenotypic measures. Existing MSA R&D database records include consumer outcomes from over 70,000 cuts and 19,000 cattle. Limitations exist however for many commercial animals that do not have extensive background data, known cohorts or DNA available for analysis.

This project will review all available Australian data that includes consumer satisfaction trait records and genotypes or process stored DNA samples to provide a base for genomic studies. This project is a collaboration between ICBF, USDA MARC and Meat and Livestock Australia (MLA). Combining datasets to form a global reference population will assist in prioritising genetic outputs from this project. This project aims to collate data from cattle with records of production phenotypes, consumer eating quality information and genomic information. This information is then to be used to determine if genomics can predict eating quality and an estimated breeding value can potentially be developed and adopted by commercial beef producers to continue to improve product quality for Australian and international beef markets.

2. Project objectives

The major objectives of the project include;

1.) A global eating quality reference dataset from which genomic estimated breeding values can be calculated for multiple breeds. Construction of the data will occur with reference to international partners and be compatible with ICAR and Interbeef. The chairman of ICAR and Interbeef is a member of the project team. The dataset will combine extensive multi-breed data from diverse but genomically connected cattle populations from ICBF and USDA MARC and MSA data.
2.) Evaluation of the effectiveness of genomic data in predicting eating quality outcomes.
3.) Comparison of the prediction accuracy of each individual country's data relative to a prediction derived from the combined international dataset.
4.) An assessment of how genomic eating quality prediction may be best applied to the MSA model, the core issue being use of a single carcass or individual muscle genomic input. Existing MSA data when combined with genomic data will enable an immediate start on this question. Further validation across international datasets will test the transportability of genomic breeding values across breeds and environments and guide the need for further research.
5.) Provide a vehicle for integration of other NLGC projects to expand Australian herd data and relationships to major global cattle populations. Linkage to global populations will allow eating quality ranking and assist in prioritising genetic imports.
6.) Effective comparison and linkage to European and USA consumer populations may be achieved through complementary MSA research utilising paired Australian sourced product with that sourced from ICBF and USDA MARC. For example a current project is consumer testing USDA MARC sourced striploin in conjunction with multiple Australian cuts of different qualities in Texas from which paired Australian samples are being sensory tested by Australian consumers.
7.) The Australian investment in funding sensory testing of ICBF and MARC product will be massively leveraged through access to the huge embedded investment and depth of data within the ICBF and USDA MARC programmes.
8.) Establishment of genomic multi-breed values.
9.) Incorporation of genetic eating quality values in the MSA model will deliver immediate commercial incentives for supply chain adoption.
10.) MSA grading data may be combined with genotyping to empower more sophisticated commercial application.
11.) The introduction of multi-breed genomic value estimates into Breedplan will dramatically increase utilisation within commercial herds rather than being restricted to purchase of purebred sires. There will also be significant potential to obtain and utilise progeny data from purchased sires which will add greatly to their EBV accuracy.
12.) Commercial incentives arising from use of multi-breed genomic value estimates will result in a significant and ongoing improvement in the Australian herd, similar to that experienced in the dairy industry.

3. Methodology

3.1 Meat sample collection & consumer testing in Ireland, USA and Australia

Arrangements for carcass grade data collection, cut acquisition, fabrication of cuts into MSA protocol consumer samples for freezing and storage, followed by allocation to individual consumer groups

(Picks) through picking, posting procedures, and sensory testing were established in Australia, the USA, and Ireland.

In the USA, the USDA Meat Animal Research Center (MARC) cattle were harvested at the Cargill plant in Schuyler, Nebraska. MARC personnel, actively engaged in this project, attended each kill and collected a 3-inch anterior portion of one striploin. This portion was chilled and transported to the MARC Meat Laboratory at Clay Center, Nebraska, for instrumental and chemical analysis, occasionally in conjunction with trained taste panel assessment.

Texas Tech University (TTU) was subcontracted by L.GEN.2000 to attend each kill, record full MSA and USDA grading inputs, and collect and label a 3-inch portion from the carcass side opposite to that used for the MARC sample for all cattle. While primary traits such as sex, carcass weight, marbling, and maturity were common to both USDA and MSA, some required conversion to a common scale (carcass weight adjusted to a fats-out AUS-MEAT equivalent in kg). MSA measures for hump height, fat distribution, and ultimate pH were additional. Live animal details, including breed composition, age, HGP status, performance, camera grading, laboratory, and genomic details, were provided by MARC into the data cloud for L.GEN.2000.

Arrangements were made with Cargill to purchase the TTU samples, which were then transported to the TTU Meat Laboratory in Lubbock, Texas, for fabrication into MSA consumer protocol samples after aging for 14 days to align with MARC laboratory treatments. An ultimate pH was recorded from each sample during fabrication. Texas Tech personnel were trained by Birkenwood in the MSA consumer test protocols. COVID restrictions posed challenges but were overcome. Further virtual training and protocol reinforcement were provided via Zoom conferencing from Australia.

In Ireland, two organizations were contracted for the kill and subsequent activities. ICBF, a core project partner, had existing arrangements with the ABP Slaney Foods factory in Bunclody. All cattle from the ICBF test center in Tully were slaughtered at Slaney, with associated arrangements for striploin (sirloin) sample preparation for ICBF objectives and trained panel evaluation after boning. Further paired samples for the NLGC project sensory samples were added to these existing arrangements. Grading and sample fabrication were conducted by Birkenwood International. Both graders were accredited MSA/IMR3GF assessors and held PhD or equivalent qualifications in Meat Science, coupled with industry and MSA consumer test experience.

Teagasc was contracted by Birkenwood to receive the frozen ICBF samples, which were fabricated and ready for consumer testing from Slaney, and to manage them through to consumer test events. Years prior, collaborative MSA collection and consumer testing had been conducted in Ireland utilizing the Ashtown Meat Laboratory and Teagasc staff. However, operational staff had since moved on, leaving only Dr. Declan Troy, the senior scientist and site manager, with direct experience. To rebuild expertise, Birkenwood conducted training in Wales and Dublin. Two senior sensory scientists traveled to Wales, where they participated in UNECE/MSA protocol sensory sample posting and a sensory event, in addition to receiving extensive protocol details and written materials. Follow-up training was conducted at Ashtown. Birkenwood International continued supervising sensory sessions to consolidate training.

Control of sample collection and subsequent activities in the USA and Ireland, as well as related Australian MSA testing through to consumer testing and data collection, was managed by Birkenwood. Files controlling sample fabrication, including unique identification codes, control files, and self-adhesive labels formatted for Letter (USA) or A4 (Ireland) printing, were provided by Birkenwood upon receipt of the cut collection detail and individual cut coding. Samples were then allocated to “Picks,” a set of 42 samples evaluated by 60 consumers, with 10 consumers evaluating

each of the 42 samples and all consumers being served seven samples. The first sample served was a presumed mid-eating-quality "Link," followed by six test samples. The test samples were served according to a 6 × 6 Latin Square design, ensuring each sample was presented in five different order positions and within five subsets of 12 consumers. Additionally, each of the six test samples was assigned from different "products" (i.e., striploins grouped by expected quality from low to high), with each product served an equal number of times in each order position and before and after each other product.

To ensure integrity, sample codes remained "blind" to the sensory organization, with sensory data linked to animal, grading, and genomic data only after receiving the consumer results. All data were double-entered, variance-checked, and electronically forwarded to Birkenwood for database inclusion. All raw data for L.GEN.2000 were stored on the UNE cloud, with appropriate sharing arrangements in place as per the head agreement.

3.2 Genotypes

The genotypes of 3298, 564, 1233, 3230 and 1312 animals were available from MEQ, NBIN, SMB, USA and IRE datasets, respectively. There were 167 duplicate samples within the historic MEQ dataset, of which only one record per animal was retained, resulting in a total of 3,131 animals from the MEQ dataset. In total, 9,470 genotyped animals were used across all datasets. Genotypes were imputed up to 709,768 SNPs (bovine high-density (HD) array) using the findhap software and a reference set of 4506 cattle from relevant breeds that were genotyped with the Bovine HD array. Figure 1 shows the distribution of animals by country, while Table 1 presents both the distribution of animals by country and groups, as well as the SNP density used for genotyping each group of animals.

Table 1. Distribution of animals by country and groups, as well as the SNP density used for genotyping each group of animals

Source	No. Genotypes	No. Animals	No. SNP
MEQ CRC	750	749	43180
MEQ ANGUS	18	18	40427
MEQ 100K	1795	1666	78646
MEQ 50K	561	560	39410
MEQ 50K tropbeef	174	138	46831
MEQ ALL	3298	3131	130842
NBIN	564	564	46831
SMB 50K	134	134	44018
SMB 100K	1099	1099	85820
USA	3230	3230	673630
IRE	1312	1312	671843

Total	9637	9470	709768

3.3 Phenotypes

In total for MEQ we had 36,061 consumer meat eating quality records for 3901 animals. These animals were from 36 master groups/cohorts. The 1893 records are from STR045 CUT and GRL COOK, while 478 records are from a different CUT or COOK.

Table 2. The number of unique animals, records, and the number of genotyped animals with phenotypes across multiple datasets

Dataset	No. unique Animals	No. Records A	No. of animals with both Phenotypes and Genotypes	No. Phenotypes and Genotypes
MEQ	3901	36061	1893	3564
SMB	1298	1296	1152	
NBIN	564	560	519	
USA	3273	3597	3197	3197
IRE	1331	1530	1304	1304

3.3.1 Quality control of phenotype data

3.3.1.1 Australia (AUS)

There is a total of 3564 records (1893, 1152 and 519 from MEQ, SMB and NBIN projects, respectively), 1 per animal with genotype, for 5 consumer meat eating quality traits CTender, CFlav, CJuicy, COall_Like and CMQ4. All records were from HGP free animals.

These records belong to 101 groups (Fig 1) with STR045 cut and GRL cooking method. The groups with less than 5 records (3 groups with a master group code of 492) were merged together so the final number of groups is 99.

The records are from meat samples aged between 3 to 52 days after slaughter, a total of 10 different durations of aging (Fig 2). The majority of records (n=3241) were from samples aged for 7 or 14 days after slaughter and the majority of groups had only records with a single days aged (Fig 3). This can potentially result in a confounding between group and days aged if both included in the model, but there are some groups with days aged of 7, 14 and 21 which may help in estimation of days aged effect.

The records are from animals slaughtered between 1997 to 2023 in a total of 14 distinct years (Fig 4). The distribution of kill years across groups is shown in Fig 5. As expected all groups except the old data from CRC have only animals killed in a single year.

The distribution of sex across different groups is shown in Fig 6. The majority of groups have only male animals but there are some groups with both male and females or only female animals. The groups with both male and female records can help with the estimation of sex effect in the model.

The boxplots of carcass weights across different groups are shown in Fig 7. There are some groups with lighter animals but in the majority of groups the range of carcass weights overlap. The histogram of carcass weight is shown in Fig 8. Records that were outside of the 2 times the interquartile range (n=35) were removed.

Fig 9 to 13 are the histograms of CTender, CFlav, CJuicy, COall_Like and CMQ4, respectively. Given that the consumer meat eating quality traits are bounded between 0 and 100, the application of IQR or SD methods to detect outliers can result in thresholds that are outside the boundaries (i.e. smaller than 0 or larger than 100). Therefore, relaxed thresholds based on the distribution of 99.9% of the records were used to remove the possible outliers. This resulted in removal of 8, 8, 6, 8 and 8 records for CTender, CFlav, CJuicy, COall_Like and CMQ4, respectively. There were 3 animals for which all of their eating quality traits were removed so in total 3526 animals remained.

Finally, there were 3521 records for CTender, CFlav, COall_Like and CMQ4 and 3523 records for CJuicy.

3.3.1.2 USA

The USA phenotype data consisted of five eating quality traits: tenderness (CTender), juiciness (CJuicy), flavor (CFlav), overall liking (COall_Like), and CMQ4. These traits were measured from striploin samples collected from 3197 cattle across 32 contemporary groups (CG) aged 14 days and one CG aged 21 days (Fig. 14). A total of 1,362 animals were hormone growth promoter (HGP) free, while 1,835 animals were treated with HGP (Fig. 15). Among these animals, there were 453 heifers and 2,744 steers (Fig. 16).

Figure 17 shows the box plot of carcass weight (ranged from 202.1 to 496.1kg). There were no outliers identified based on the criterion of being greater or lower than two times the interquartile range (IQR method).

Using relaxed thresholds based on the distribution of 99.9% of the records to identify and remove the potential outliers resulted in the removal of 8 records for CTender, CJuicy, COall_Like and CMQ4, and 7 records for CFlav. Fig 18 shows the box plots for the different traits.

3.3.1.3 Ireland (IRE)

The IRE phenotype data consisted of five meat quality traits: CTender, CJuicy, CFlav, COall_Like, and CMQ4. These traits were measured from striploin samples collected from

1304 cattle across 29 contemporary groups (CG) aged 14 days and one CG aged 16 days (Fig. 19). All animals were hormone growth promoter (HGP) free and there were 152, 328, and 824 bulls, heifers and steers, mostly confounded by the contemporary group.

Figure 20 shows the box plot of carcass weight. Carcass weight ranged from 185.2 to 471kg and 7 animals were identified as outliers and were removed based on the criterion of being greater or lower than two times the interquartile range (IQR method).

Using relaxed thresholds based on the distribution of 99.9% of the records to identify and remove the potential outliers resulted in the removal of 4 different records for each trait. Fig 21 shows the box plots for the different traits.

3.4 Statistical Analyses

Two approaches- GBLUP and BayseR- were used for variance component estimation and breeding value prediction.

GBLUP

Univariate. A mixed linear univariate animal model was fitted for each of the 5 meat eating quality traits from each country separately using blupf90+.

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

Where **y** is the vector of phenotypes; **b** includes the estimates of fixed cross classified contemporary group, SEX and HGP effects, and covariates of days aged after slaughter, carcass weight, first 4 principal components and heterosis; **u** is the vector of genomic estimated breeding value (GEBV) of animals; and **e** is the residual term. **X** and **Z** are incident matrices relating observations to fixed and random effects, respectively. It was assumed that $v(u) = G\sigma_a^2$ where G was the genomic relationship matrix based on VanRaden (2008) and σ_a^2 is the additive genetic variance. The models including full data were used to estimate variance components and heritabilites.

In addition, all phenotypes from the 3 countries for each meat eating quality trait were combined to estimate variance components a heritabilites by adding a fixed cross classified country effect to the above model.

Bivariate. A mixed liner bi-variate animal model was fitted for each of the 5 sensory eating quality traits and 4 carcass quality traits (Marbling, Rib fat, EMA and OSS) by treating the sensory (SENS) and carcass (CARC) eating quality records as different traits.

$$\begin{bmatrix}\mathbf{y}_{\mathbf{SENS}} \\ \mathbf{y}_{\mathbf{CARC}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{\mathbf{SENS}} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{\mathbf{CARC}}\end{bmatrix}\begin{bmatrix}\mathbf{b}_{\mathbf{SENS}} \\ \mathbf{b}_{\mathbf{CARC}}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{\mathbf{SENS}} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{\mathbf{CARC}}\end{bmatrix}\begin{bmatrix}\mathbf{a}_{\mathbf{SENS}} \\ \mathbf{a}_{\mathbf{CARC}}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_{\mathbf{SENS}} \\ \mathbf{e}_{\mathbf{CARC}}\end{bmatrix}$$

where $\mathbf{y}_{\mathbf{SENS}}$ and $\mathbf{y}_{\mathbf{CARC}}$ are vectors of phenotypes for sensory and carcass meat eating quality traits, respectively; $\boldsymbol{\beta}_{\mathbf{SENS}}$ and $\boldsymbol{\beta}_{\mathbf{CARC}}$ are vector of estimates for fixed effects including fixed cross classified contemporary group, SEX and HGP effects, and covariates of days aged after slaughter (only for SENS), carcass weight, first 4 principal components and heterosis; $\mathbf{u}_{\mathbf{SENS}}$ and $\mathbf{u}_{\mathbf{CARC}}$ are vectors of

genomic estimated breed values; $\mathbf{e}_{\mathbf{SENS}}$ and $\mathbf{e}_{\mathbf{CARC}}$ are vectors of random residual terms; $\mathbf{X}_{\mathbf{SENS}}$ and $\mathbf{X}_{\mathbf{CARC}}$ and $\mathbf{Z}_{\mathbf{SENS}}$ and $\mathbf{Z}_{\mathbf{CARC}}$ are incidence matrices relating observations to the fixed and random additive genetic animal effect, respectively.

It was assumed that $\begin{bmatrix}\mathbf{u}_{\mathbf{SENS}}\\ \mathbf{u}_{\mathbf{CARC}}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{u_{\mathbf{SENS}}} & \sigma_{u_{\mathbf{SENS}},u_{\mathbf{CARC}}}\\ \sigma_{u_{\mathbf{CARC}},u_{\mathbf{SENS}}} & \sigma^2_{u_{\mathbf{CARC}}}\end{bmatrix} \otimes \mathbf{G}\right)$, with genetic variances on diagonal and genetic covariances as off-diagonals; and $\begin{bmatrix}\mathbf{e}_{\mathbf{SENS}}\\ \mathbf{e}_{\mathbf{CARC}}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\mathbf{W}_{\mathbf{SENS}} & 0\\ 0 & \mathbf{W}_{\mathbf{CARC}}\end{bmatrix}\begin{bmatrix}\sigma^2_{e_{\mathbf{SENS}}} & \sigma_{e_{\mathbf{SENS}},e_{\mathbf{CARC}}}\\ \sigma_{e_{\mathbf{CARC}},e_{\mathbf{SENS}}} & \sigma^2_{e_{\mathbf{CARC}}}\end{bmatrix}\right)$, with residual variances on diagonal and residual covariances as off-diagonals;

Only Australian animals with records for both sensory and carcass eating quality traits (n=2548) were used for this analysis.

BayseR

Univariate. In the BayesR approach (Erbe et al. 2012), the following prediction equation was used to estimate the SNP effects for each of the five meat quality traits for each country.

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Zg} + \boldsymbol{e} \ldots$$

where $\boldsymbol{y}$ is a vector of phenotypes, $\boldsymbol{b}$ includes the estimates of fixed class effects of contemporary group, SEX and HGP treatments, and covariates of days aged after slaughter, carcass weight, first 4 principal components (PCs) and heterosis, $\boldsymbol{g}$ is a vector of $\boldsymbol{m}$ SNP effects with distribution $\boldsymbol{g} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{\sigma}^2_{\boldsymbol{i}})$. In turn, the genetic variance of the trait ($\boldsymbol{\sigma}^2_{\boldsymbol{g}}$) was assumed $\boldsymbol{\sigma}^2_{\boldsymbol{i}} = \{\mathbf{0},\ \mathbf{10^{-4}}\boldsymbol{\sigma}^2_{\boldsymbol{g}},\ \mathbf{10^{-3}}\boldsymbol{\sigma}^2_{\boldsymbol{g}},\ \mathbf{10^{-2}}\boldsymbol{\sigma}^2_{\boldsymbol{g}}\}$. This setup allows the BayesR model (Moser et al., 2015) to have a more flexible SNP effect distribution which is a mixture of four possible normal distributions: $\boldsymbol{N}(\mathbf{0}, \mathbf{0}), \boldsymbol{N}(\mathbf{0},\ \mathbf{10^{-4}}\boldsymbol{\sigma}^2_{\boldsymbol{g}}), \boldsymbol{N}(\mathbf{0}, \mathbf{10^{-3}}\boldsymbol{\sigma}^2_{\boldsymbol{g}}),\ \boldsymbol{N}(\mathbf{0}, \mathbf{10^{-2}}\boldsymbol{\sigma}^2_{\boldsymbol{g}})$, all with a mean of 0 but with different variances. $\boldsymbol{Z}$ is a matrix $\boldsymbol{n\ x\ m}$ of standardized genotypes and $\boldsymbol{e}$ is a vector of random residuals with $\boldsymbol{e} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{\sigma}^2_{\boldsymbol{e}})$. In addition, phenotypes for each meat-eating quality trait from the three countries were combined to estimate variance components and heritability, with a fixed class effect of country included in the model.

Notice that in contrast to GBLUP models (Yang, Lee, Goddard, & Visscher, 2011), the BayesR model simultaneously provides estimates for the $\boldsymbol{m}$ SNP effects ($\widehat{\boldsymbol{g}} = \boldsymbol{E}[\boldsymbol{g}]$), the genetic ($\boldsymbol{\sigma}^2_{\boldsymbol{g}} = \boldsymbol{E}[\boldsymbol{\sigma}^2_{\boldsymbol{g}}]$), and residual variances ($\boldsymbol{\sigma}^2_{\boldsymbol{e}} = \boldsymbol{E}[\boldsymbol{\sigma}^2_{\boldsymbol{e}}]$), and the SNP-based heritability ($\boldsymbol{h} = \boldsymbol{E}[\boldsymbol{\sigma}^2_{\boldsymbol{g}}]/(\boldsymbol{E}[\boldsymbol{\sigma}^2_{\boldsymbol{e}}] + \boldsymbol{E}[\boldsymbol{\sigma}^2_{\boldsymbol{g}}]$. By default, in the literature (Moser et al., 2015), these estimates are expected values of the parameters over the MCMC chain, e.g. mean of the posterior distribution $\boldsymbol{E}[.]$ for each parameter. The model was estimated using Marcov Chain Monte Carlo (MCMC) with Gibbs Sampling in GCTB with 40,000 iterations of which the first 5,000 are discarded as burn-in.

3.5 Cross-validation.

To assess the accuracies of genomic evaluations for AUS animals using both single-trait (GBLUP and BayseR) and multi-traits (GBLUP) models, four 5-fold cross validation strategies were implemented.

RANKED1- In the first strategy, AUS animals were sorted based on their first principal component (PC1) and assigned into 1 of the 5 groups of approximately equal size such that in each group there were animals with smaller PC1 values as well as those with larger PC1 values.

RANKED2- In the second strategy, AUS animals were sorted based on PC1 and divided into 5 groups of approximately equal size. The top 20% of animals were placed in the first group, the second 20% in the next group, and so on, with the bottom 20% grouped together in the final group.

YOBIRT -In the third strategy, AUS animals were sorted based on their year of birth (YOB) and assigned into 5 groups of approximately equal size such that in each group there were both older animals as well as more recent animals.

RANDOM -In the fourth strategy grouping of animals was done randomly such that all groups have approximately equal size.

In each rotation of the cross validation the phenotypes of 1 group was masked and the remaining 4 groups were used to estimate the GEBV of the group without phenotypes. The accuracy of genomic prediction for each group was calculated as the Pearson correlation between GEBV obtained from cross-validation and raw/corrected phenotypes of validation animals obtained from full models above, dived by the square root of the trait heritability (Hayes et al. 2009). We also used method LR for evaluating the accuracies by comparing the GEBV from whole data using GBLUP approaches, obtained from full models above, and GEBV from partial data obtained from cross-validations where the phenotypes were masked. Accuracies were averaged across 5 groups and standard error was calculated as the standard deviation divided by the square root of the number of groups i.e. SE=STD/√5. The following 2 references were used to evaluate the predictive abilities of evaluation model for AUS animals.

4. Results

4.1 Population structure

The genetic structure of the 9,470 genotyped animals used for calculating the GRM is visualized by plotting PC1 against PC2. The first principal component (PC1) clearly separated Bos taurus and Bos indicus animals, while the second principal component (PC2) distinguished different Bos taurus breeds. The IRE animals were primarily of Bos taurus origin and formed a tight cluster. The USA animals included individuals with both Bos taurus and Bos indicus backgrounds; however, none of the USA animals clustered with the AUS animals that have

pure indicus ancestry. The AUS animals were the most genetically diverse, representing highly divergent breeds, and were scattered across the PC plot.

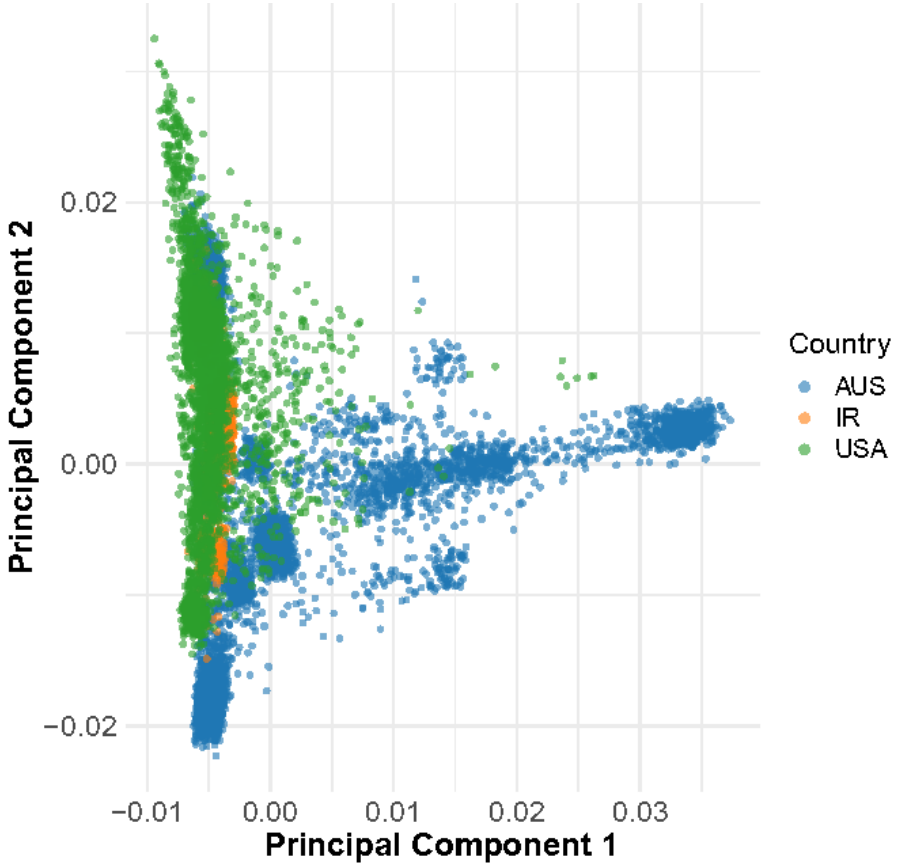

Figure 1. Principal component analysis (PCA) results for genotyped animals (Legends indicate the contributing countries, reflecting the genetic diversity of the dataset. NB: Serious overlap for IRE, USA and AUS on PC2)

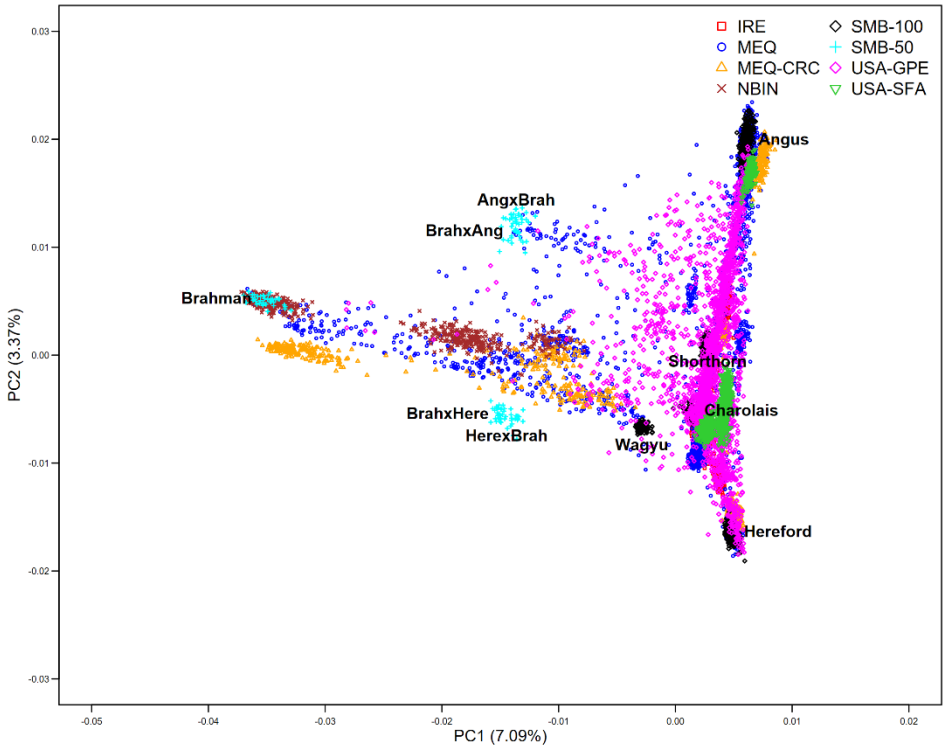

Figure 2. Principal component analysis (PCA) results for genotyped animals (Legends indicate the contributing countries and text on figure represents breeds.)

4.2 Univariate models for estimation of genetic parameters

Table 3 and 4 present the estimates of variance components and heritabilities (± standard errors) for the five meat quality traits, obtained from a single-trait GBLUP and BayseR models, respectively. Among all countries, tenderness exhibited the highest additive genetic variance, leading to the highest heritabilities across the five meat quality traits. The estimated heritabilites were highest in AUS compared to USA and IRE, with the largest difference observed for tenderness. The lowest heritabilities were found for juiciness in AUS and IRE, and for flavour in the USA. When combining data from all three countries (ALL), the heritability estimates were intermediate, falling between the within-country heritability estimates.

Table 3. Variance components and heritabilities for five meat quality traits using single trait GBLUP model for each country, as well as using combined records from three countries as a single trait.

Trait	Variances	AUS	USA	IRE	ALL
CTender	σ_a^2	56.47 ± 8.91	28.18 ± 5.99	21.24 ± 9.53	34.62 ± 3.97
	σ_e^2	124.48 ± 7.15	111.74 ± 5.36	109.69 ± 8.92	121.78 ± 3.50
	σ_p^2	180.95	139.92	130.93	156.40
	h^2	0.31 ± 0.05	0.20 ± 0.04	0.16 ± 0.07	0.22 ± 0.02
CJuicy	σ_a^2	26.92 ± 6.49	22.47 ± 5.93	10.53 ± 7.64	20.33 ± 3.15

	σ_e^2	117.57 ± 5.72	124.32 ± 5.60	101.66 ± 7.58	120.38 ± 3.10
	σ_p^2	144.49	146.79	112.19	140.71
	h^2	0.19 ± 0.04	0.15 ± 0.04	0.09 ± 0.07	0.14 ± 0.02
CFlav	σ_a^2	22.63 ± 4.95	7.65 ± 3.50	16.00 ± 7.22	13.99 ± 2.27
	σ_e^2	84.59 ± 4.27	95.21 ± 3.75	85.03 ± 6.80	90.94 ± 2.28
	σ_p^2	107.22	102.86	101.03	104.93
	h^2	0.21 ± 0.04	0.07 ± 0.03	0.16 ± 0.07	0.13 ± 0.02
COall_Like	σ_a^2	34.74 ± 6.32	13.67 ± 4.25	17.59 ± 7.95	21.06 ± 2.82
	σ_e^2	97.15 ± 5.24	101.86 ± 4.26	92.82 ± 7.47	101.15 ± 2.68
	σ_p^2	131.89	115.53	110.42	122.21
	h^2	0.26 ± 0.04	0.12 ± 0.04	0.16 ± 0.07	0.17 ± 0.02
CMQ4	σ_a^2	34.99 ± 6.07	15.16 ± 3.85	16.75 ± 7.26	34.62 ± 3.97
	σ_e^2	89.89 ± 4.97	82.10 ± 3.65	83.85 ± 6.79	121.78 ± 3.50
	σ_p^2	124.88	97.26	100.59	156.40
	h^2	0.28 ± 0.05	0.16 ± 0.04	0.17 ± 0.07	0.22 ± 0.02

Table 4. Variance components and heritabilities for 5 meat quality traits using single trait BayseR model for each country, as well as using combined records from three countries as a single trait.

Trait	Variances	AUS	USA	IRE	ALL
CTender	σ_a^2	68.41 (16.37)	25.66 (5.76)	22.91 (8.20)	43.88 (9.29)
	σ_e^2	121.16 (7.01)	113.86 (5.21)	107.68 (8.21)	120.53 (3.36)
	h^2	0.36 (0.06)	0.18 (0.04)	0.17 (0.06)	0.26(0.04)
CJuicy	σ_a^2	35.03 (12.72)	23.05 (5.36)	15.11 (4.94)	27.49 (6.10)
	σ_e^2	115.65 (5.70)	123.74 (5.29)	97.68 (6.01)	119.23 (3.04)
	h^2	0.23 (0.06)	0.16 (0.03)	0.13 (0.04)	0.19 (0.03)
CFlav	σ_a^2	30.09 (8.48)	7.56 (2.10)	15.73 (6.15)	18.49 (3.80)
	σ_e^2	82.22 (3.98)	95.33 (3.02)	84.90 (6.37)	90.11 (3.81)
	h^2	0.27 (0.06)	0.07(0.02)	0.15 (0.06)	0.17 (0.03)
COall_Like	σ_a^2	43.75 (13.88)	11.58 (4.11)	17.63 (6.74)	28.29 (5.42)
	σ_e^2	95.49 (5.24)	103.74 (4.34)	92.31 (6.88)	99.61 (2.62)

	h^2	0.31 (0.07)	0.10 (0.03)	0.16 (0.06)	0.22 (0.03)
CMQ4	σ_a^2	43.76 (13.35)	13.25 (3.36)	16.43 (5.29)	28.24 (5.97)
	σ_e^2	88.09 (4.84)	83.71 (3.45)	83.59 (5.66)	87.93 (2.31)
	h^2	0.33 (0.07)	0.14 (0.03)	0.16 (0.05)	0.24 (0.04)

4.3 Cross-validation for Australia

Table 5 and 6 present the accuracies of 5-fold cross-validation for the five meat quality traits, using GBLUP Model with AUS only or AUS+USA+IRE reference populations, respectively. Table 6 represents the accuracies of 5-fold cross-validation for the five meat quality traits, using BayseR model with AUS only or AUS+USA+IRE reference populations, specifically for the RANKED1 and YOBIRTH strategies.

For both the GBLUP and BayseR models, when using the AUS-only reference, juiciness had the highest prediction compared to the other meat-eating quality traits across the different strategies (Table 5 and 7). The accuracies, calculated as the correlation between GEBV and raw phenotype divided by the square root of heritability and averaged across five folds, were similar for the remaining four traits, ranging from 0.44 and 0.72 (GBLUP model: RANKED1, YOBIRTH and RANDOM validation strategies) and from 0.36 to 0.60 (BayseR model: RANKED1 and YOBIRTH validation strategies).

Using the GBLUP model, among the different validation strategies, the RANKED2 strategy consistently resulted in the lowest prediction accuracies for all traits, with values ranging from 0.26 to 0.31.

Table 5. Genomic prediction accuracies from GBLUP averaged across 5 folds for AUS only reference.

RANKED1	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.44 (0.02)	0.72 (0.03)	0.55 (0.02)	0.49 (0.01)	0.49 (0.01)
cor(res,gebv)/h	0.35 (0.03)	0.34 (0.02)	0.30 (0.03)	0.34 (0.03)	0.34 (0.02)
LR_bias	0.01 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.03)	0.00 (0.01)
LR_slope	0.98 (0.04)	1.01 (0.03)	0.98 (0.04)	1.00 (0.04)	0.99 (0.03)
LR_accuracy	0.62 (0.01)	0.69 (0.00)	0.64 (0.01)	0.64 (0.01)	0.63 (0.01)
RANKED2	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.26 (0.14)	0.38 (0.20)	0.33 (0.18)	0.32 (0.17)	0.31 (0.16)
cor(res,gebv)/h	0.22 (0.05)	0.25 (0.07)	0.18 (0.06)	0.22 (0.06)	0.22 (0.05)
LR_bias	0.15 (0.11)	0.06 (0.15)	0.07 (0.08)	0.13 (0.12)	0.14 (0.09)
LR_slope	1.01 (0.15)	1.20 (0.18)	1.03 (0.14)	1.08 (0.16)	1.03 (0.14)

LR_accuracy	0.44 (0.05)	0.51 (0.07)	0.44 (0.07)	0.46 (0.06)	0.45 (0.06)
YOBIRTH	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.45 (0.04)	0.72 (0.02)	0.57 (0.03)	0.51 (0.04)	0.5 (0.04)
cor(res,gebv)/h	0.37 (0.01)	0.33 (0.02)	0.33 (0.03)	0.36 (0.01)	0.36 (0.00)
LR_bias	0.01 (0.06)	0.00 (0.04)	-0.01 (0.05)	-0.01 (0.05)	0.00 (0.06)
LR_slope	1.01 (0.04)	1.02 (0.04)	1.03 (0.07)	1.02 (0.04)	1.02 (0.03)
LR_accuracy	0.63 (0.01)	0.69 (0.01)	0.65 (0.00)	0.64 (0.01)	0.64 (0.01)
RANDOM	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.44 (0.03)	0.71 (0.05)	0.56 (0.05)	0.49 (0.04)	0.49 (0.04)
cor(res,gebv)/h	0.36 (0.03)	0.33 (0.04)	0.32 (0.03)	0.34 (0.03)	0.35 (0.03)
LR_bias	0.02 (0.07)	0.00 (0.04)	-0.01 (0.02)	0.01 (0.05)	0.00 (0.04)
LR_slope	1.01 (0.08)	1.04 (0.09)	1.03 (0.07)	1.03 (0.07)	1.02 (0.07)
LR_accuracy	0.62 (0.01)	0.69 (0.01)	0.64 (0.01)	0.64 (0.01)	0.63 (0.01)

Adding international data from the USA and IRE to the reference population increased the prediction accuracies for all traits across all validation strategies, for both GBLUP and BayesR models (Table 6 and 7), highlighting the added value of a larger and more diverse reference population. The prediction accuracies followed a trend similar to the AUS-only reference population, with RANKED2 yielding the lowest accuracies (ranging from 0.39 to 0.47), while the other three strategies produced similar accuracies, ranging from 0.64 to 0.84.

Using GBLUP model with the RANDOM validation strategy, the accuracies from the AUS+USA+IRE reference population showed notable improvements compared to the AUS-only reference. For tenderness, accuracy increased by 0.20 (from 0.44 to 0.64); for juiciness, by 0.12 (from 0.71 to 0.83); for flavor, by 0.18 (from 0.56 to 0.74); and for overall liking and MQ4, by 0.19 (from 0.49 to 0.68).

Using the BayseR model with the RANKED1 validation strategy, adding the international data (USA+IRE) to the reference population, resulted in accuracy improvements of 0.22 and 0.23, for tenderness and MQ4, respectively, with a smaller increase of 0.01 for juiciness (from 0.58 to 0.59).

These findings underscore the benefits of incorporating international data to enhance prediction accuracies for meat eating quality traits in Australia.

Table 6. Genomic prediction accuracies from GBLUP averaged across 5 folds for AUS+USA+IRE reference.

RANKED1 AUS_USA_IRE	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.64 (0.02)	0.84 (0.02)	0.73 (0.02)	0.68 (0.01)	0.68 (0.02)
cor(res,gebv)/h	0.57 (0.03)	0.62 (0.01)	0.61 (0.02)	0.63 (0.01)	0.62 (0.02)
LR_bias	0.02 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.02)	0.01 (0.01)
LR_slope	1.04 (0.03)	1.04 (0.02)	1.07 (0.02)	1.06 (0.01)	1.05 (0.02)
LR_accuracy	0.83 (0.01)	0.89 (0.00)	0.89 (0.00)	0.87 (0.00)	0.86 (0.00)
RANKED2 AUS_USA_IRE	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.39 (0.18)	0.47 (0.27)	0.47 (0.27)	0.47 (0.22)	0.45 (0.22)
cor(res,gebv)/h	0.37 (0.15)	0.41 (0.20)	0.39 (0.20)	0.44 (0.17)	0.41 (0.17)
LR_bias	-0.13 (0.11)	-0.14 (0.18)	-0.15 (0.16)	-0.15 (0.14)	-0.13 (0.12)
LR_slope	1.12 (0.14)	1.26 (0.24)	1.26 (0.25)	1.22 (0.16)	1.17 (0.16)
LR_accuracy	0.59 (0.09)	0.66 (0.08)	0.64 (0.09)	0.65 (0.08)	0.62 (0.09)
YOBIRTH AUS_USA_IRE	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.65 (0.04)	0.84 (0.03)	0.74 (0.02)	0.69 (0.04)	0.69 (0.04)
cor(res,gebv)/h	0.58 (0.02)	0.62 (0.02)	0.62 (0.02)	0.64 (0.02)	0.63 (0.02)
LR_bias	0.01 (0.03)	0.00 (0.03)	-0.01 (0.03)	-0.01 (0.03)	0.00 (0.03)
LR_slope	1.05 (0.03)	1.04 (0.03)	1.07 (0.03)	1.06 (0.03)	1.05 (0.03)
LR_accuracy	0.83 (0.01)	0.89 (0.00)	0.89 (0.00)	0.87 (0.00)	0.86 (0.01)
RANDOM AUS_USA_IRE	**TENDER**	**JUICY**	**FLAVOR**	**OVERALL**	**MQ4**
cor(pheno,gebv)/h	0.64 (0.02)	0.83 (0.04)	0.74 (0.05)	0.68 (0.04)	0.68 (0.04)
cor(res,gebv)/h	0.58 (0.03)	0.62 (0.04)	0.61 (0.04)	0.64 (0.04)	0.63 (0.03)
LR_bias	0.02 (0.05)	0.00 (0.03)	-0.01 (0.02)	0.00 (0.04)	0.00 (0.04)
LR_slope	1.05 (0.04)	1.05 (0.04)	1.08 (0.04)	1.06 (0.04)	1.06 (0.04)
LR_accuracy	0.83 (0.00)	0.89 (0.00)	0.89 (0.00)	0.87 (0.00)	0.86 (0.00)

Table 7. Genomic prediction accuracies (correlation of GEBV and phenotype dived by square root of trait heritability) from BayseR averaged across 5 folds.

Trait	RANKED1		YOBIRT	
	AUS	AUS+USA+IRE	AUS	AUS+USA+IRE
TENDER	0.42 (0.07)	0.64 (0.03)	0.36 (0.07)	0.65 (0.05)
JUICY	0.58 (0.06)	0.59 (0.06)	0.60 (0.06)	0.60 (0.04)
FLAVOR	0.39 (0.03)	0.55 (0.03)	0.41 (0.04)	0.57 (0.06)
OVERALL	0.45 (0.03)	0.56 (0.01)	0.50 (0.07)	0.53 (0.07)
MQ4	0.41 (0.03)	0.64 (0.03)	0.39 (0.03)	0.64 (0.03)

4.4 Cross validation for USA and Ireland

Adding international data from the Australia and IRE to the USA reference population and likewise Australia and USA data to the IRE reference population increased the prediction accuracies for all traits across all validation strategies using BayesR models (Figure 3 and 4 respectively) when 2023 data was used as the validation. The addition of all data increased genomic prediction accuracies for the USA by 0.25 or more. Juiciness went from 0.23 to 0.5 while flavour went from 0.43 to 0.85 with the inclusion of Irish and Australian data. Meanwhile, the addition of all data increased genomic prediction accuracies for Ireland by 0.17 or more. Flavour went from 0.41 to 0.58 while tenderness went from 0.09 to 0.55 with the inclusion of USA and Australian data. This highlights the added value of a larger and more diverse reference population for all countries to achieve higher genomic predictions for all traits. These findings underscore the benefits of incorporating international data to enhance prediction accuracies for meat eating quality traits in all 3 countries involved in this project.

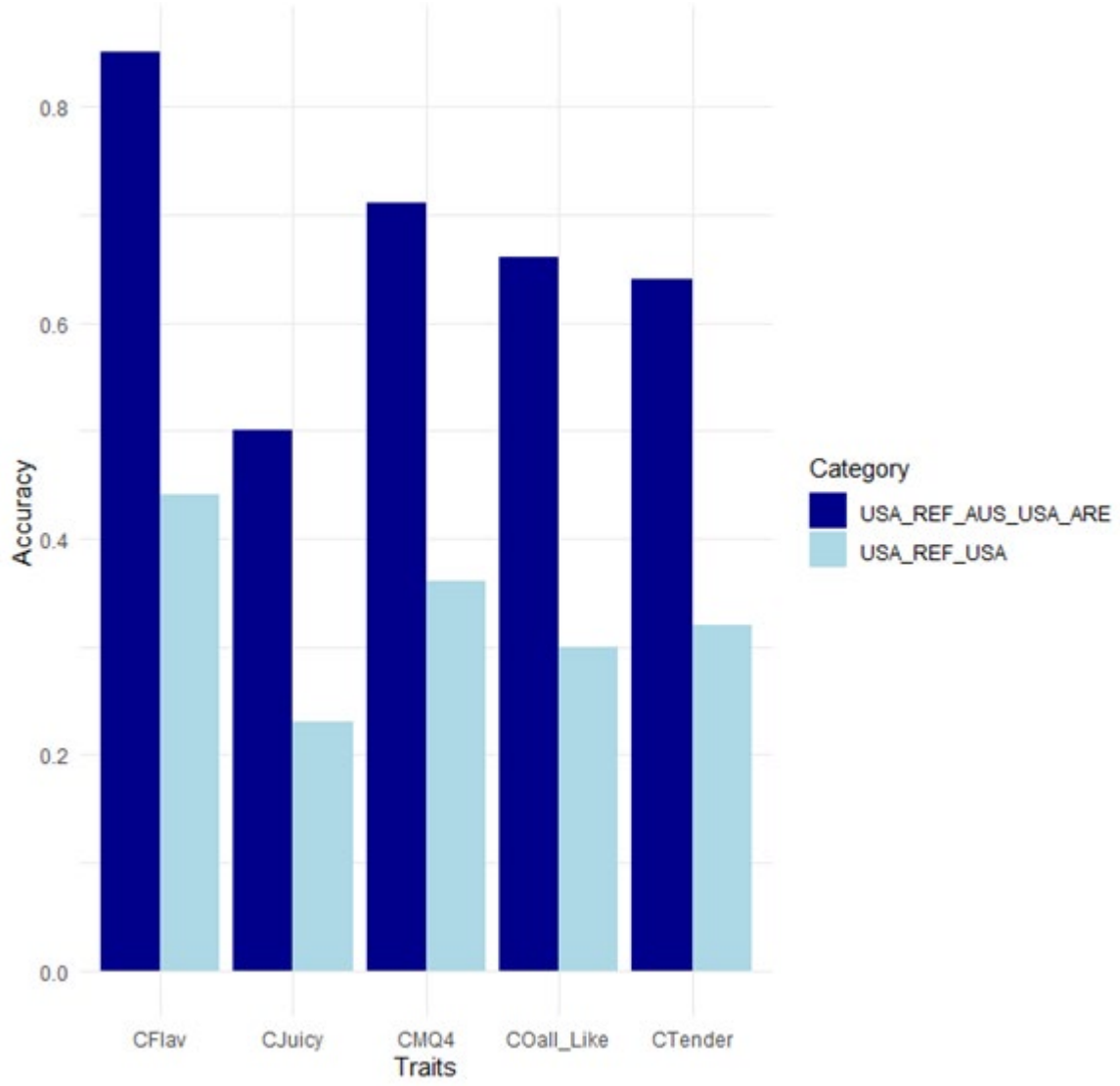

Figure 3: Cross-validation of USA genomic predictions from the BayesR analysis with (dark blue columns) and without (light blue columns) the Australian and Irish phenotypes

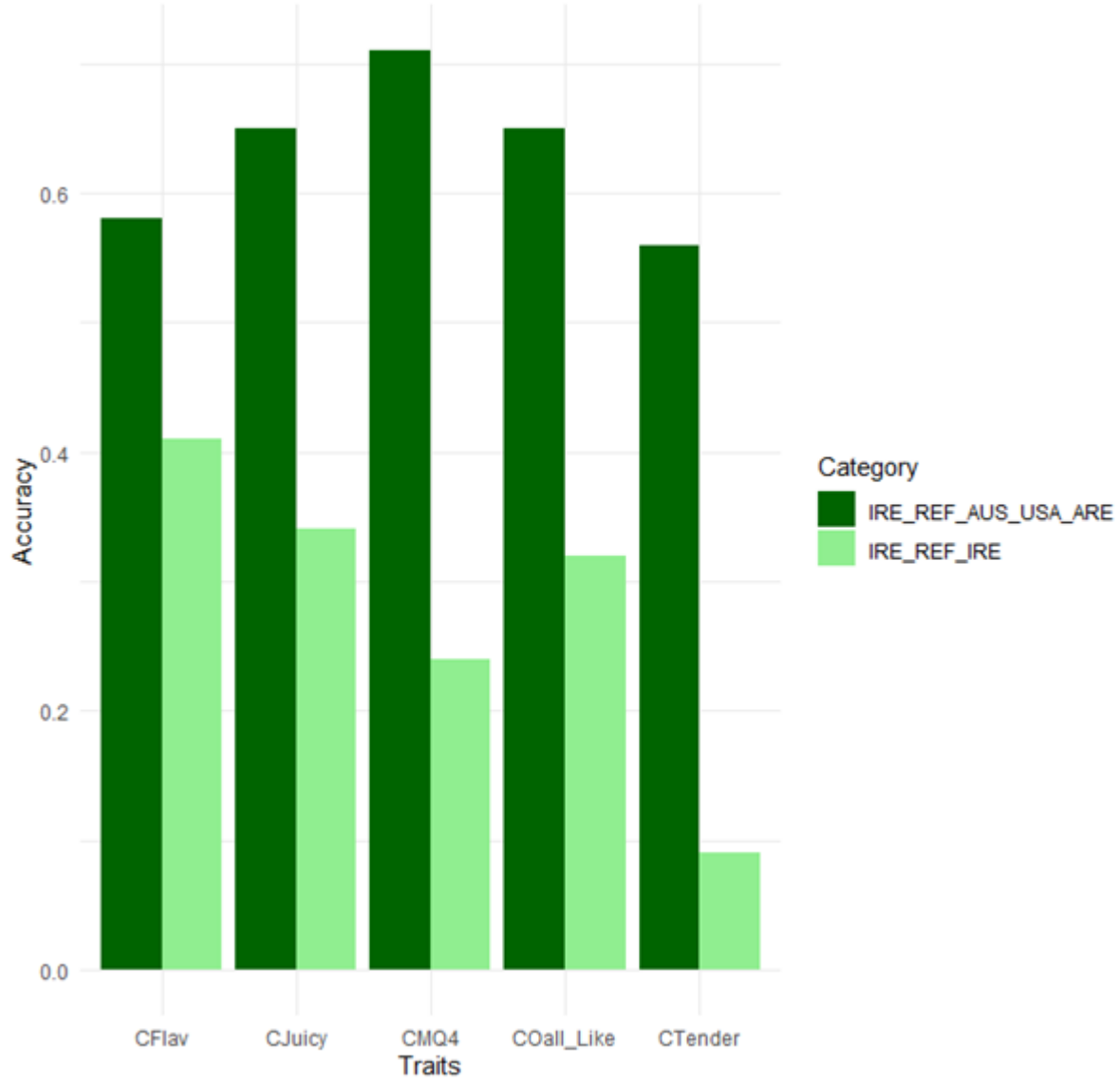

Figure 4: Cross-validation of Irish genomic predictions from the BayesR analysis with (dark blue columns) and without (light blue columns) the Australian and USA phenotypes

4.5 Bivariate models for estimations of genetic correlations

Table 8 contains the estimates of genetic correlations (± standard errors) between sensory and carcass eating quality traits, using bivariate GBLUP models. Between carcass eating quality traits, ossification had the lowest absolute values of genetic correlations with sensory eating quality traits and marbling had the highest genetic correlations with sensory eating quality traits. Among the carcass eating quality traits, ossification exhibited the lowest absolute values of genetic correlations with sensory eating quality traits, while marbling showed the highest genetic correlations with sensory eating quality traits. This highlights the stronger genetic association between marbling and sensory traits compared to other carcass eating quality traits. The signs of genetic correlations were positive for three carcass eating quality traits, indicating favourable genetic relationships with sensory eating quality traits. However, the genetic correlation for eye muscle area was negative, suggesting that a decrease in sensory eating quality traits is associated with an increase in eye

muscle area. The standard errors were quite large, reflecting the small sample size used in the estimation. However, it can be concluded that sensory and carcass eating quality traits are not genetically identical traits.

Table 8. The estimated genetic correlation (±SE) between sensory and carcass meat quality traits.

Trait	EMA	RIB	OSS	MARBL
CTender	-0.27 ± 0.10	0.15 ± 0.11	-0.05 ± 0.11	0.14 ± 0.11
CJuicy	-0.29 ± 0.14	0.13 ± 0.14	0.01 ± 0.15	0.47 ± 0.15
CFlav	-0.34 ± 0.13	0.16 ± 0.14	0.10 ± 0.14	0.55 ± 0.15
COall_Like	-0.31 ± 0.11	0.19 ± 0.12	0.05 ± 0.13	0.35 ± 0.12
CMQ4	-0.29 ± 0.10	0.15 ± 0.12	0.02 ± 0.12	0.32 ± 0.12

4.6 Accuracies of genomic predictions for specific breeds within a multi-breed analysis

To estimate genomic prediction accuracies for specific breeds within the multi-breed evaluation, we used the correlation between GEBVs and phenotypes as a proxy for accuracy.

We analysed a target population of 1,040 animals from the Southern Multi-Breed (SMB) project with known breed information. These animals belonged to two cohorts (R and S) and represented five temperate beef breeds.

The same statistical models as used before for multi-breed evaluation, incorporating the first four principal components (PCs), and GBLUP and BayseR methodologies were used. A two-fold cross-validation approach was performed: phenotypic records from one cohort were masked, while the other cohort was included in the reference population, alongside either the remaining animals from either AUS-only or AUS_USA_IRE datasets. Prediction accuracy for each breed listed in Table 9 was calculated.

Table 9. Number of Animals from different breeds in SMB R and S cohorts.

Breed	**Number of animals in cohort**	
	R	S
Angus	188	164
Charolais	52	68
Hereford	129	149
Shorthorn	73	94
Wagyu	91	32

Overall, the correlations were generally low —often close to zero or even negative —across all traits and breeds using both GBLUP and BayseR methodologies. Slightly higher accuracies were observed using the AUS_USA_IRE reference compared to the AUS-only reference. GEBV distributions showed strong breed clustering: Charolais and Shorthorn tended to have the lowest GEBVs, Wagyu the highest, and other breeds fell in between. Within-breed GEBVs showed a narrow spread and clustered along a linear trend, explaining the low breed-specific cross-validation accuracies. These patterns suggest that current model do not fully account for breed effects.

To determine whether including additional PCs in the model could mitigate breed effects and improve prediction accuracies, we tested models incorporating the first 15 or 20 PCs. While this resulted in marginal improvements, within-breed accuracies remained substantially lower than those from multi-breed evaluations. Across most breeds, prediction accuracy was consistently higher for the R cohort than for the S cohort, regardless of the number of PCs used. MDS plots suggest that this difference may stem from better alignment between the R cohort and the reference population.

To further explore whether a more closely related reference population could improve prediction accuracy, we conducted focused analyses on Angus animals using several scenarios.

1) **Similarity-based reference selection (average similarity > 0.01):** Angus animals were used in a two-fold cross-validation, with reference animals preselected based on an average genomic similarity threshold (>0.01) to the target fold. BayesR methodology was used for the prediction.
2) **Minimum similarity threshold (≥ 0.25):** Angus animals were used in a two-fold cross-validation, with reference animals preselected based on having at least 0.25 similarity to any target fold. BayesR methodology was used for the prediction.
3) **Cluster-based reference (k-means on PCs):** The full population (AUS, USA, and IRE) was clustered into 10 groups using k-means based on the first 10 PCs. A total of 352 Angus animals, all in the same cluster, were split into two folds for cross-validation, with the remaining animals from that cluster used as the reference population. Two scenarios were tested—using only AUS animals and using all three populations from the cluster as reference population. BayesR methodology was used for the prediction.
4) **PC-defined reference sets:** Reference groups were defined using PC1 and PC2 to exclude indicus and non-Angus taurine breeds. This resulted in two reference groups: Angus and Taurus. Subsets were further split into Australia-only (Angus_AUS and Taurus_AUS) and international (Angus_all and Taurus_all) groups. The target group in all cases consisted of SMB Angus animals (cohorts R and S). GBLUP methodology was used for the prediction.

Across all scenarios, correlations remained low, consistent with previous breed-specific cross-validations. However, better results were observed when using the Cluster-based reference that included international data (Figure 5). In this scenario, prediction accuracy —measured as the correlation between GEBVs and phenotypes divided by the square root of heritability—ranged from 0.13 to 0.27 for tenderness and flavour.

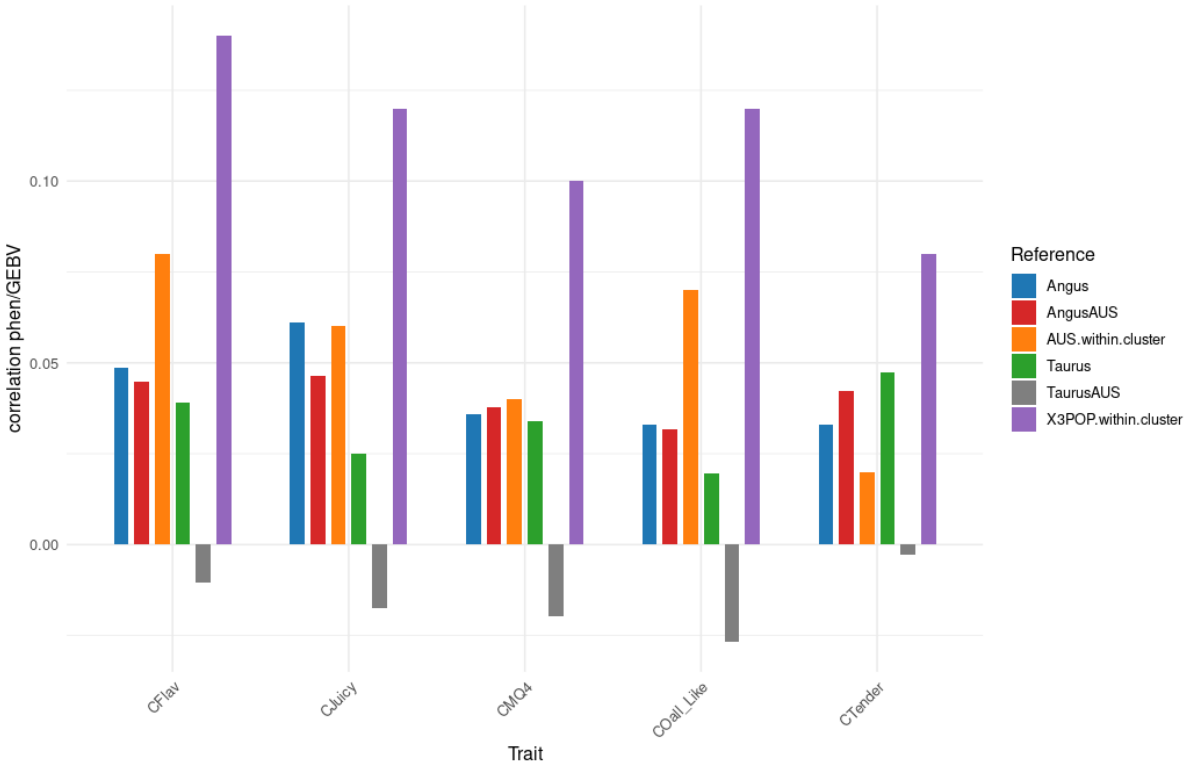

Figure 5. Average correlation between GEBVs and phenotypes for Angus animals across two-fold cross-validation. Results are presented for various reference population scenarios, including PC-defined sets—using only Australian animals (Angus_AUS and Taurus_AUS) or all international animals (Angus and Taurus)—and cluster-based references using either Australian animals within the cluster (AUS.within.cluster) or all international animals within the cluster (X3POP.within.cluster).

In conclusion, despite testing multiple reference populations, breed-specific prediction accuracies remained low. This consistent pattern across scenarios reinforces the importance of increasing reference data for key breeds to better capture the haplotype diversity unique to each breed.

4.7 Implications for reference population design:

The comparison between across-breed and within-breed validation provided insights for reference population design. The across-breed validation had a relatively high accuracy because it effectively utilised genetic variation from both across-breed and crossbred animals. In contrast, within-breed cross-validation had lower accuracy due to small numbers of pure-bred animals and a diverse number of breed crosses in the reference population. This difference has important implications for how we structure reference populations for genetic prediction.

For the prediction of breeding value within breeds reference populations need be carefully designed with breed-specific considerations in mind. Each breed has specific reference population needs to achieve optimal prediction accuracy. However, the relationship between purebred and crossbred populations offers valuable opportunities for mutual information exchange. Crossbreds can be predicted from purebred data and vice versa, provided the parent breeds are adequately represented in the dataset. Similarly, crossbred animals contribute valuable information about their parent purebreds, creating a two-way flow of genetic information that enhances overall prediction

capabilities. The number of effective records for prediction depends on the breed content of the cross for pure bred prediction. For example, a simplified comparison could be a reference population for predicting a single breed could be composed of 10,000 animals recorded for the traits of interest from that breed. If the reference composed of first cross animals with another breed (i.e. breed A x B) then double the number of animals (20,000) would potentially be required to have the same number of haplotypes represented for the within breed prediction. In general, combining data from different breeds is yet to result in increases in accuracy for within breed selection unless breeds are closely related or composites including the same breeds.

The effectiveness of any reference population depends on four interconnected factors: population size, genetic diversity, relatedness to target animals, and trait heritability. When animals in the reference population are more closely related to the prediction targets, accuracy increases substantially. This relationship between genetic closeness and prediction accuracy forms the foundation for strategic population design. The value of a reference population can be systematically assessed through three key questions: how many records are available, how much each phenotypic record reveals about breeding value, and how genetically relevant the measured animals are to the prediction targets.

5. Conclusion

5.1 Key findings

- The heritabilities of all 5 eating quality traits (Tenderness, Juiciness, Flavour, Overall liking and MQ4) are moderately heritable traits
- There is significant variation between animals in the 5 eating quality traits, hence there is merit in selection for improvements in eating quality
- For the AUS-only reference, juiciness had the highest prediction at 0.69, while the accuracies for the remaining four traits were similar, ranging between 0.62 and 0.65.
- The accuracies from the AUS+USA+IRE reference population showed notable improvements compared to the AUS-only reference, increasing by 0.21 for tenderness (from 0.62 to 0.83), by 0.20 for juiciness (from 0.69 to 0.89), by 0.25 for flavour (from 0.64 to 0.89), by 0.23 for overall liking (from 0.64 to 0.87), and by 0.23 for MQ4 (from 0.63 to 0.86). This indicates that combining international data gives much greater accuracy of prediction for all 4 eating quality traits.

5.2 Benefits to industry

Combining international data gives much greater accuracy of prediction for all 5 eating quality traits. This benefit was not only limited to Australia but was seen across the 3 populations, strongly demonstrating the benefit of combining forces to improve the genomic prediction accuracy of hard to measure, expensive traits like tenderness, juiciness, flavour, overall liking and MQ4. The roll out of an eating quality breeding value will generate improvements in the eating quality of beef internationally which shows significant variation.

6 Future Research & Recommendations

- Genomic predictions for eating quality are demonstrating high enough heritability and accuracy to explore commercial multi-breed products for delivery through BREEDPLAN or other genetic analyses.
- Careful consideration is required into the design of future reference populations to ensure they are developed to predict the future target animals of interest, whilst also utilising data from animals consumer tested for other projects.
- The multibreed prediction has low predictive accuracy for within breeds. Therefore, the application of the multibreed algorithm developed in this project is not suitable for use within purebred populations. The international partners (USDA and ICBF) will have access to the prediction equation developed via commercial licences.
- Permitted use of the prediction equation will be for the purpose of generating genomic breeding value products on the breeds and breed types involved in the project. Its use will be restricted to providing a service within country.
- Modelling is required to understand the numbers of eating quality records that need to be collected by each country annually, biannually or tri-annually to maintain constant improvement of the genomic prediction for eating quality.
- Understand the appetite from USDA and ICBF for further collaboration, further collection of sensory phenotypes. There is opportunity to establish a collaboration based on routine data collection and routine updates to the equation using the numbers modelled from the point above. This would be arranged as an update to the licence prepared for the above or under a separate licence.
- Combining data from multiple countries improved prediction accuracy. This should be investigated for other key hard to measure and expensive to record traits.
- Combining data is very beneficial for all countries that contributed. The expense to generate enough eating quality data individually is cost prohibitive to the development of a breeding value.
- Generate an MSA database that is not excel based to allow phenotypes and genotypes to be stored.

7 References

Aguilar I, Tsuruta S, Masuda Y, Lourenco D, Legarra A, Misztal, I. BLUPF90 suite of programs for animal breeding with focus on genomics. 10th World Congr Genet Appl Livest Prod 2018.

Clark SA, Hickey JM, Daetwyler HD, van der Werf JH. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. Genet Select Evol 2012, 44(1), 1-9.

Erbe M, Hayes BJ, Matukumalli LK, Goswami S, Bowman PJ, Reich CM, Mason BA, Goddard ME. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J Dairy Sci. 2012;95(7):4114-29. doi: 10.3168/jds.2011-5019.

Gee, A., Porter, M., Coffey, D., Polkinghorne, R. (1998). Design and Protocol for Steak Grilling Trials

Hayes BJ, Bowman PJ, Chamberlain AC, Verbyla K, Goddard ME. Accuracy of genomic breeding values in multi-breed dairy cattle populations. Genet Select Evol 2009, 41(1), 1-9.

Massart J, Sjögren RJO, Egan B, Garde C, Lindgren M, Gu W, Ferreira DMS, Katayama M, Ruas JL, Barrès R, O'Gorman DJ, Zierath JR, Krook A. Endurance exercise training-responsive miR-19b-3p improves skeletal muscle glucose metabolism. Nat Commun. 2021 Oct 12;12(1):5948. doi: 10.1038/s41467-021-26095-0.

Moser G, Lee SH, Hayes BJ, Goddard ME, Wray NR, Visscher PM. Simultaneous discovery, estimation and prediction analysis of complex traits using a Bayesian mixture model. PLoS Genet 2015, 11(4), e1004969.

Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011, 7;88(1):76-82. doi: 10.1016/j.ajhg.2010.11.011.

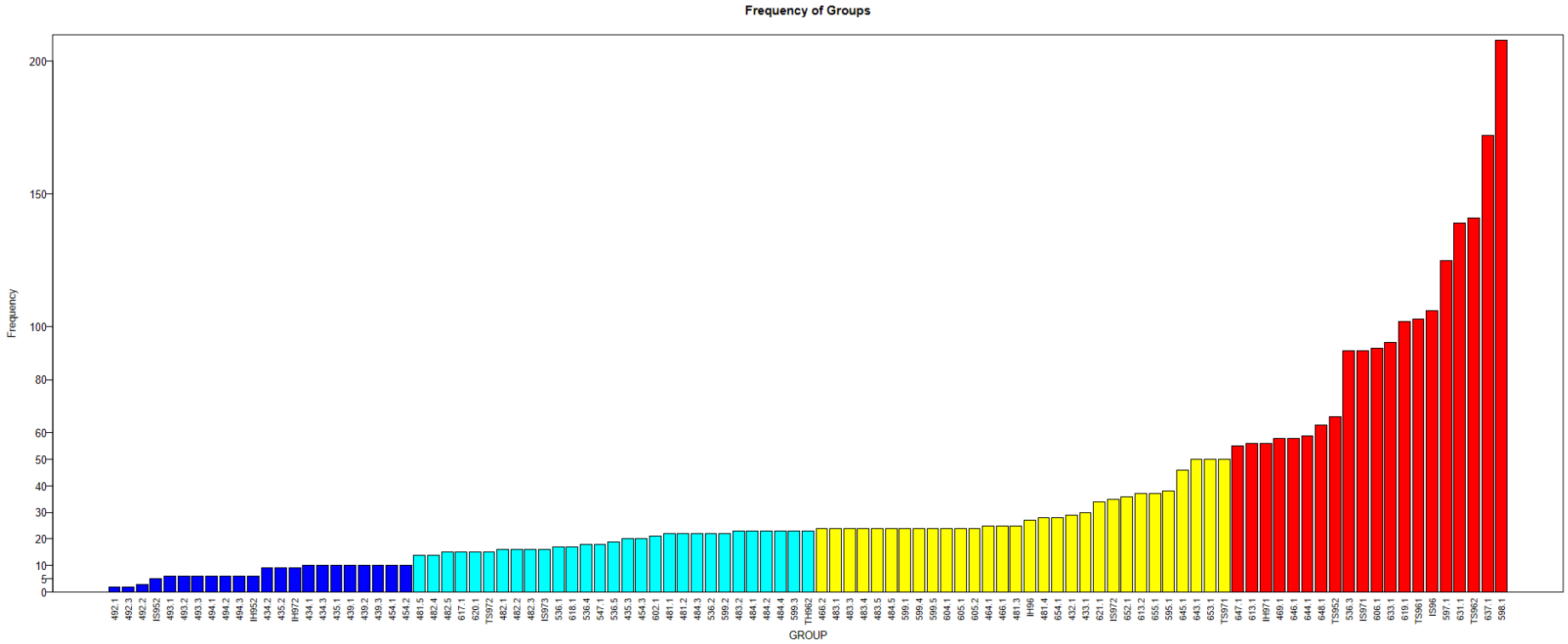
Frequency of Groups
Frequency
200
150
100
80
60
50
40
30
20
10
5
0
GROUP
492.1
492.3
492.2
IS952
493.1
493.2
493.3
494.1
494.2
494.3
IH952
434.2
435.2
IH972
434.1
434.3
435.1
439.1
439.2
439.3
454.1
454.2
481.5
482.4
482.5
617.1
620.1
TS972
482.1
482.2
482.3
IS973
536.1
618.1
536.4
547.1
536.5
435.3
454.3
602.1
481.1
481.2
484.3
536.2
599.2
483.2
484.1
484.2
484.4
599.3
TH962
466.2
483.1
483.3
483.4
483.5
484.5
599.1
599.4
599.5
604.1
605.1
605.2
464.1
466.1
481.3
IH96
481.4
654.1
432.1
433.1
621.1
IS972
652.1
613.2
655.1
595.1
645.1
643.1
653.1
TS971
647.1
613.1
IH971
469.1
646.1
644.1
648.1
TS952
536.3
IS971
606.1
633.1
619.1
TS961
IS96
597.1
631.1
TS962
637.1
598.1

Fig 1.

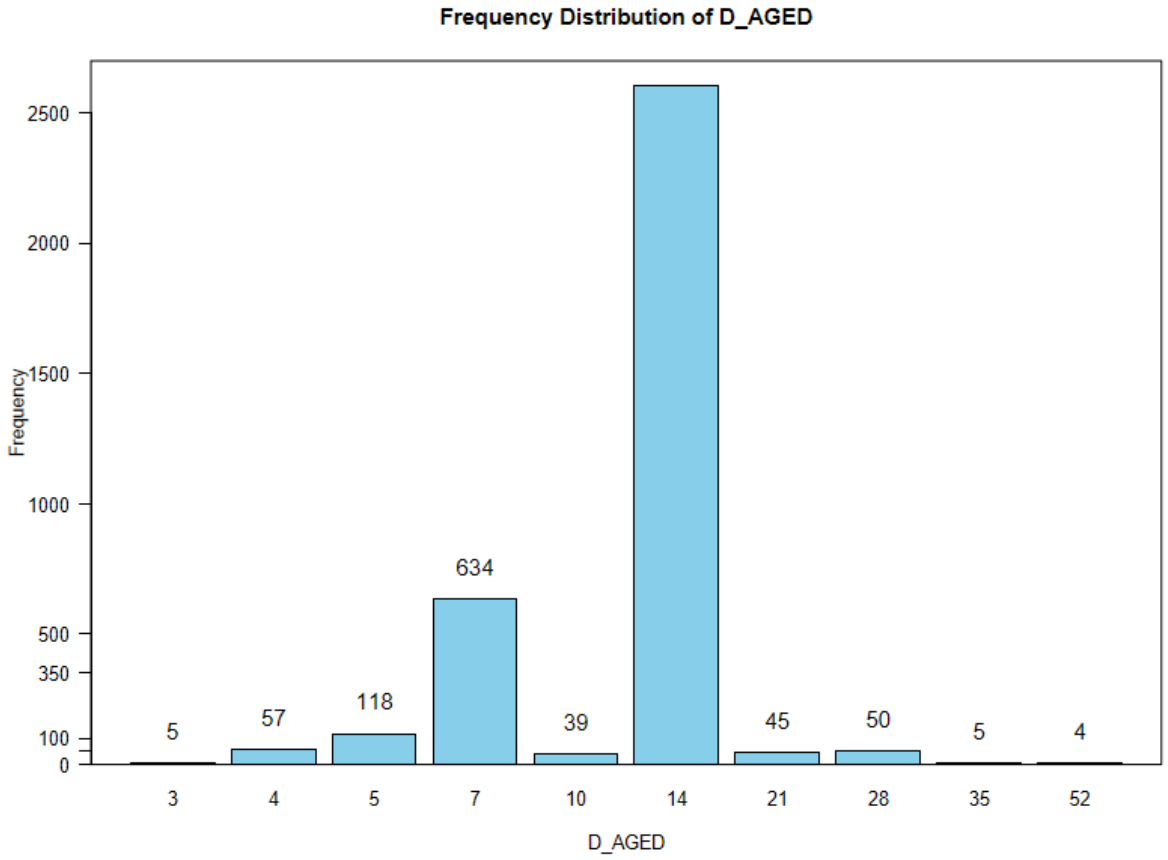
Frequency Distribution of D_AGED
Frequency
2500
2000
1500
1000
500
350
100
0
5
57
118
634
39
45
50
5
4
3
4
5
7
10
14
21
28
35
52
D_AGED

Fig 2.

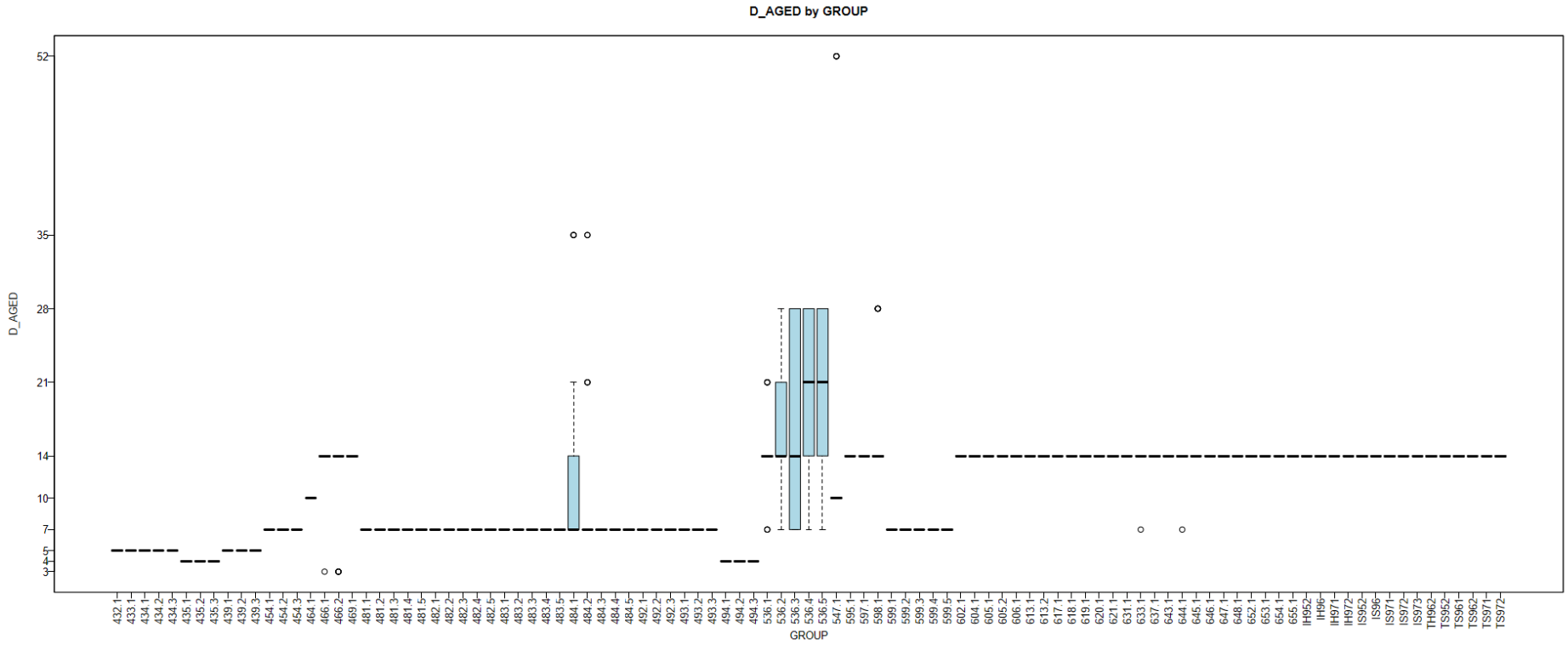

Fig 3.

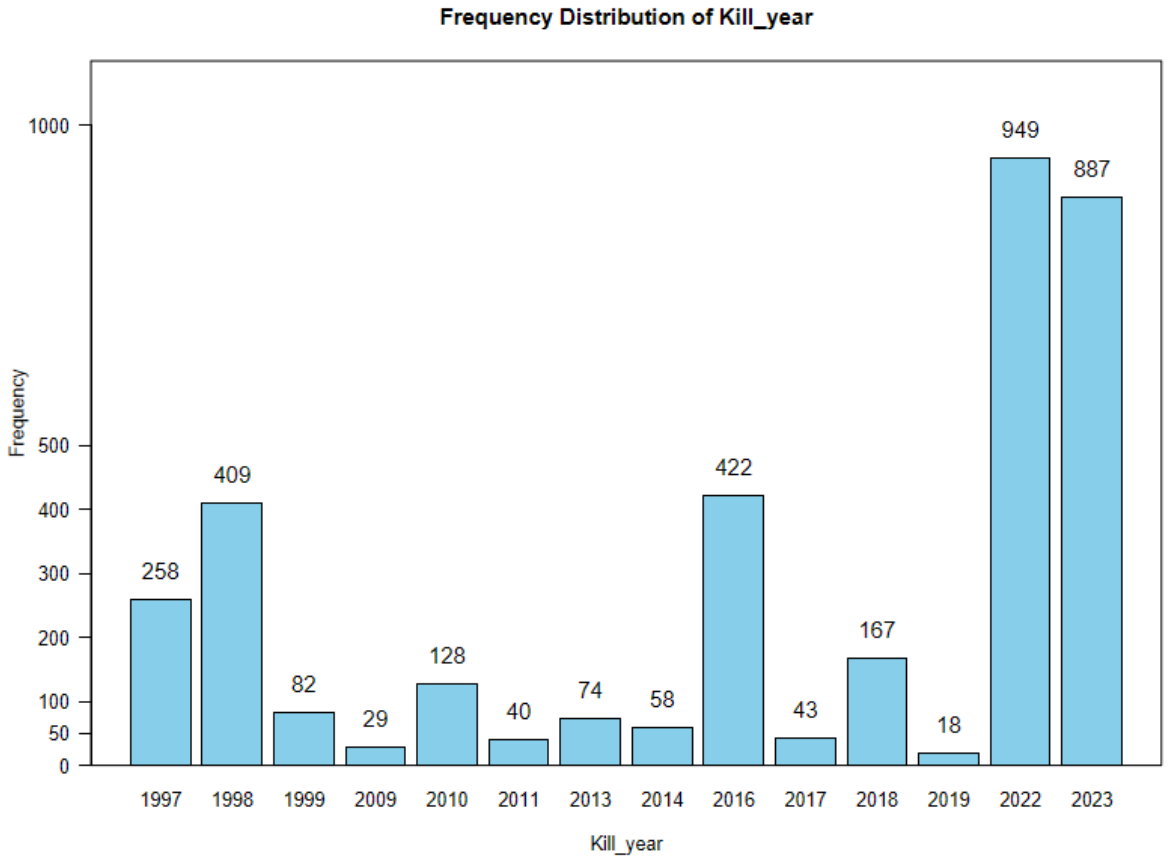
Frequency Distribution of Kill_year
Frequency
1000
500
400
300
200
100
50
0
258
409
82
29
128
40
74
58
422
43
167
18
949
887
1997
1998
1999
2009
2010
2011
2013
2014
2016
2017
2018
2019
2022
2023
Kill_year

Fig 4.

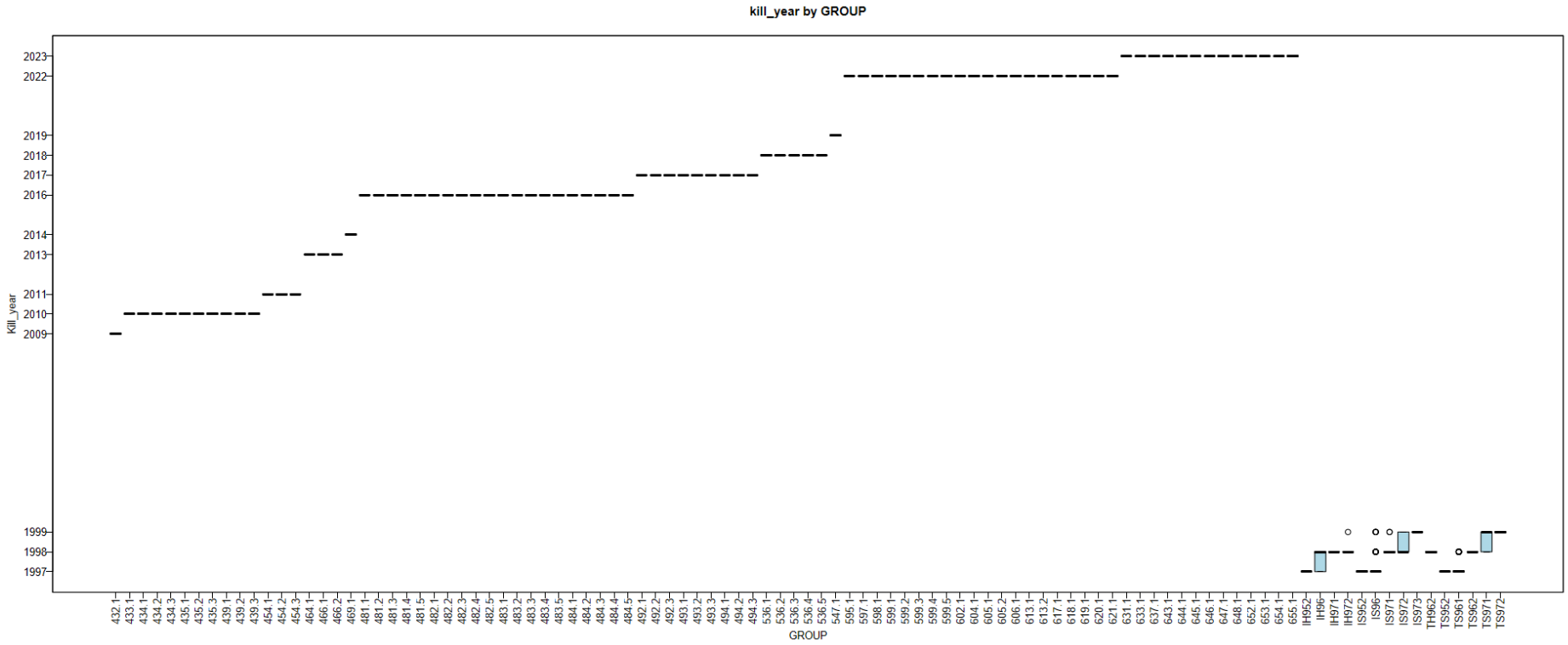
kill_year by GROUP
Kill_year
2023
2022
2019
2018
2017
2016
2014
2013
2011
2010
2009
1999
1998
1997
GROUP
432 1
433 1
434 1
434 2
434 3
435 1
435 2
435 3
439 1
439 2
439 3
454 1
454 2
454 3
464 1
466 1
466 2
469 1
481 1
481 2
481 3
481 4
481 5
482 1
482 2
482 3
482 4
482 5
483 1
483 2
483 3
483 4
483 5
484 1
484 2
484 3
484 4
484 5
492 1
492 2
492 3
493 1
493 2
493 3
494 1
494 2
494 3
536 1
536 2
536 3
536 4
536 5
547 1
595 1
597 1
598 1
599 1
599 2
599 3
599 4
599 5
602 1
604 1
605 1
605 2
606 1
613 1
613 2
617 1
618 1
619 1
620 1
621 1
631 1
633 1
637 1
643 1
644 1
645 1
646 1
647 1
648 1
652 1
653 1
654 1
655 1
IH952
IH96
IH971
IH972
IS952
IS96
IS971
IS972
IS973
TH962
TS952
TS961
TS962
TS971
TS972

Fig 5.

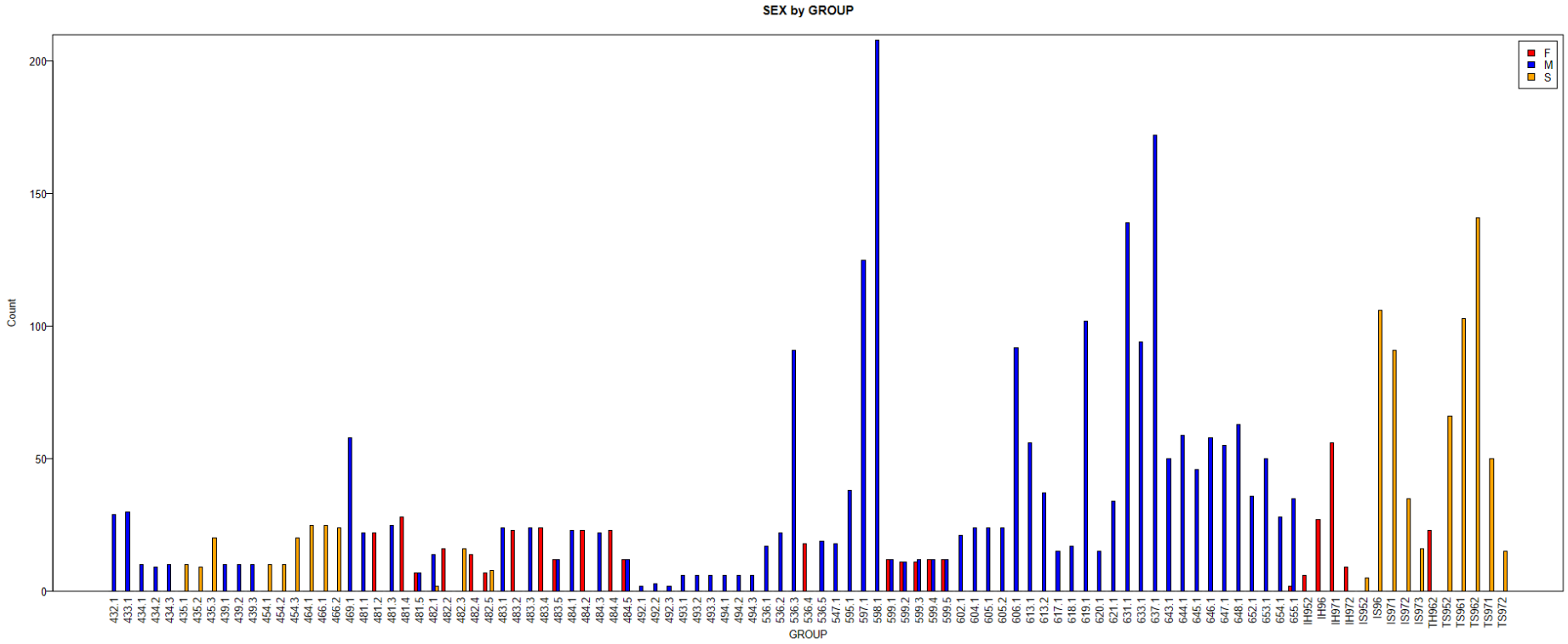
SEX by GROUP
F
M
S
Count
0
50
100
150
200
GROUP
432 1
433 1
434 1
434 2
434 3
435 1
435 2
435 3
439 1
439 2
439 3
454 1
454 2
454 3
464 1
466 1
466 2
469 1
481 1
481 2
481 3
481 4
481 5
482 1
482 2
482 3
482 4
482 5
483 1
483 2
483 3
483 4
483 5
484 1
484 2
484 3
484 4
484 5
492 1
492 2
492 3
493 1
493 2
493 3
494 1
494 2
494 3
536 1
536 2
536 3
536 4
536 5
547 1
595 1
597 1
598 1
599 1
599 2
599 3
599 4
599 5
602 1
604 1
605 1
605 2
606 1
613 1
613 2
617 1
618 1
619 1
620 1
621 1
631 1
633 1
637 1
643 1
644 1
645 1
646 1
647 1
648 1
652 1
653 1
654 1
655 1
IH952
IH96
IH971
IH972
IS952
IS96
IS971
IS972
IS973
TH962
TS952
TS961
TS962
TS971
TS972

Fig 6.

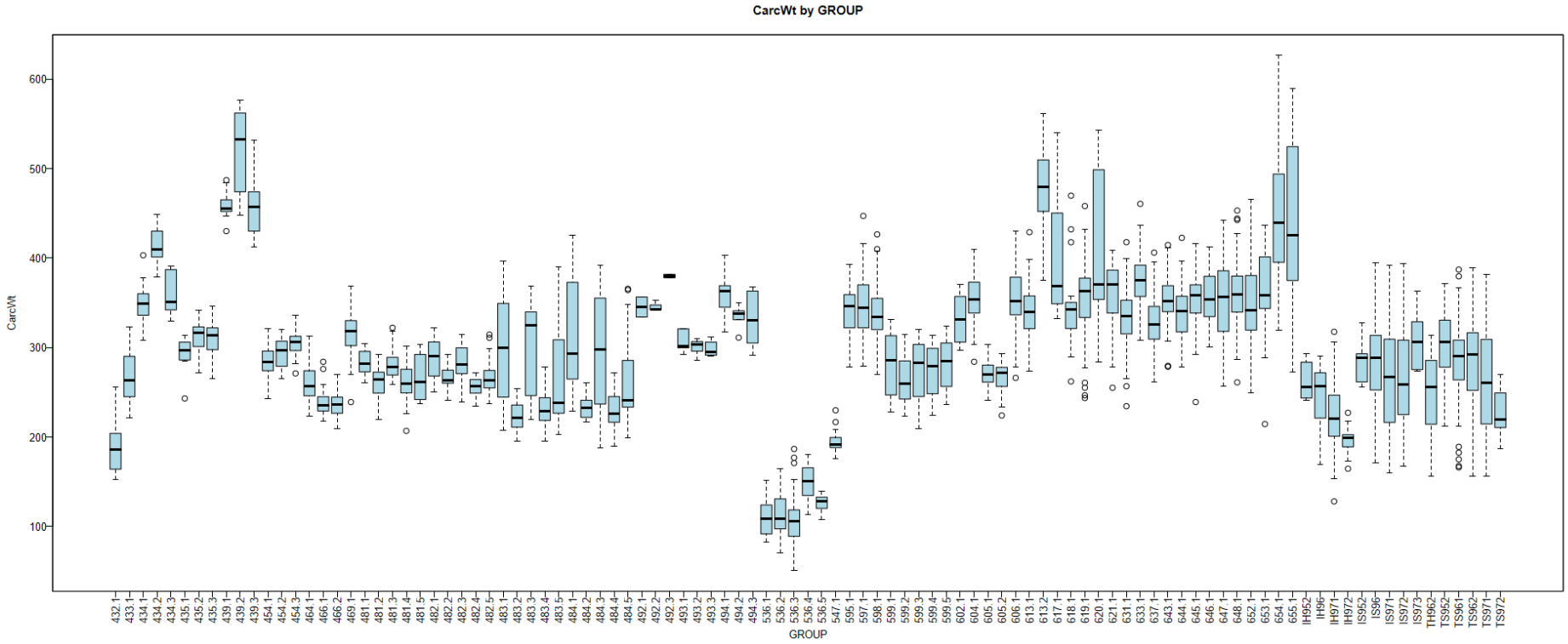
CarcWt by GROUP
CarcWt
600
500
400
300
200
100
GROUP
432 1
433 1
434 1
434 2
434 3
435 1
435 2
435 3
439 1
439 2
439 3
454 1
454 2
454 3
464 1
466 1
466 2
469 1
481 1
481 2
481 3
481 4
481 5
482 1
482 2
482 3
482 4
482 5
483 1
483 2
483 3
483 4
483 5
484 1
484 2
484 3
484 4
484 5
492 1
492 2
492 3
493 1
493 2
493 3
494 1
494 2
494 3
536 1
536 2
536 3
536 4
536 5
547 1
595 1
597 1
598 1
599 1
599 2
599 3
599 4
599 5
602 1
604 1
605 1
605 2
606 1
613 1
613 2
617 1
618 1
619 1
620 1
621 1
631 1
633 1
637 1
643 1
644 1
645 1
646 1
647 1
648 1
652 1
653 1
654 1
655 1
IH952
IH96
IH971
IH972
IS952
IS96
IS971
IS972
IS973
TH962
TS952
TS961
TS962
TS971
TS972

Fig 7.

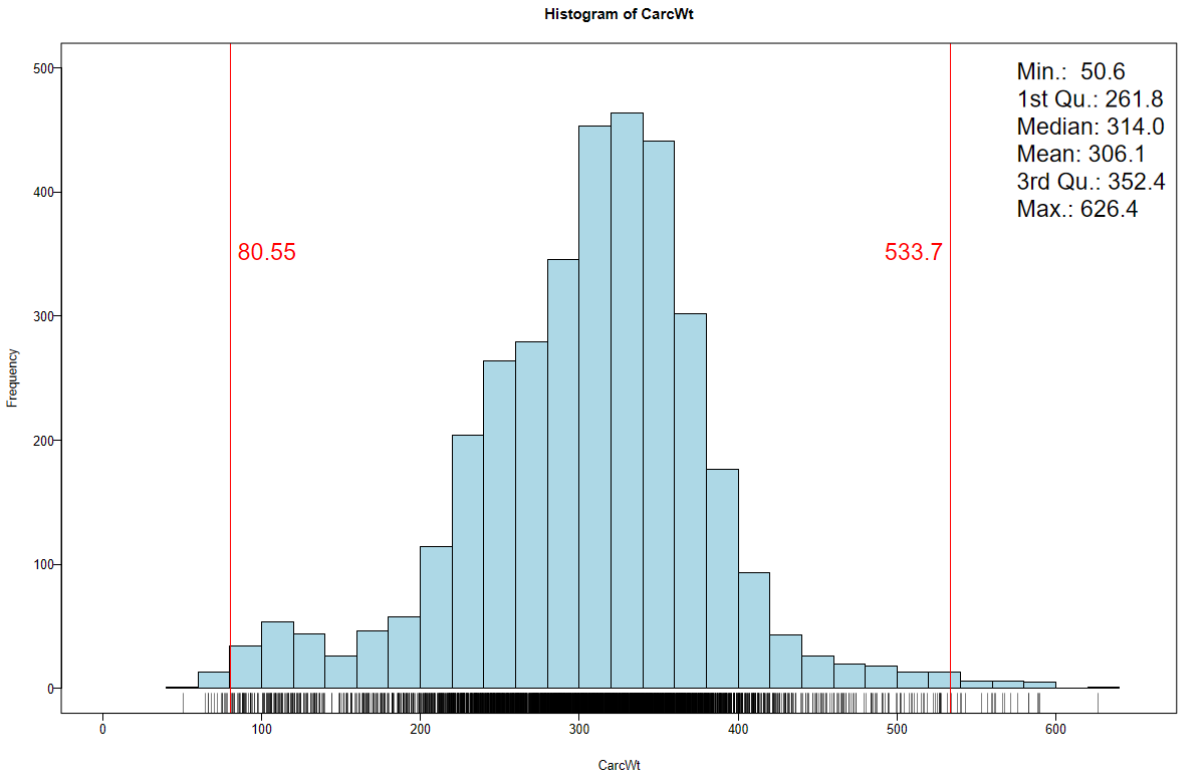
Histogram of CarcWt
Min.: 50.6
1st Qu.: 261.8
Median: 314.0
Mean: 306.1
3rd Qu.: 352.4
Max.: 626.4
80.55
533.7
Frequency
CarcWt

Fig 8.

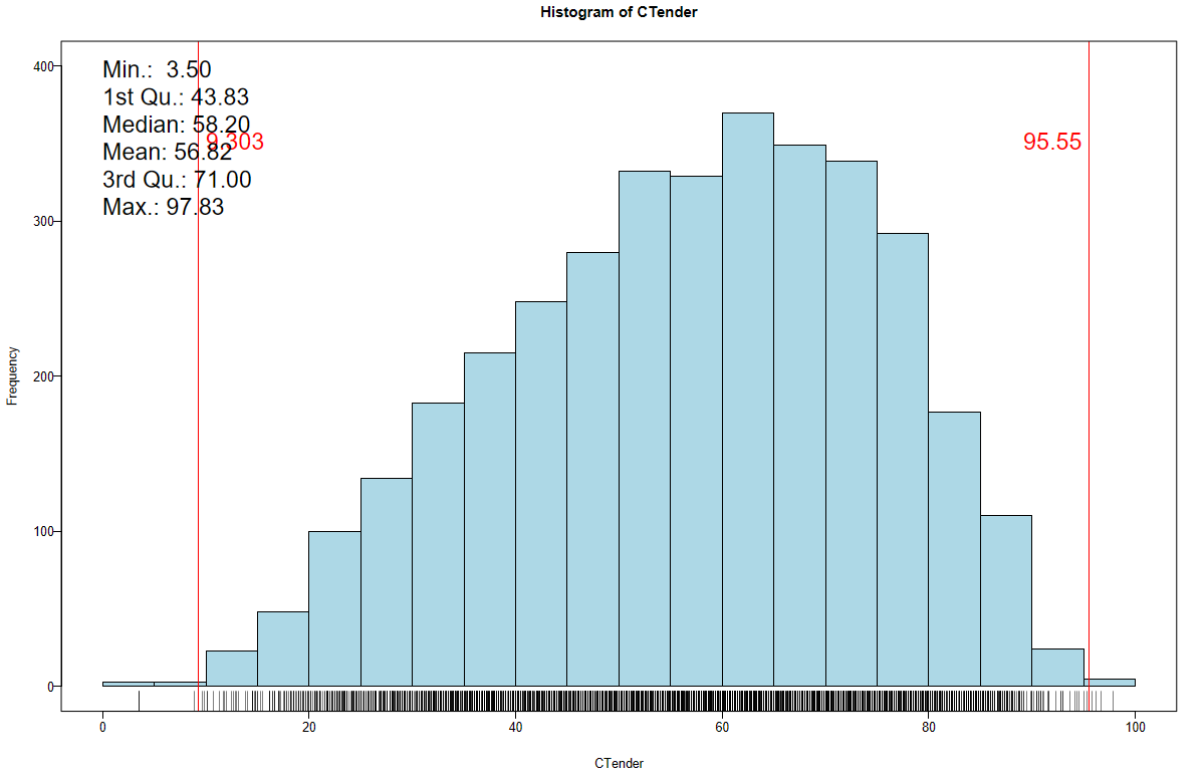
Histogram of CTender
Min.: 3.50
1st Qu.: 43.83
Median: 58.20
Mean: 56.82
3rd Qu.: 71.00
Max.: 97.83
9.303
95.55
Frequency
CTender

Fig 9.

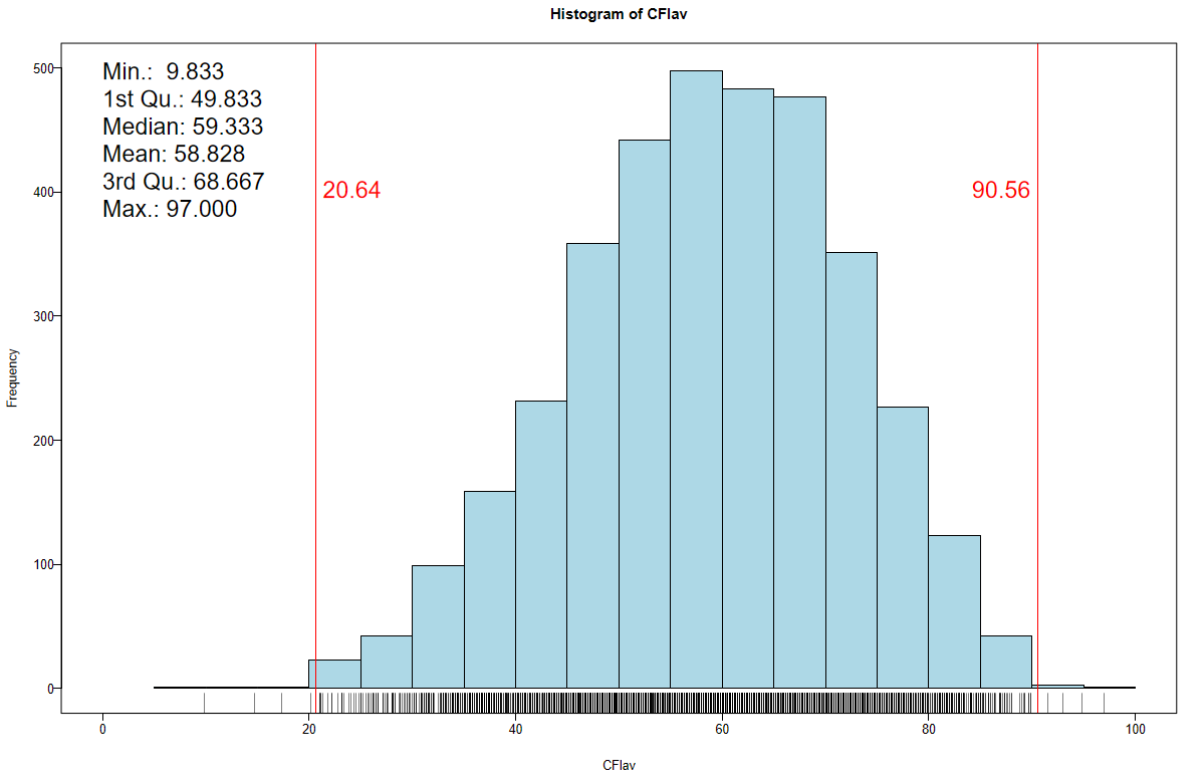
Histogram of CFlav
Min.: 9.833
1st Qu.: 49.833
Median: 59.333
Mean: 58.828
3rd Qu.: 68.667
Max.: 97.000
20.64
90.56
Frequency
CFlav

Fig 10.

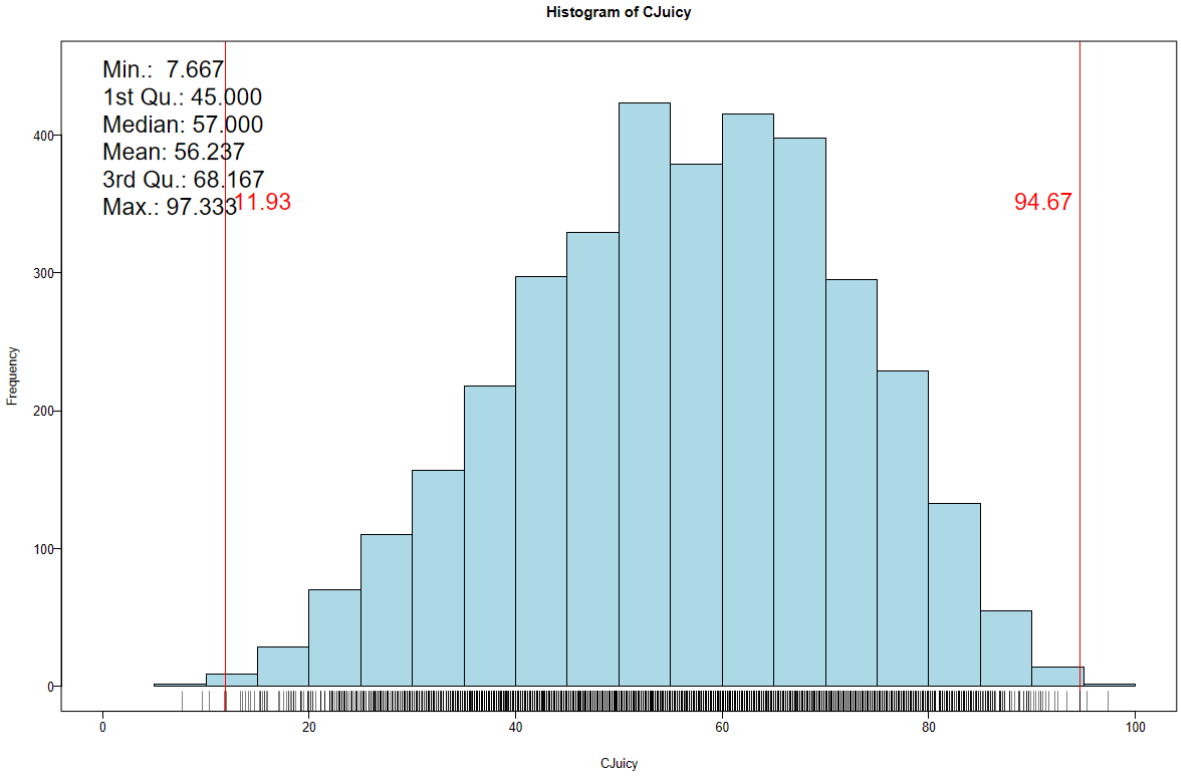
Histogram of CJuicy
Min.: 7.667
1st Qu.: 45.000
Median: 57.000
Mean: 56.237
3rd Qu.: 68.167
Max.: 97.333
11.93
94.67
Frequency
CJuicy

Fig 11.

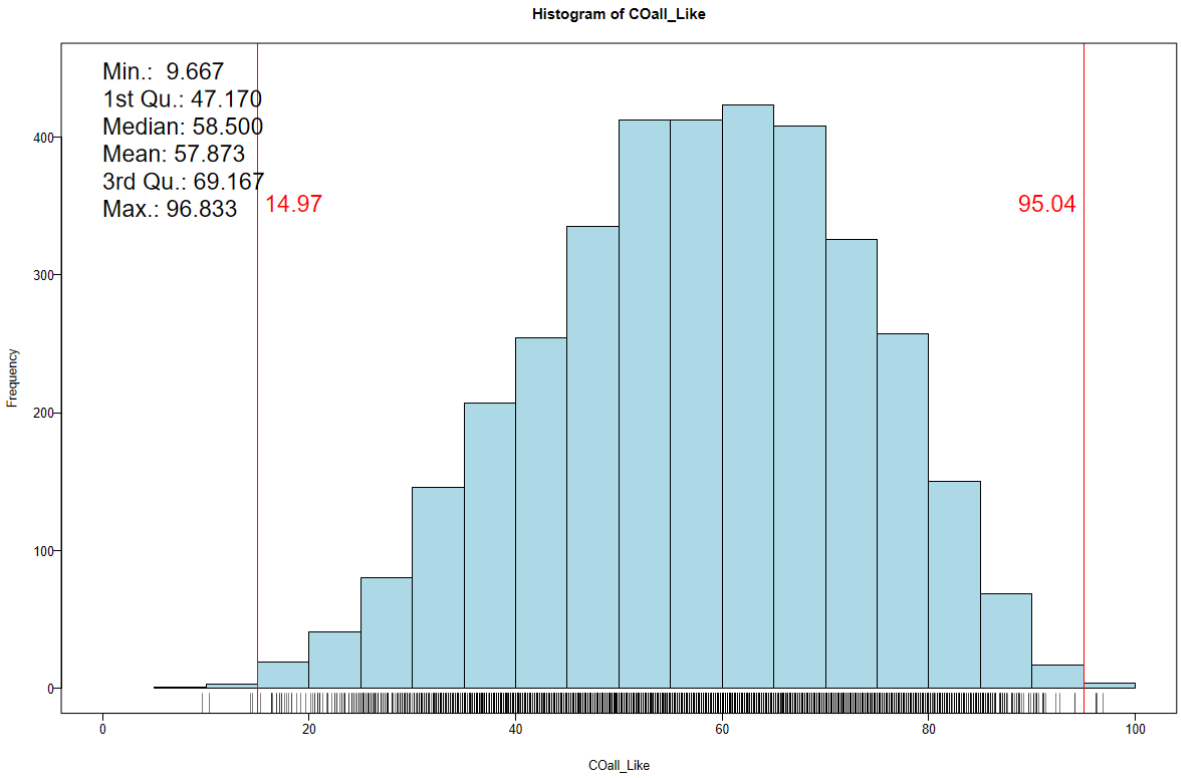
Histogram of COall_Like
Min.: 9.667
1st Qu.: 47.170
Median: 58.500
Mean: 57.873
3rd Qu.: 69.167
Max.: 96.833
14.97
95.04
Frequency
COall_Like

Fig 12.

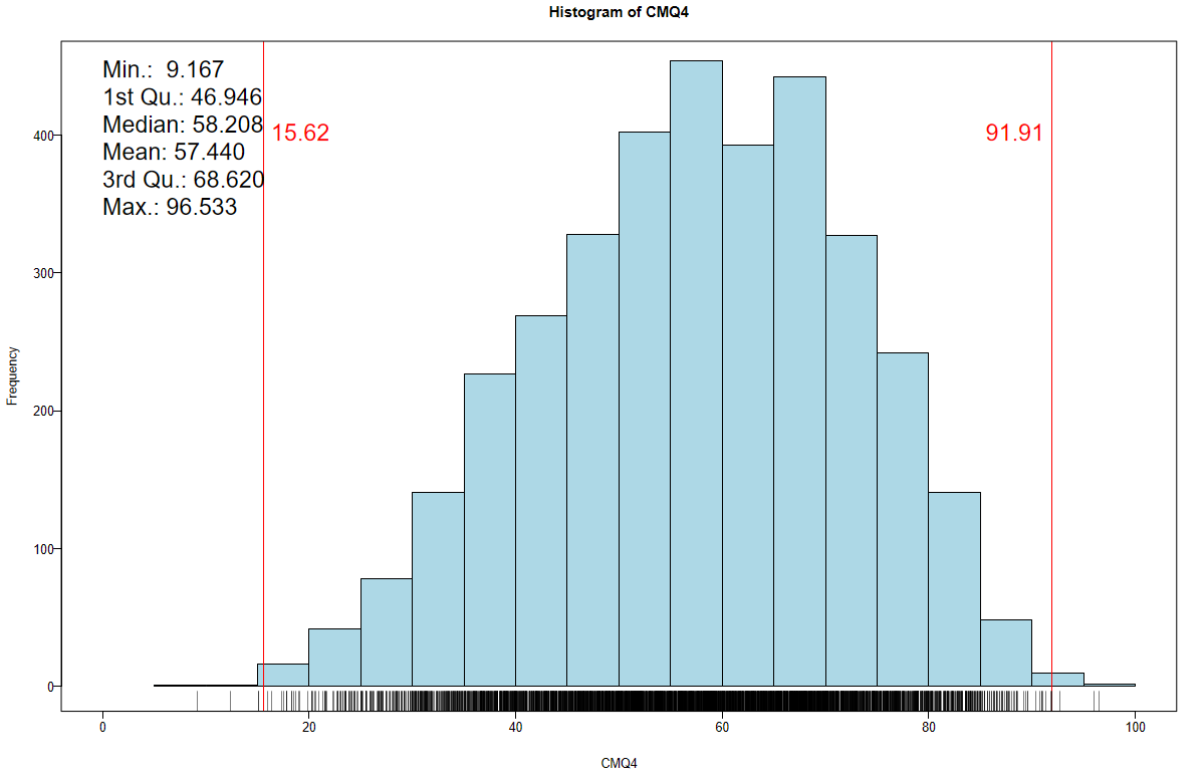
Histogram of CMQ4
Min.: 9.167
1st Qu.: 46.946
Median: 58.208
Mean: 57.440
3rd Qu.: 68.620
Max.: 96.533
15.62
91.91
Frequency
CMQ4

Fig 13.

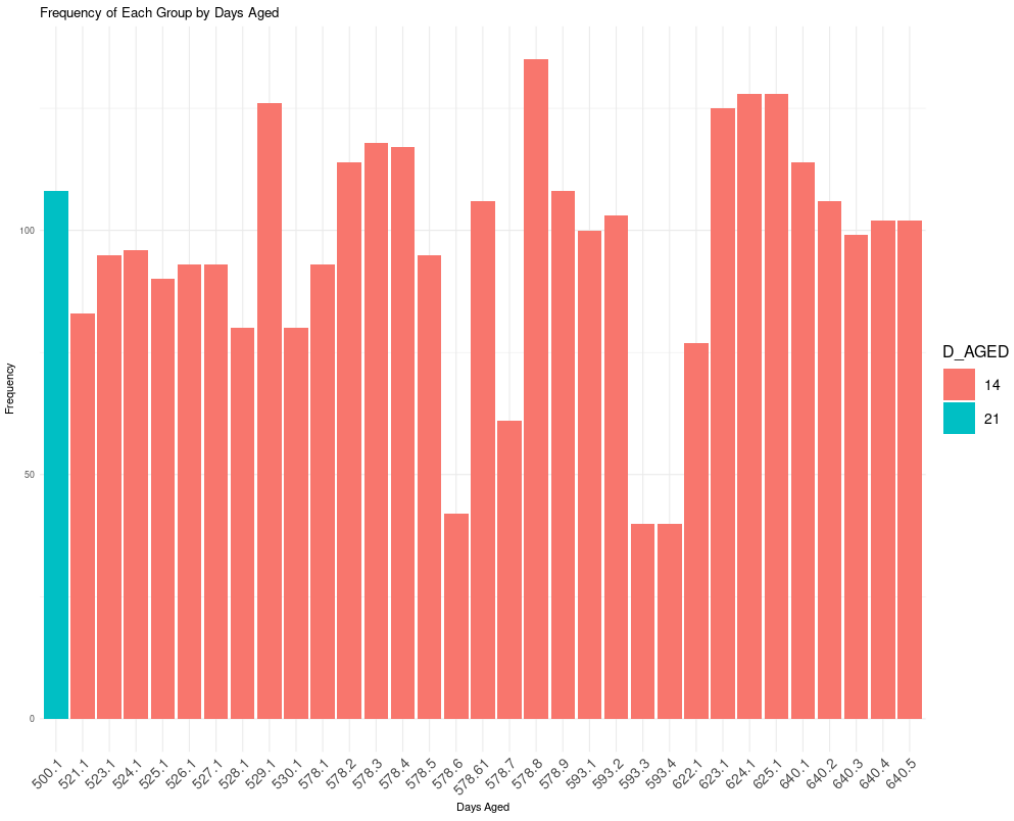
Frequency of Each Group by Days Aged
Frequency
100
50
0
500.1
521.1
523.1
524.1
525.1
526.1
527.1
528.1
529.1
530.1
578.1
578.2
578.3
578.4
578.5
578.6
578.61
578.7
578.8
578.9
593.1
593.2
593.3
593.4
622.1
623.1
624.1
625.1
640.1
640.2
640.3
640.4
640.5
Days Aged
D_AGED
14
21

Fig. 14

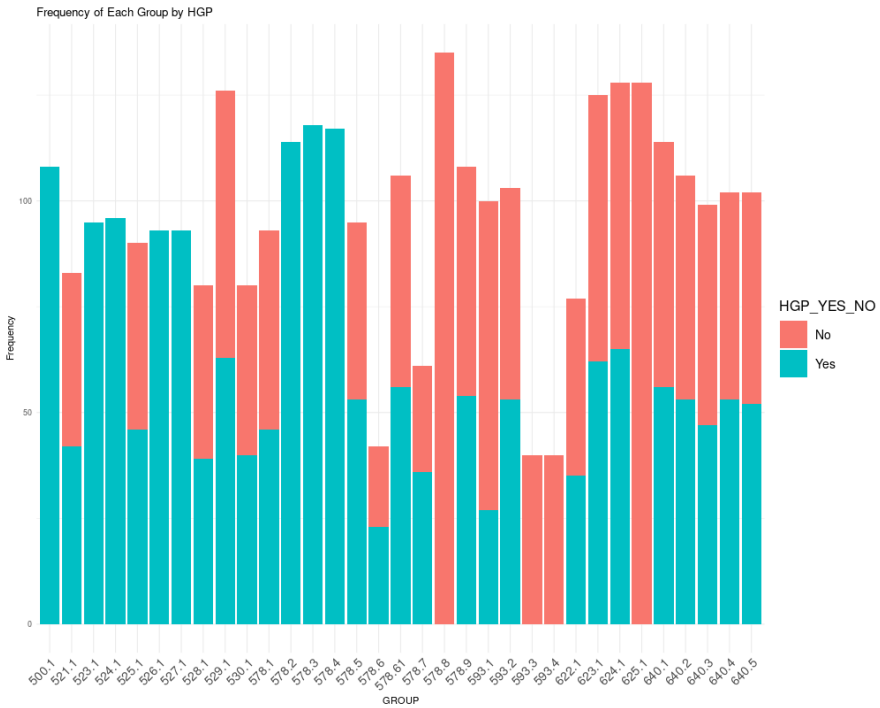
Frequency of Each Group by HGP
Frequency
100
50
0
500.1
521.1
523.1
524.1
525.1
526.1
527.1
528.1
529.1
530.1
578.1
578.2
578.3
578.4
578.5
578.6
578.61
578.7
578.8
578.9
593.1
593.2
593.3
593.4
622.1
623.1
624.1
625.1
640.1
640.2
640.3
640.4
640.5
GROUP
HGP_YES_NO
No
Yes

Fig. 15

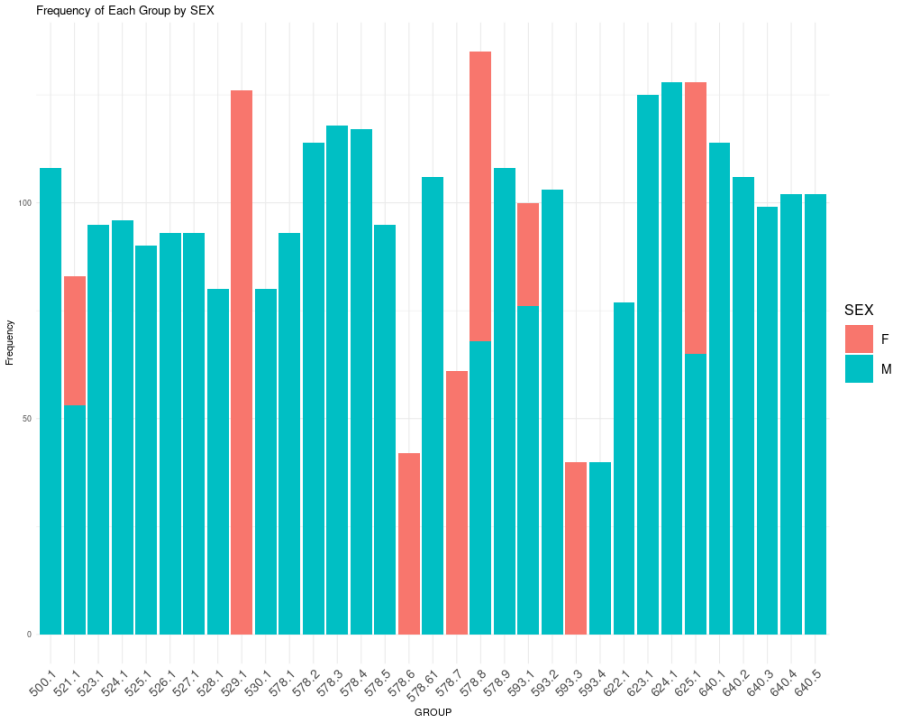

Fig. 16

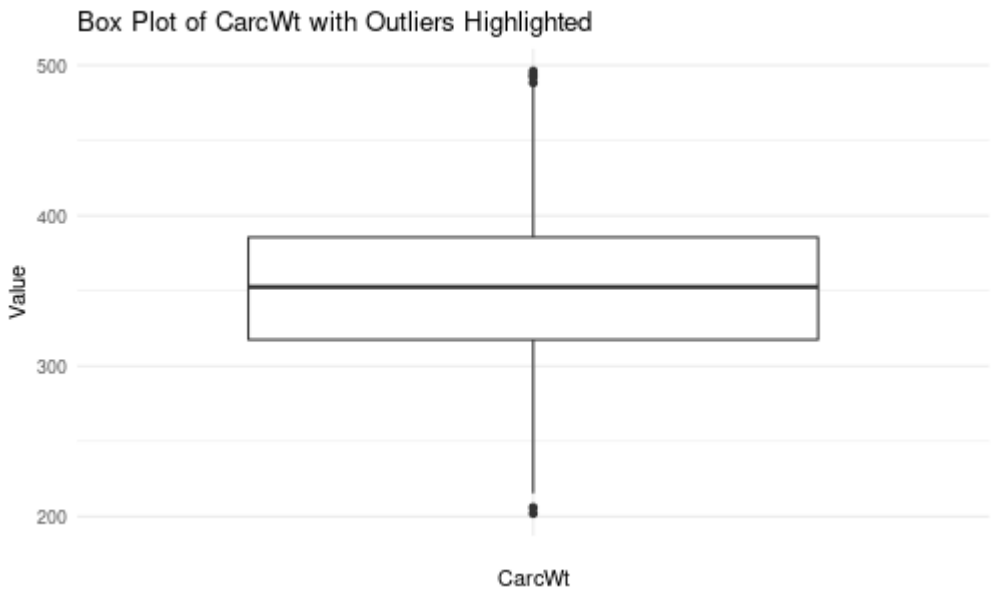

Fig. 17

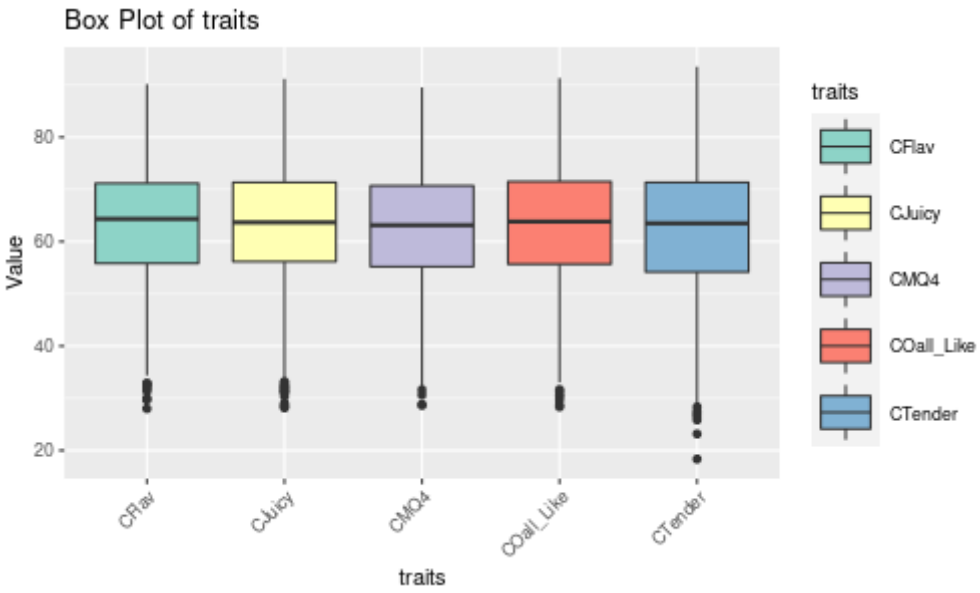
Box Plot of traits
Value
80
60
40
20
CRlav
CJuicy
CMQ4
COall_Like
CTender
traits
traits
CRlav
CJuicy
CMQ4
COall_Like
CTender

Fig. 18

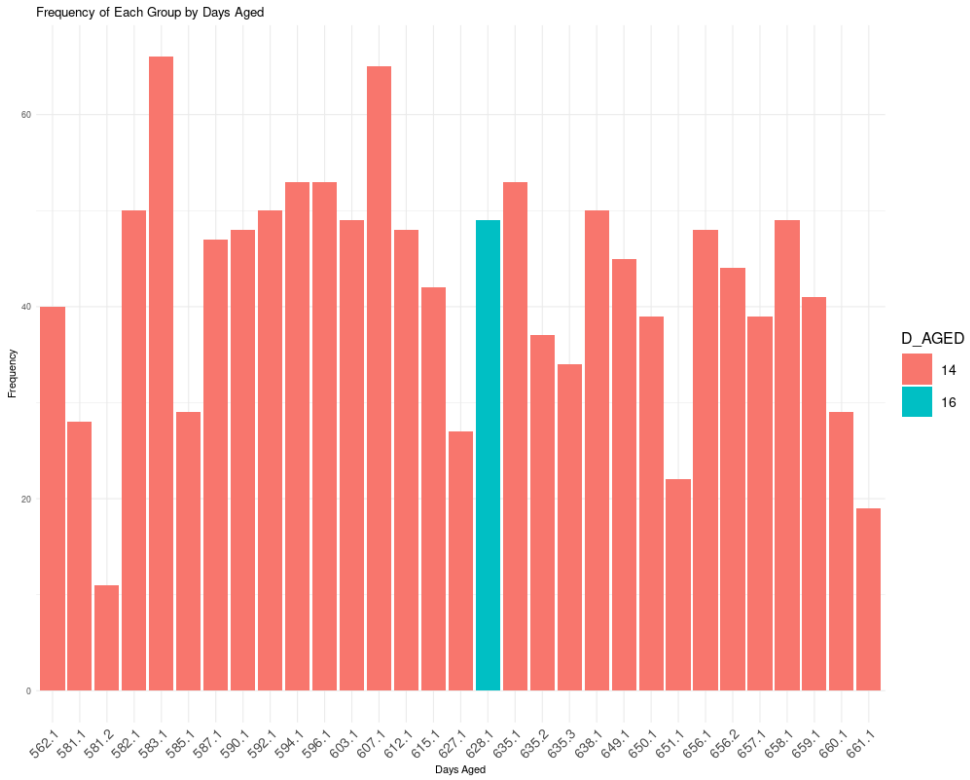
Frequency of Each Group by Days Aged
Frequency
Days Aged
D_AGED
14
16

Fig. 19

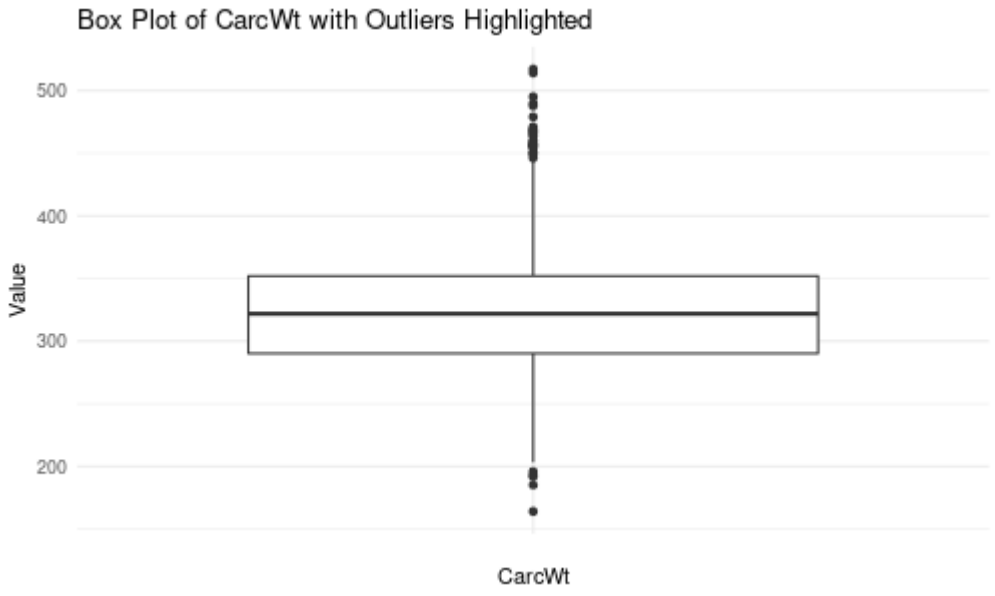
Box Plot of CarcWt with Outliers Highlighted
Value
CarcWt

Fig. 20

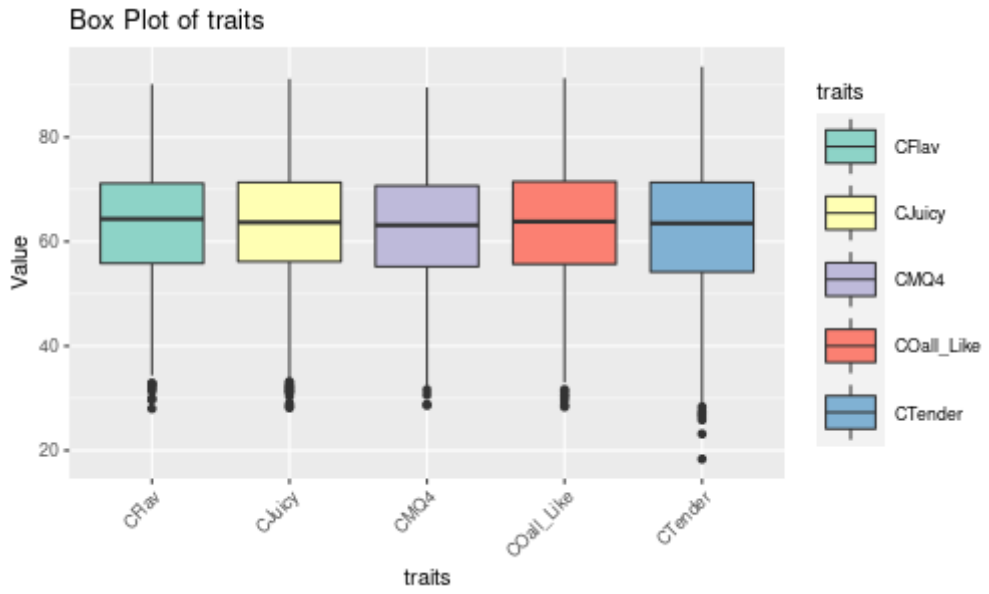
Box Plot of traits
Value
80
60
40
20
CRav
CJuicy
CMQ4
COall_Like
CTender
traits
traits
CFlav
CJuicy
CMQ4
COall_Like
CTender

Fig. 21