University of Natural Resources
and Life Sciences, Vienna

Department of Sustainable Agricultural Systems
Division of Livestock Science

Master Thesis

Mapping genes responsible for coat colour patterns of the Gir cattle breed in Brazil

Victoria Elisabeth Riffert
0940871

Supervisor: Univ. Prof. Dipl.-Ing. Dr. Johann Sölkner
Department of Sustainable Agricultural Systems
Division of Livestock Sciences

Co- Supervisors: Dr. Marcos Vinicius Barbosa da Silva
Empresa Brasileira de Pesquisa Agropecuária, Embrapa Gado de Leite, Rua Eugênio do Nascimento 610, Dom Bosco, 36038-330 Juiz de Fora, MG, Brasil

Dr. Adam Taiti
Department of Sustainable Agricultural Systems
Division of Livestock Sciences

Dr. Solomon Antwi Boison
Department of Sustainable Agricultural Systems
Division of Livestock Sciences

Edlbach, October 2015

Content

Summary

Gir is a cattle breed in Brazil, which belongs to the zebus and is used for milk production. Gir animals were imported from India to Brazil in the years 1870-1962, in 1938 the Brazilian Breeders Association (ABCZ) started running the herd book. There appears a wide spectrum of different coat colour and coat colour patterns in the breed.

The aim of this study was to find the genes that are responsible for the expression of the different coat colour patterns in the Gir cattle of Brazil. Genotypic data of 474 bulls genotyped with Illumina Bovine HD and 1688 cows genotyped with the Illumina BovineSNP50 v2 (50) were available. Coat colour and pattern phenotypes of 2123 animals were recorded. Twelve different phenotypic groups were defined, basically divided into different subcategories of animals with dark skin, sprinkled animals and monochrome animals.

A genome wide association study was performed, applying the program PLINK v1.07. Different phenotypic groups were compared to the monochrome animals that served as control-group. Genomic regions with significant signals were entered into the gene databank ensemble.org to detect the genes located in the area. The genes were followed up with a literature search.

For different forms of spotting there were highly significant regions on Chromosome six and 21 found. The region on Chromosome six harbours the KIT gene which is already known to be responsible for spotting in many mammals. In the region on Chromosome 21 no gene related to colour patterns was found and it is thought to be a copy of KIT.
Animals with dark skin differ from the monochrome ones on Chromosome 18, in a region where MC1R is located. MC1R is known to be linked to pigmentation.
The comparison of the subcategories among each other and to the control-group showed different combinations of the three significant regions, indicating interaction of the genes.
More signals with significant differences were found, but they did not harbour genes already known to be responsible for colour or colour patterns.

In conclusion, several very strong genetic signals for coat colour and pattern differentiation were found in this study. A very significant region on Chr. 21 is not known to be involved in the genetics of coat colour patterns. A study using whole genome sequence data is suggested as a follow up.

Introduction

Coat colour and patterns have always been a distinguishing characteristic for cattle breeds. Coat colour and patterns is a complex set of traits including interaction and epistasis of loci. One single trait can be inherited by dominant, recessive, codominant or quantitative processes (Schmutz 2012)

Since colour is a characteristic, people want to know about the responsible loci for that trait to work on opportunities to trace back products like cheese through the somatic cells to a certain breed. That can be useful for mono-breed labels (Fontanesi, Scotti, and Russo 2010).

The Gir cattle breed was imported from India to Brazil from 1870 to 1962. All together less than 700 animals were imported. From an initially dual-purpose breed the animals were selected by the breeders for dairy traits. In 1938 the Brazilian Zebu Breeders Association (ABCZ) started running the Herd book for Zebu breeds. The Brazilian Dairy Gir Breeding Program (PNMGL) was founded in 1985 and started a progeny testing program. As the breed started in Brazil with a comparatively low number of animals and it is rapidly increasing number of head of cattle the breeders have to be aware of the possibility of inbreeding depression (Santana et al. 2014).

Gir is a breed that is not clearly defined by its colour. There occurs a wide range coat colour and pattern phenotypes in the breed. The animals show different base colours as well as different coat colour patterns. There is no selection by colour by the breeding organisation and the decision can be done based on the preference of the breeders. The heritability of the different patterns is not known. Two crossed monochrome animals can have a sprinkled offspring and so the Gir- herds can be very colourful. JOSÈ OTÀVIO LEMOS dedicated a whole bilingual book to the topic, which is of big interest for people interested in zebu cattle (Otàvio Lemos 2012).

The aim of this thesis is to find significant regions harbouring genes responsible for the different coat colour patterns occurring in the Gir cattle breed of Brazil.

Review of Literature

Variations in coat colour and spotting can be attributed to mutations of the wild type. The ancestor of our cattle breeds, the aurochs, had a brownish colour which can be seen as the wild type of cattle. Some breeds like Jersey or Longhorn show this colour. Variations from the wild type can occur due to removal of pigmentation or lightening and cause a red or solid black phenotype. The white colour of breeds like Chianina and some Zebu cattle or Brown Swiss is a product of loss of pigmentation.
Some loci are known to be responsible for coat colour expression.
The locus causing most of the variation in coat colour is the extension (*E*) locus on chromosome 18 with the three haplotypes E^D, E^+ *and e.* The *E* locus regulates the production of tyrosinase. Low levels of tyrosinase lead to production of phaeomelanin and therefore the expression of red coat colour. Dark coat colour is caused by eumelanin due to high levels of tyrosinase
Agouti (*A*) impacts the expression of the wild type. The impact of agouti is known in dogs, mice and horses. In cattle it is said to be responsible for the yellow belly of Limousin as well as it affects the wild type of Brown Swiss.
Animals that have pigmentation neither in the skin nor in the eyes are called albino. When an animal has the haplotype *c/c* on the *C* locus it is a true albino.
Dilution appears in Simmental, Gelbvieh and Charolais. The haplotype of deluted Charolias is Dc^+ which is completely dominant over the wild type.
The *S* locus is responsible for white spotting in cattle. The order of dominance at the *S* locus is S^H (Hereford pattern) = S^p (line-backed) > S^+ (Wild type without spotting) > *s* (irregular white spotting).
The symbol *Bl* stands for blaze which leads in combination with S^+ to solid coloured animals with a white blaze on the face.
Also roan, colour-sided and belted animals as well as animals with pigmented legs occur in cattle (Olson 1999).

There were numerous studies focussing on coat colour and patterns. Several genes are involved in the coat colour and patterns expression in cattle. Genes responsible for pigmentation are including hormone receptors, signalling hormones and transcription factors. Base pair substitutions, insertions and deletions, copy number variation and promoter mutations can be the result of a mutation (Schmutz 2012).

The following genes listed in Table 1 are known to be responsible for basic coat colour in cattle (Bos Taurus):

Gene	Coat colour
MC1R	black, brownish and grey (E^D), red (*e*, recessive), both expressions possible in the wild type (E^+)
TYRP1	Chocolate, brown (*b*), dun brown (*b*,*b* in combination with E^D)
TYR	Albinism
DEFB 103	Variant red (copy number variation of five SNPs in the `5 UTR)

Table 1: Genes responsible for basic coat colour in cattle (Schmutz 2012)

In Holstein a mutation of COPA (Chromosome 3), a region that has not been related to colour before, was found to cause the dominant red phenotype (Dorshorst et al. 2015).
In Highland cattle there are six base colours from silver over yellow to black known. SCHMUTZ and DREGER found MC1R is interacting with PML. There were six possible haplotypes for PML found causing the different phenotypes (Schmutz and Dreger 2013). PML and AP3B2 are also discussed to be responsible for different shades of red in Fleckvieh (Mészáros et al. 2015).
Several studies had the aim to find genes responsible for spotting in different cattle breeds. In Holstein cattle the proportion of black in the coat was studied using a genome-wide association study. The gene KIT (chromosome six), the MITF locus on chromosome 22 and a region on chromosome eight at 64 Mb near PAX5 were found to explain 24% of the variance in the proportion of black. The missing 76% are split onto small effects of many genes (Hayes et al. 2010).

The different facial markings in Fleckvieh were traced back to the gene KIT. For Circum-ocular pigmentation and spot on the cheek the highest signal was found on chromosome 6 (KIT). No significant signal was detected to the phenotype spot on the eye (Mészáros et al. 2015).

Recently it has been found that the colour sidedness in Brown Swiss, Belgian blue. Pustertaler Sprinzen and several other cattle types, is induced by a translocation of the KIT gene to chromosome 29 (Durkin et al. 2012). After that finding, more research in that field was done. A similar effect was found in Galloway, mismarked animals were homozygous for the KIT gene insertion on BTA29 (Cs29/Cs29), well-marked and strongly marked White Galloway cattle were heterozygous (Cs29/ wt29) (Brenig et al. 2013).

Material and Methods

Genotypic data

There was SNP-Data of the Gir cattle breed provided by Embrapa Dairy Cattle (Jiuz de Fora, Brazil), a breeding program in Brazil. The data consists of 474 bulls genotyped with Illumina Bovine HD and 1688 cows genotyped with the Illumina BovineSNP50 v2 (50).

Quality control was done in the program PLINK v1.07 (Purcell et al. 2007) for the Illumina Bovine HD SNP-Data. SNP markers that had a call rate smaller than 90%, those with a minor allele frequency was lower than 0.02 and SNP with a p-value lower than 0.0000001 for the Hardy-Weinberg equilibrium test were excluded from the analyses. From the dams only data with a call rate higher than 0.9 was kept in the analysis.

Then the 50K data of the cows was imputed using FImpute v2.2 (Sargolzaei, Chesnais, and Schenkel 2014) to have a more comprehensive database. The data was imputed following the methodology of BOISON et al. (Boison et al. 2015).

Phenotypic data

For this study phenotypic data of 1678 cows and 445 bulls that were classified into twelve groups by the breeding organization (ABCZ) was used. Table 2 shows the possible phenotypes in the Gir cattle breed. There is the Portuguese name, the English translation and the English description given. As the English translation does not describe the whole meaning of the Portuguese names the Portuguese names are kept in this study. To visualize the descriptions of the animals better, there is a picture for each phenotype provided in Figure 1-12.

	Coat colour Portuguese	**Translation English**	**Description**
1	Amarela	yellow	whole body yellow
2	Amarela Chitada	yellow calico	yellow base colour, white sprinkles
3	Amarela Gargantilha	yellow choker	yellow body, dewlap sprinkled
4	Chita de Amarela	calico yellow	light base colour, yellow sprinkles
5	Chita Clara	calico clear	light-coloured base colour, with few spots (legs, tail)
6	Chita de Vermelha	calico red	light-coloured base colour, red sprinkles

7	Moura Clara	dark skin, coat colour light	Dark skin, head dark (light parts of the head possible, inner ears dark), light-coloured body
8	Moura Escura	dark skin, coat colour dark	Dark skin, dark inner ears, dark coat
9	Moura de Vermelha	dark skin, head red	dark skin, dark inner ears, red head, light-coloured body
10	Vermelha	red	whole body red
11	Vermelha Chitada	red calico	red base colour, white sprinkles
12	Vermelha Gargantilha	red choker	red body, dewlap sprinkled

Table 2: Phenotypic groups of the Gir cattle (Otàvio Lemos 2012)

Figure 1: Phenotype Amarela

Figure 2: Phenotype Amarela Chitada

Figure 3: Phenotype Amarela Gargantilha

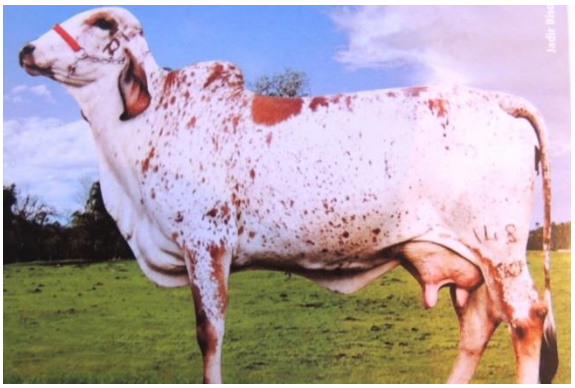

Figure 4: Phenotype Chita de Amarela

Figure 5: Phenotype Chita Clara

Figure 6: Phenotype Chita de Vermelha

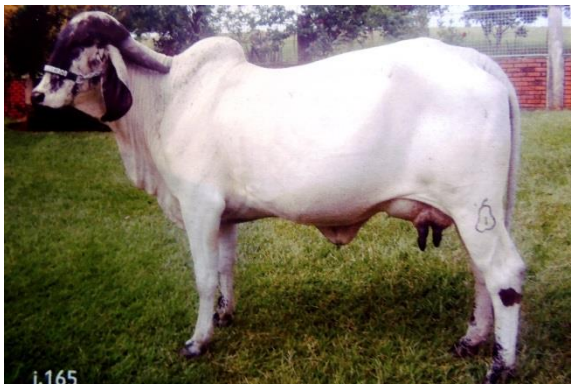

Figure 7: Phenotype Moura Clara

Figure 8: Phenotype Moura Escura

Figure 9: Phenotype Moura de Vermelha

Figure 10: Phenotype Vermelha

Figure 11: Phenotype Vermelha Chitada

Figure 12: Phenotype Vermelha Gargantilha

Genome wide association analysis

Different phenotypes were merged in groups to find regions of the genome responsible for different expressions in the phenotype. The red and yellow animals were put together and functioned as a control-group as they are monochrome and do not show any colour patterns.

The analyses were calculated in the programs PLINK v1.07 and R version 3.1.2.
For every comparison a Single SNP Analysis was calculated. The model included the phenotype and the genotype. Sex and dominance effects did not show an effect, so they were not included into the model. Due to strong relatedness of the animals a correction for eigenvectors was done.

The result was the effect size and the significance (p-value) for each SNP. The results were plotted using the package “qqman” (Turner 2014), therefore the negative log (-log10) of the p-value was calculated. To find the SNPs that are associated with the traits of interest the Bonferroni threshold was calculated (significance level $p<0.05$). As there were 496727 SNP in the analysis the significance line was at 6.99715 (-log10 (0.05/496727)).

To visualize the differences between the expected and observed p-values QQ-plots were created for every comparison.

There was a list of the significant SNPs for each comparison created. The significant SNP with lowest map position and the significant SNP with highest map position were selected. Their positions on the genome were taken and 500.000 Bp were added left and right of that region. If there was more than one significant region on a Chromosome, this was performed separately for each region.

The regions found this way were entered into the gene-databank ensemble.org and that way the genes lying in these regions found. There was a literature search made in “google.scholar” and “boku:LITsearch” with the found genes and the keywords “cattle”, “colour”, “pigment” etc.

Results

Animals with dark skin

Moura

As they all show the dark inner ears due to a pigmented skin the three phenotypic groups “Moura Escura”, “Moura de Vermelha” and “Moura Clara” were put together and compared with the control-group, that consisted of the animals that were phenotyped as “Amarela” and “Vermela”. There were 50 animals in the group “Moura” and 220 in the control-group.

Figure 13 shows a significant region on Chromosome 18. To visualize the peak better Figure 14 provides the Manhatten plot of Chromosome 18. There were nine significant SNPs on Chromosome 18 in the region of about 11.000.000 BP and 15.000.000 BP respectively, where the gene MC1R is located. The QQ-plot of Figure 15 shows a big difference in the samples.

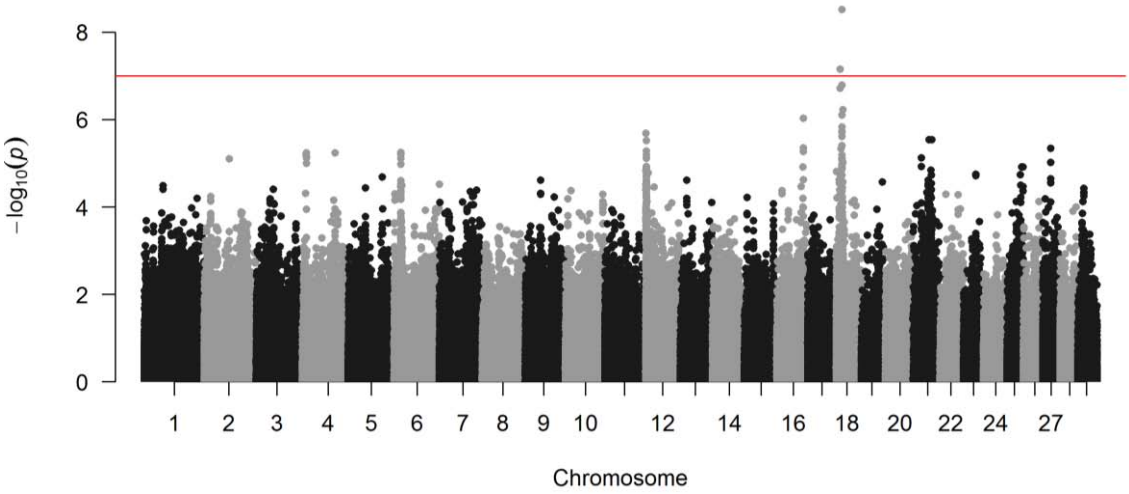

Figure 13: Moura vs. Monochrome Manhattan Plot

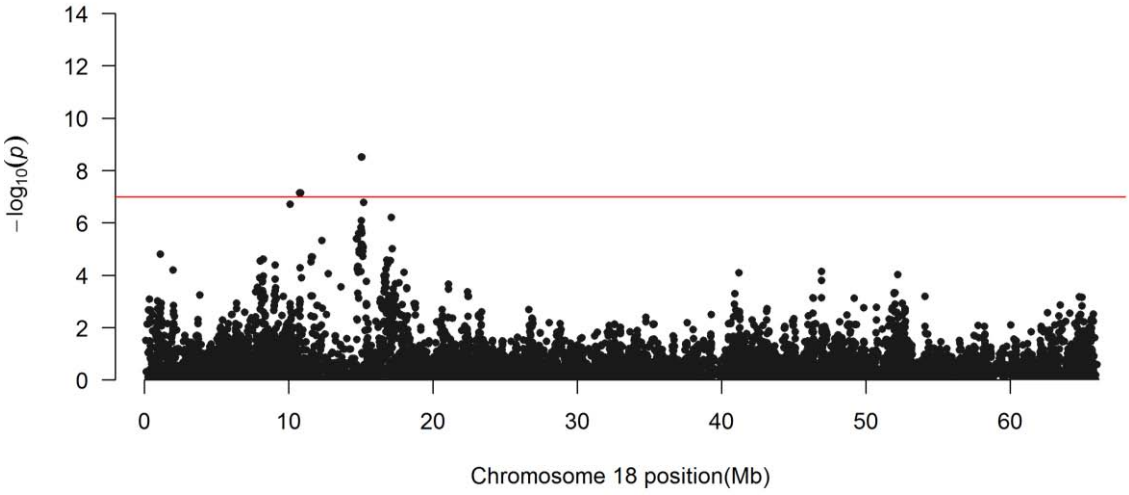

Figure 14: Moura vs. Monochrome Manhattan Plot Chr. 18

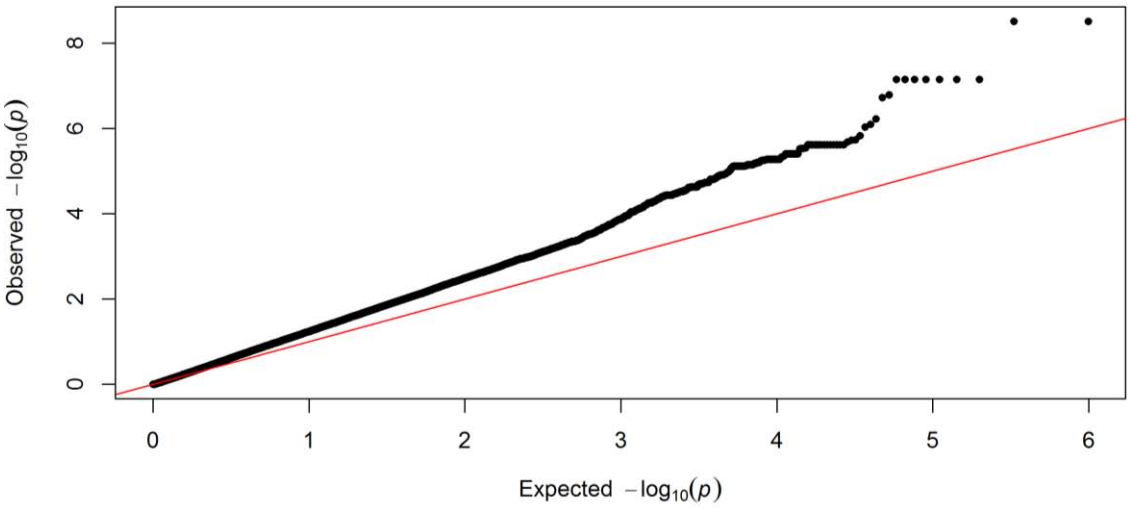

Figure 15: Moura vs. Monochrome QQ-Plot

Moura Escura

There was a comparison between the twelve animals phenotyped as “Moura Escura” and the control-group calculated. As to prove in Figure 15 no SNP appeared above the significance line (6.99715). Figure 17 provides the QQ-plot for the “Moura Escura” group, which shows a big difference between observed and calculated p-values.

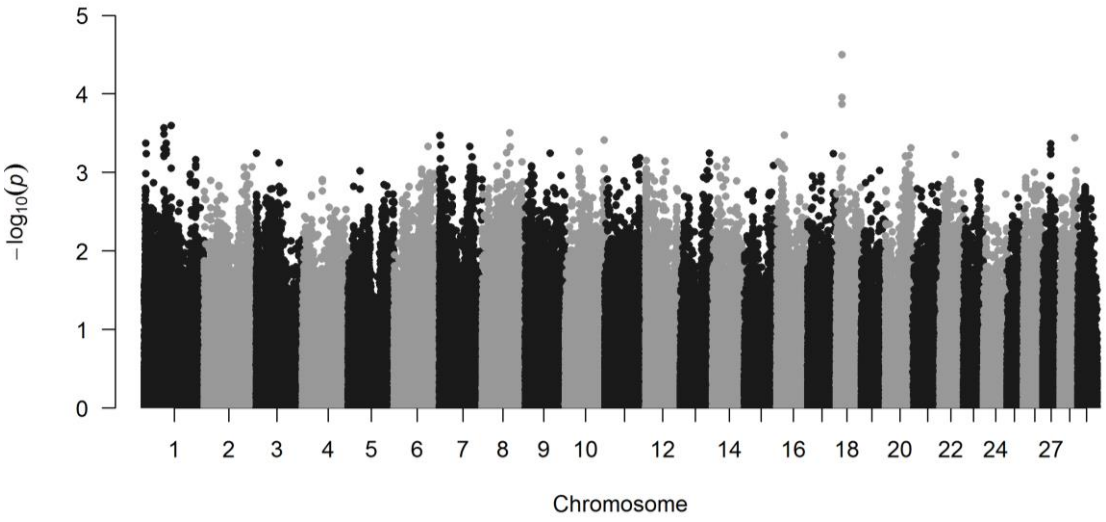

Figure 16: Moura Escura vs. Monochrome Manhattan Plot

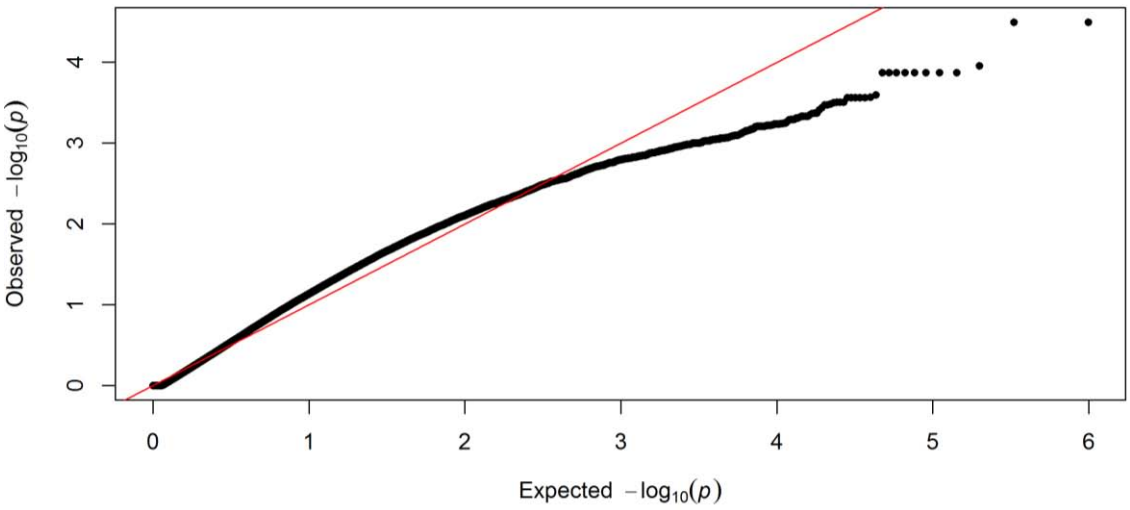

Figure 17: Moura Escura vs. Monochrome QQ-Plot

Moura Clara

The group “Moura Clara” was composed of 34 animals and compared against the 220 animals of the control-group. The Manhattan plot of Figure 18 shows significant SNPs on the same regions as in the Moura group where the MC1R gene is located. The zoom into Chromosome 18 was made with Figure 19. One can see differences in the compared groups on the QQ-plot presented in Figure 20.

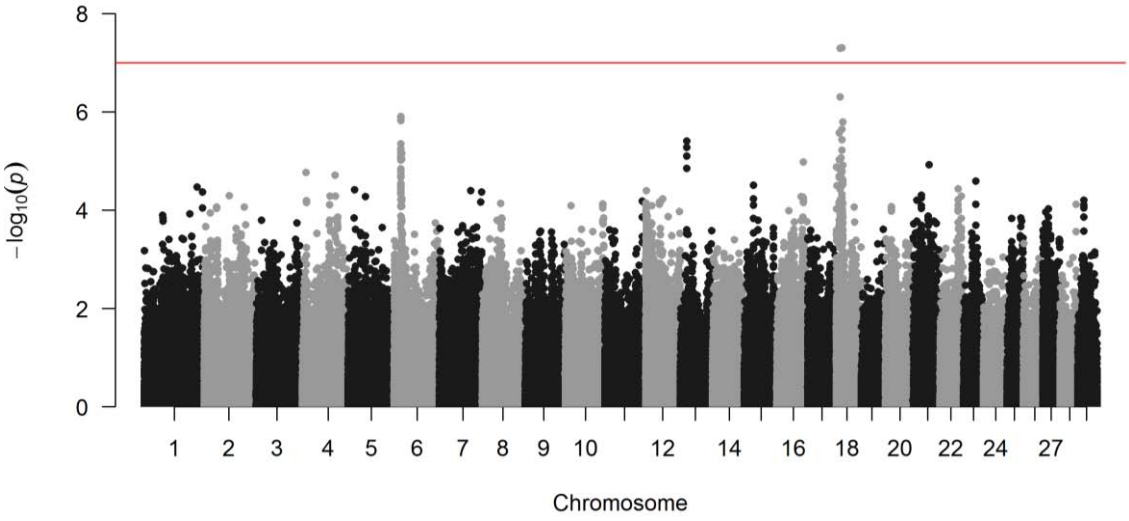

Figure 18: Moura Clara vs. Monochrome Manhattan Plot

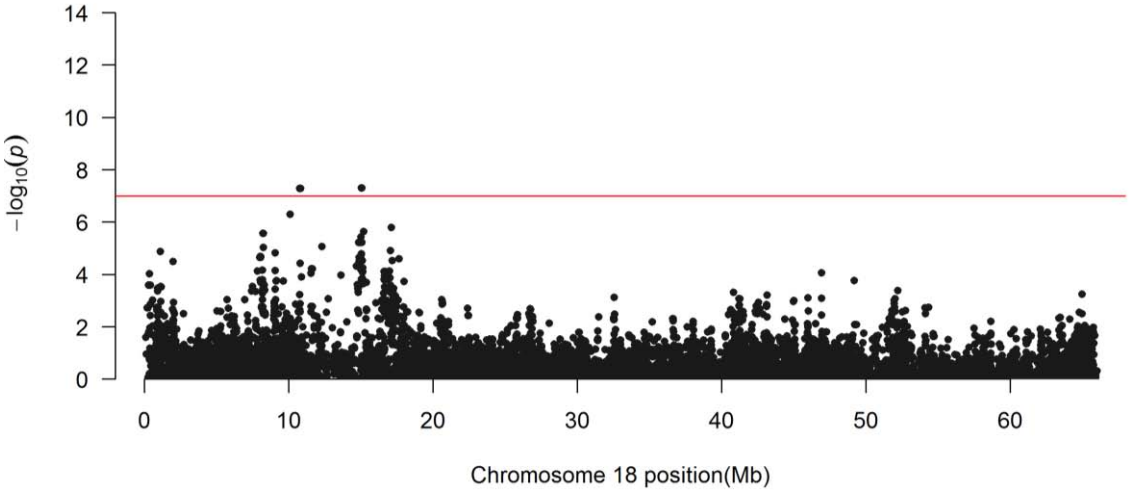

Figure 19: Moura Clara vs. Monochrome Manhattan Plot Chr. 18

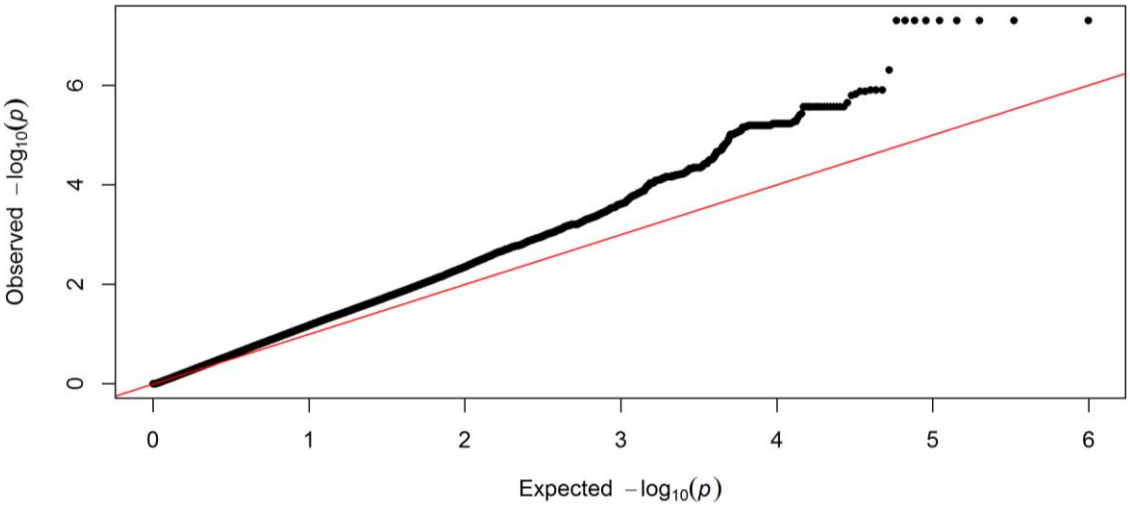

Figure 20: Moura Clara vs. Monochrome QQ-Plot

Moura Clara and Moura Vermelha

Due to small sample size and only little difference in the phenotype “Moura Vermelha” with four animals was put together with the group “Moura Clara” that consisted of 34 animals. These groups were compared with the control-group. A view on Figure 21 displays: there came up the same significant regions on Chromosome 18 as in the “Moura Clara” group. Again a zoom on Chromosome 18 is provided with Figure 22. The found gene in that region is MC1R. Figure 23 demonstrates differences between calculated and expected p-values in the QQ-plot.

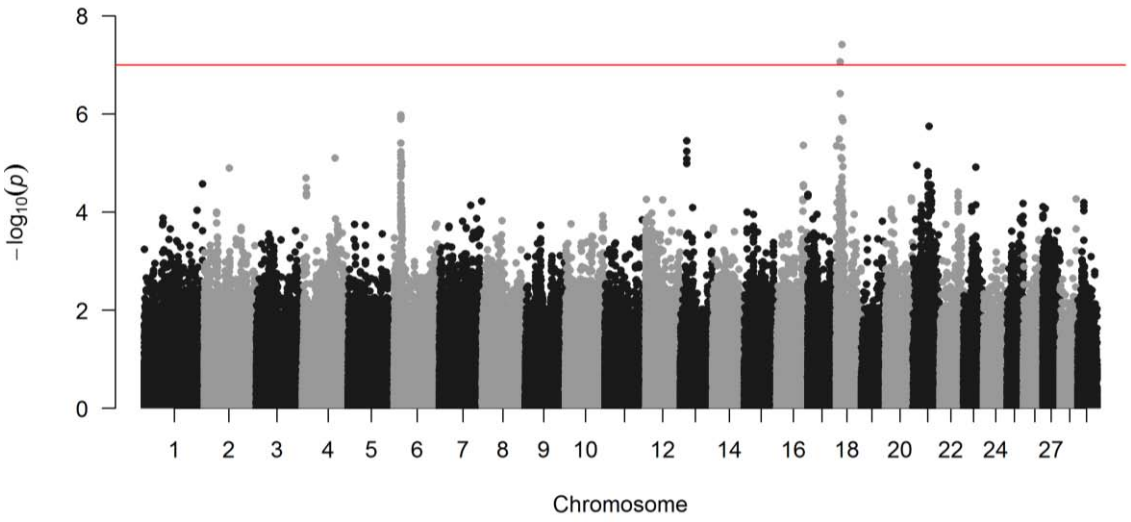

Figure 21: Moura Clara/ Vermelha vs. Monochrome Manhattan Plot

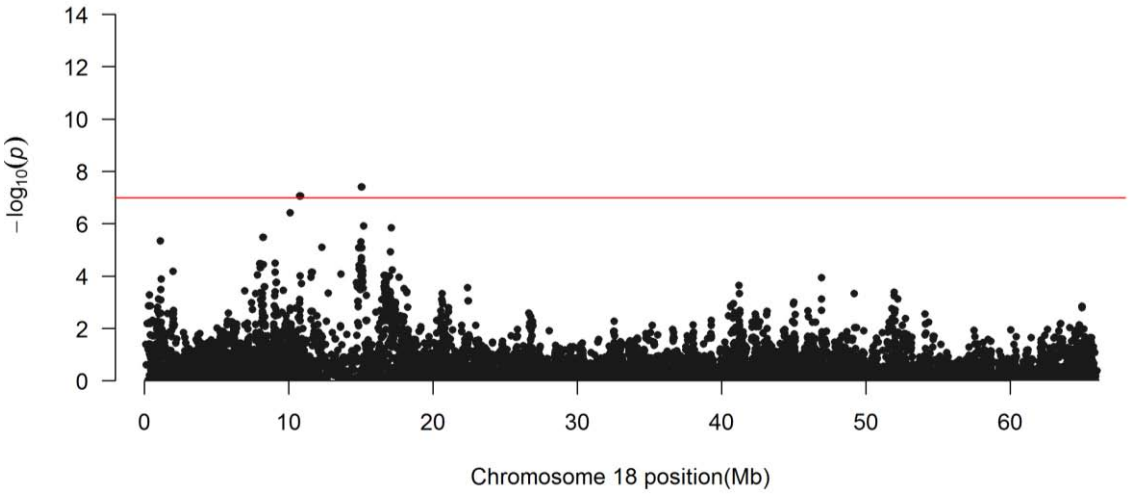

Figure 22: Moura Clara/ Vermelha vs. Monochrome Manhattan Plot Chr. 18

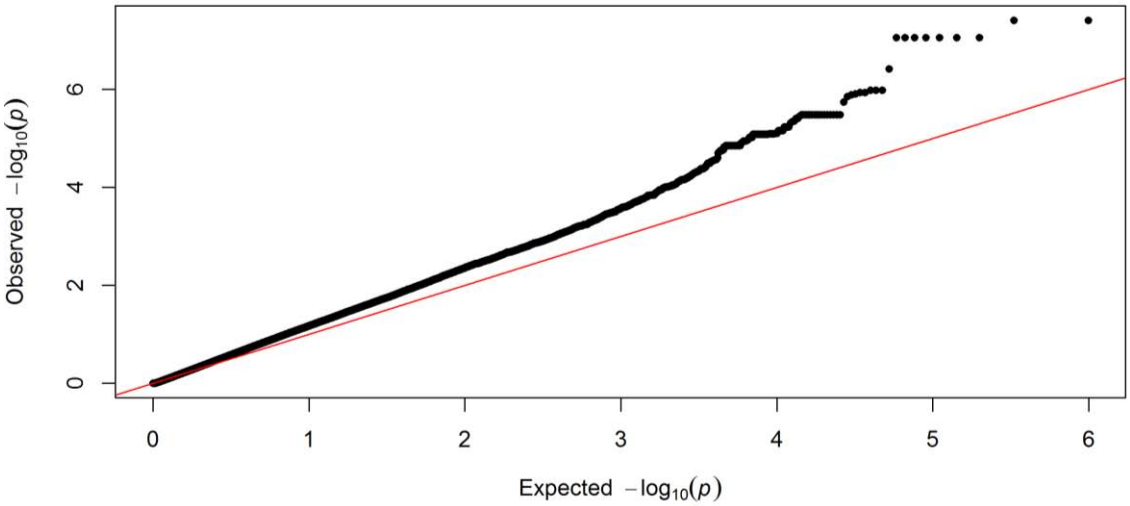

Figure 23: Moura Clara/ Vermelha vs. Monochrome QQ-Plot

Moura Escura vs. Moura Clara/ Vermelha

There was a big difference in the phenotype between the “Moura Escura” and the two other groups, “Moura Clara” and “Moura Vermelha”. For that reason and because there were only few animals of that phenotype (50 head of cattle), a Fisher`s exact test was done. The results can be seen in Figure 24. No differences were found.

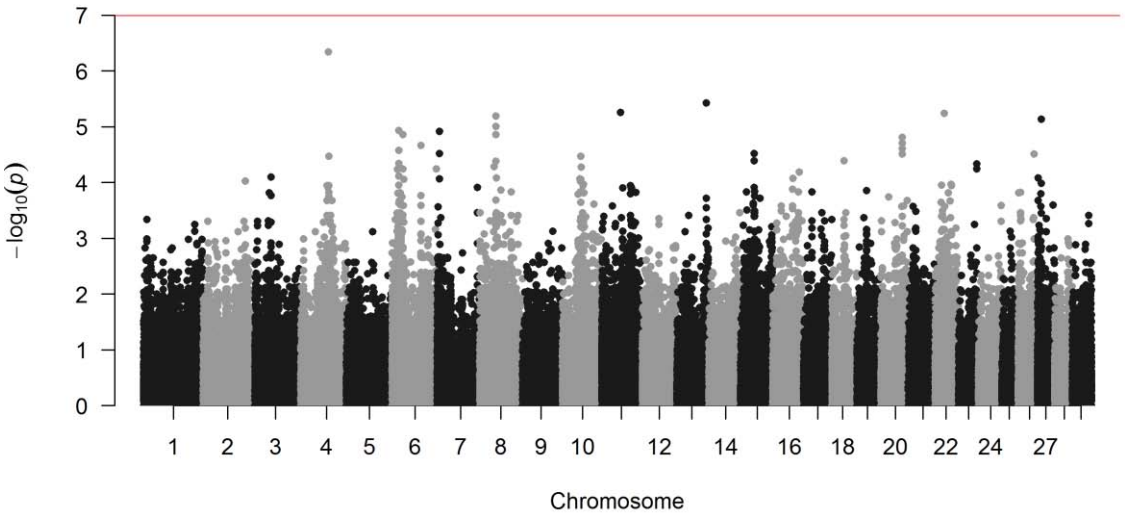

Figure 24: Moura Escura vs. Moura Clara/Vermelha Fisher`s exact test

Sprinkled animals

Chita Clara

The 267 animals with white base colour and few spots on the legs and the tail (called “Chita Clara”) were compared with the control-group. The Manhattan plot in Figure 25 illustrates that two highly significant regions were identified on Chromosome six and 21. On Chromosome six the significant SNPs are located between 68.500.000 and 74.500.000 Bp, which can be looked up in Figure 26. In this region the KIT gene is located. The region that passed the significance line in Chromosome 21 lies between 35.800.000 and 54.000.000 Bp. Figure 27 shows a zoom of this peak. There was no gene related to colour found in that area, the most significant SNPs occurred at around 40.000.000 Bp where no gene was found. The Manhattan plot for Chromosome 22 in Figure 28 shows that one SNP passed the significance line at Bp 54017830. There the genes CCCR6 and CCR9 were found. A big difference between the groups shows the QQ-plot of Figure 29.

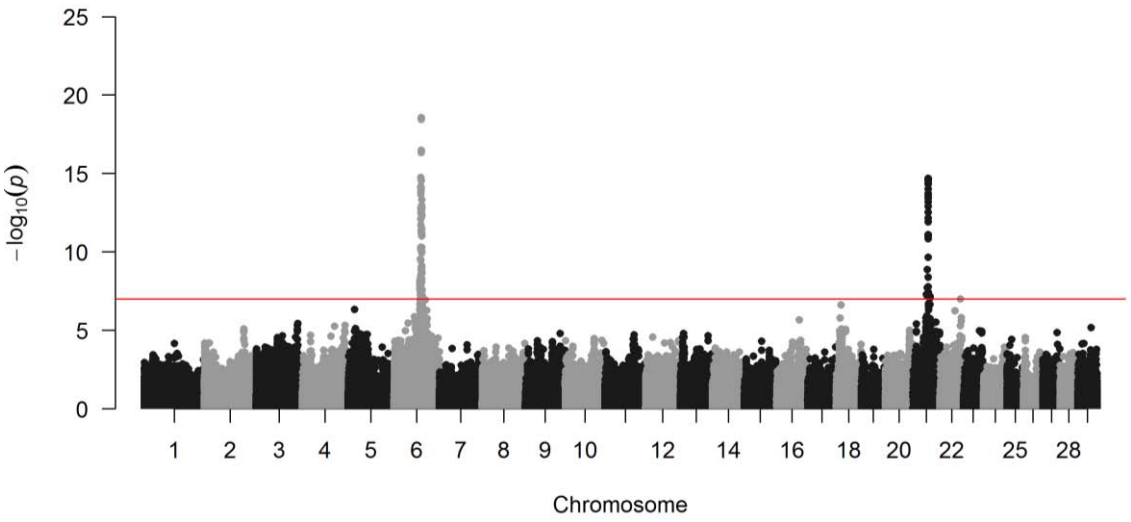

Figure 25: Chita Clara Manhattan Plot

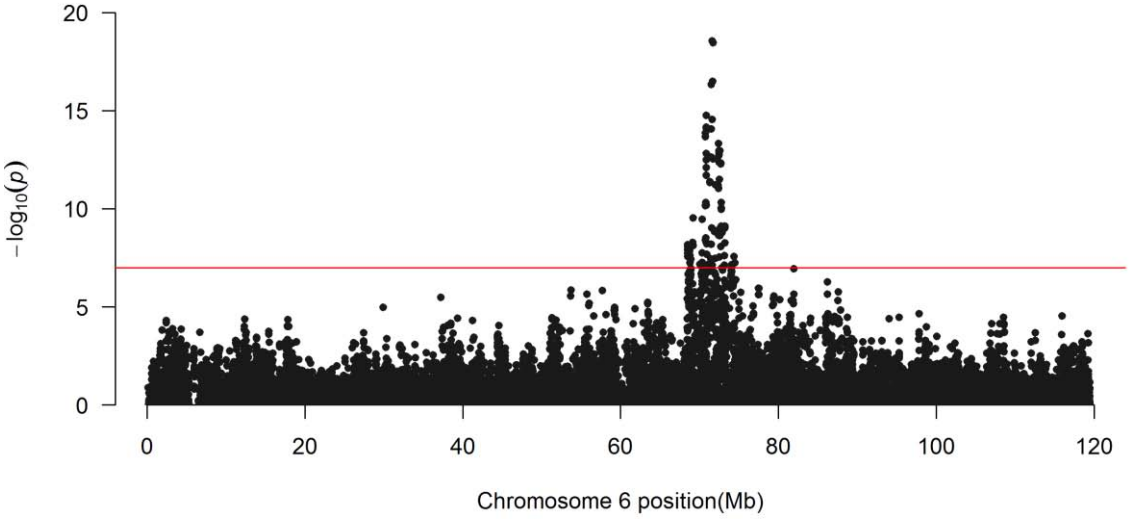

Figure 26: Chita Clara Manhattan Plot Chr. 6

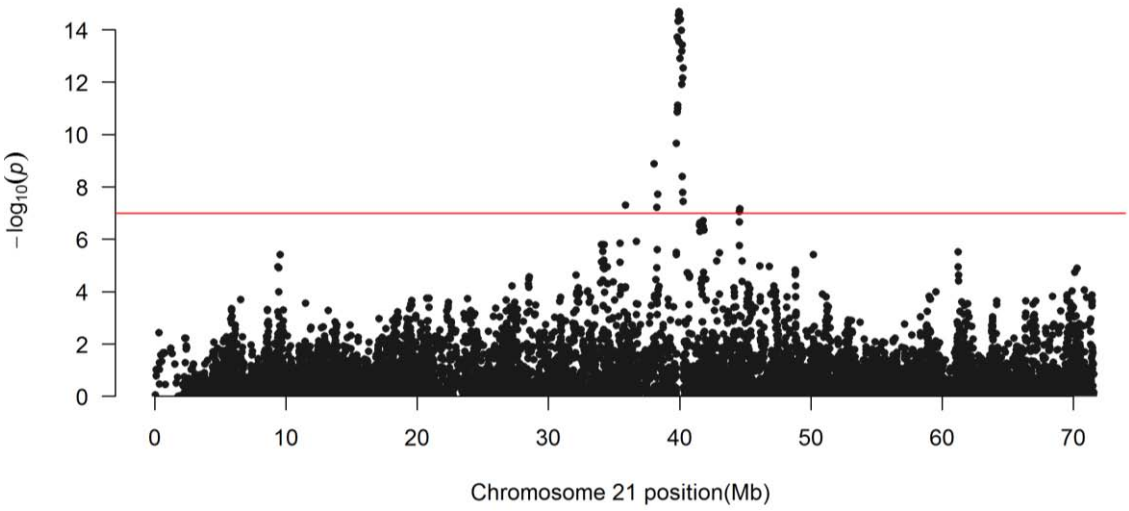

Figure 27: Chita Clara Manhattan Plot Chr. 21

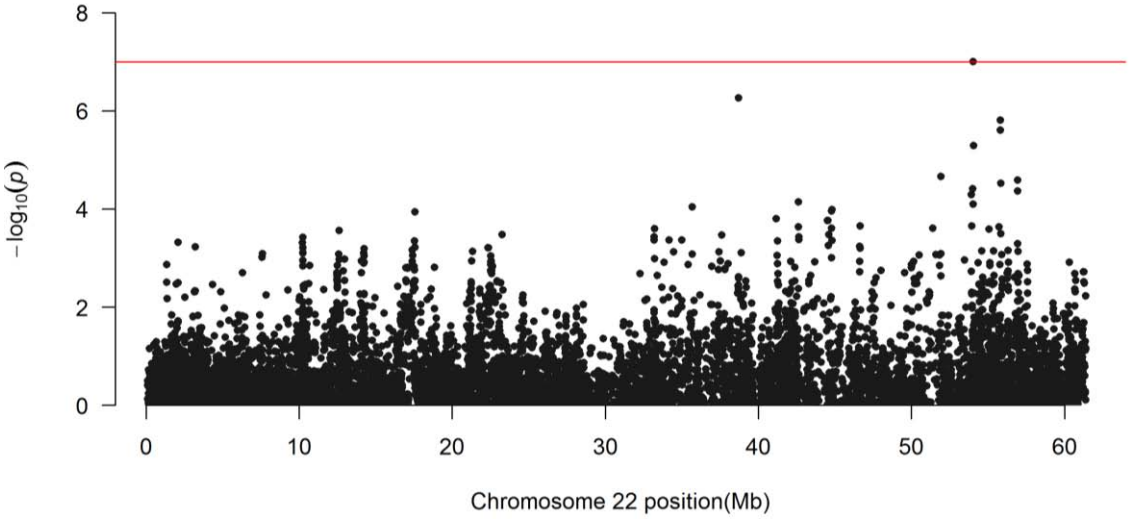

Figure 28: Chita Clara Manhattan Plot Chr. 22

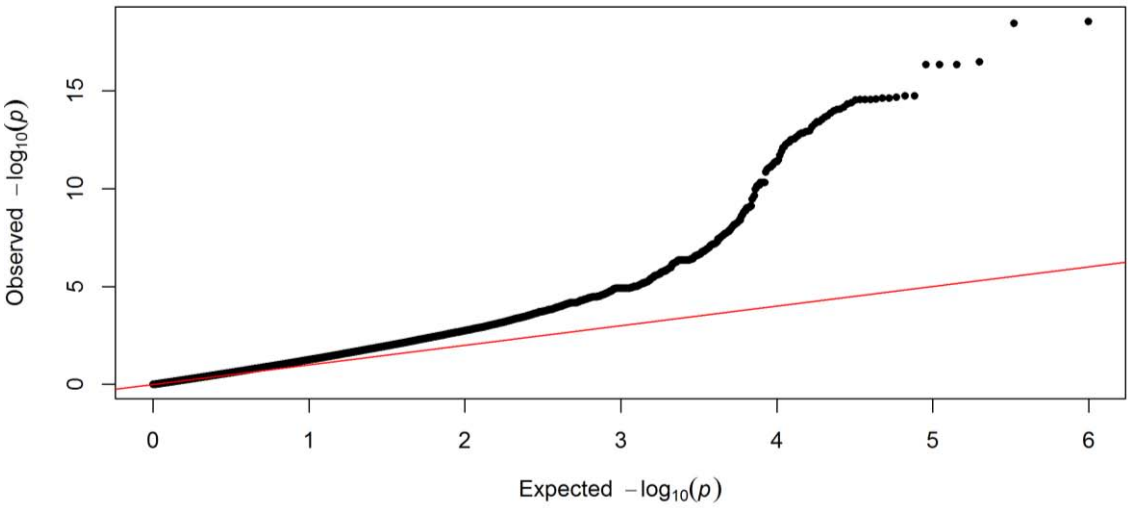

Figure 29: Chita Clara QQ-Plot

Chita de Vermelha and Chita de Amarela

As they show the same colour pattern, the groups “Chita de Vermelha” and “Chita de Amarela”, together 452 animals, were compared against the control-group. The Manhattan plots visualize significant SNPs on Chromosome six, 18 and 21. Figure 30 provides the Manhattan plot for the whole genome. The Manhattan plot for Chromosome six is shown in Figure 31. The gene KIT is again located in the significant region of Chromosome six, but the region is wider and ranges from 63.000.000 to 82.000.000. A close look at Figure 32 screens the significant region on Chromosome 18. There came up two significant SNPs at around 14.000.000 Bp. On Chromosome 18 there are few significant SNP located around 15.000.000 Bp where MC1R lies. Figure 33 shows a small peak at 10.000.000 Bp and a big one between 33.000.000 and 50.000.000 Bp was located on Chromosome 21. No gene related to coat colour was found in that region. The QQ-plot shows big differences between calculated and expected phenotypes, which is visualized in Figure 34.

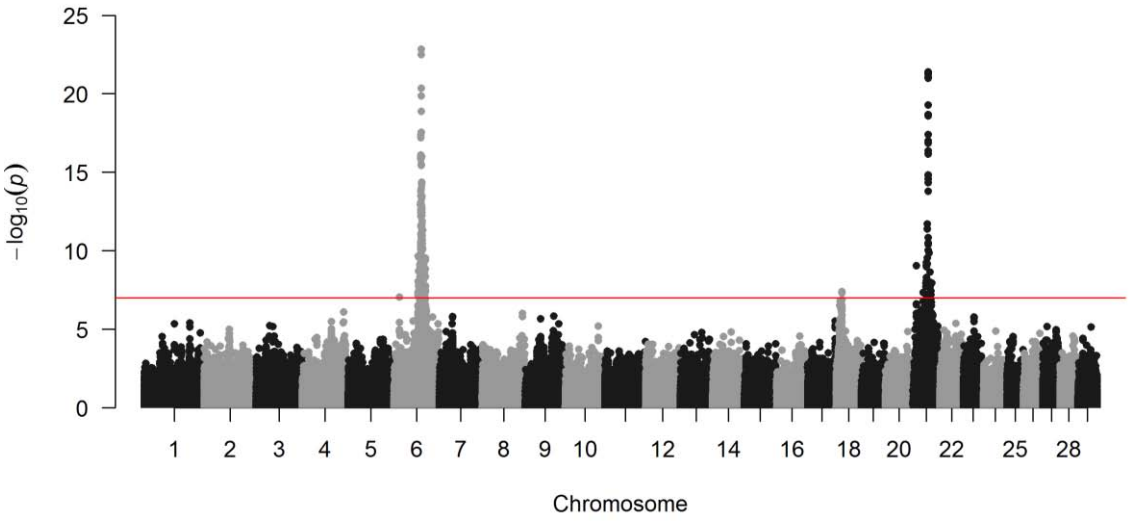

Figure 30: Chita de Amarela/ Vermelha Manhattan Plot

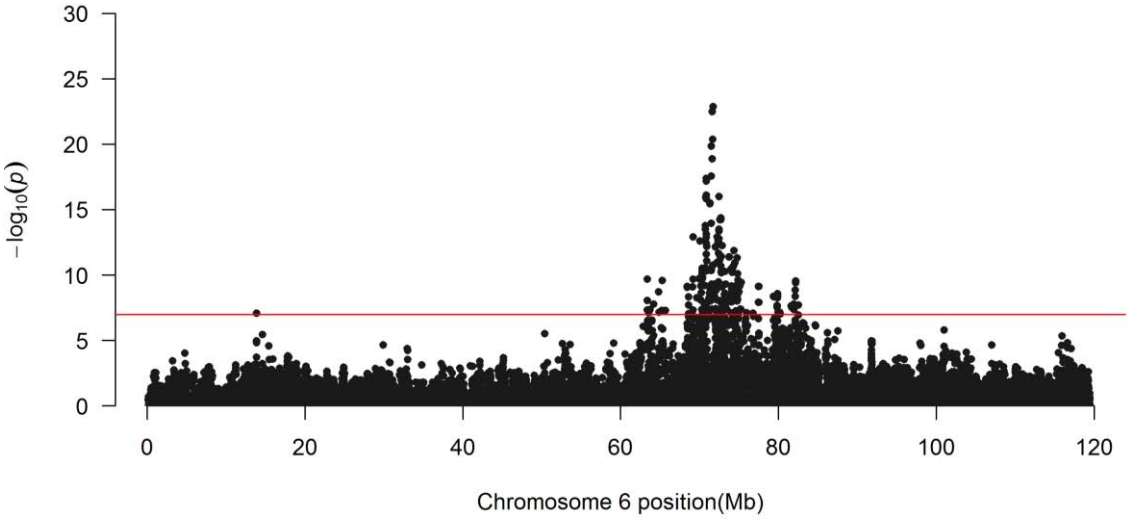

Figure 31: Chita de Amarela/ Vermelha Manhattan Plot Chr. 6

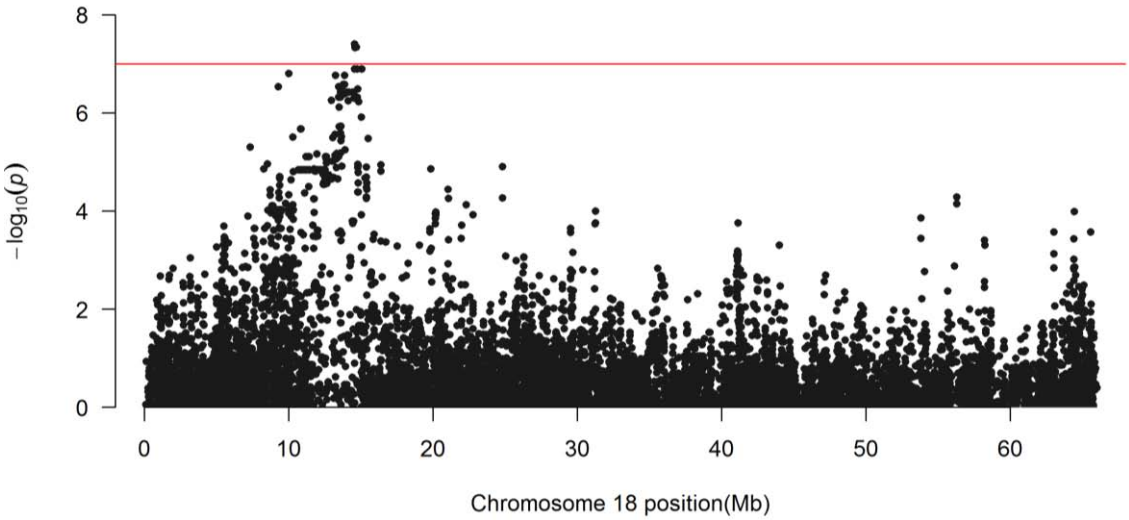

Figure 32: Chita de Amarela/ Vermelha Manhattan Plot Chr. 18

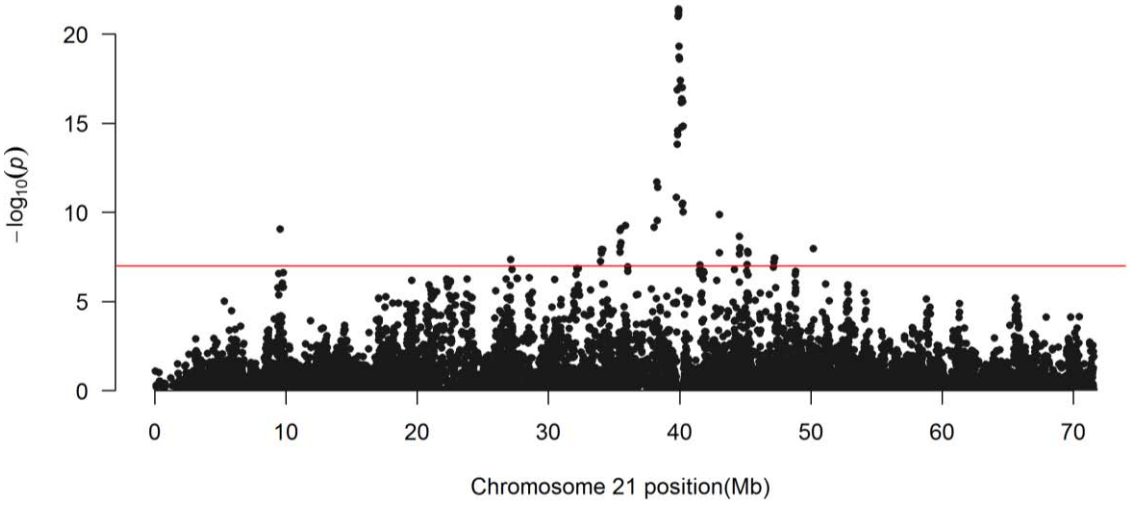

Figure 33: Chita de Amarela/ Vermelha Manhattan Plot Chr. 21

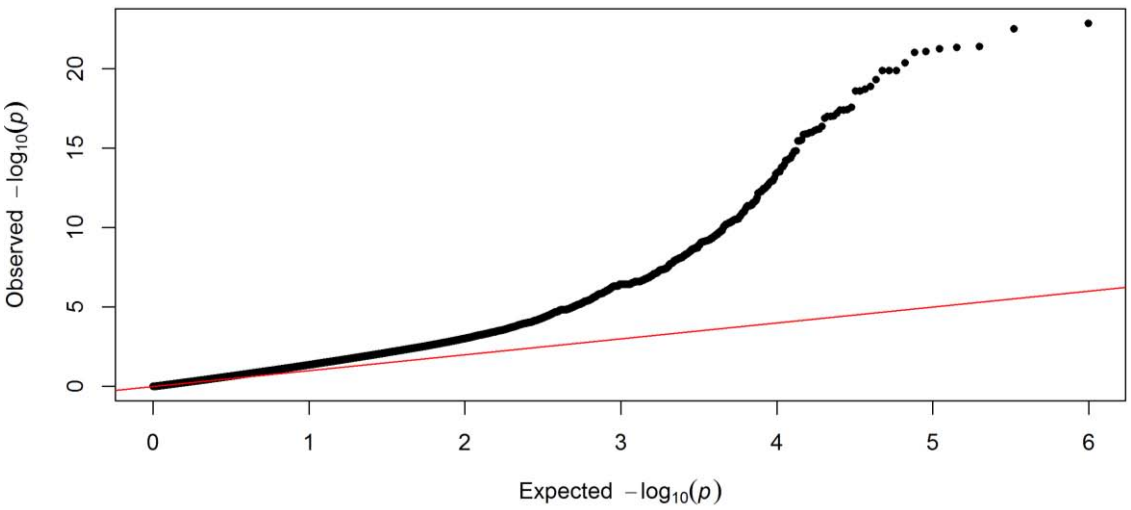

Figure 34: Chita de Amarela/ Vermelha QQ-Plot

Amarela Chitada and Vermelha Chitada

The phenotypic groups “Amarela Chitada” and “Vermelha Chitada” consisting of 805 animals were compared with the control-group. The Manhattan plot of Figure 35 shows a high (35.000.000 - 45.000.000 Bp) peak and two small ones (20.000.000 and 50.000.000 Bp) on Chromosome 21 where no genes related to colour were found as well as a small one on Chromosome nine at 40.000.000 Bp where FIG4 lies The peaks are shown in a more detailed graph in Figure 36 for Chromosome nine and for Chromosome 21 in Figure 37. Big differences in the compared groups can be seen in Figure 38, which provides the QQ-plot for the group “Amarela Chitada” and “Vermelha Chitada”.

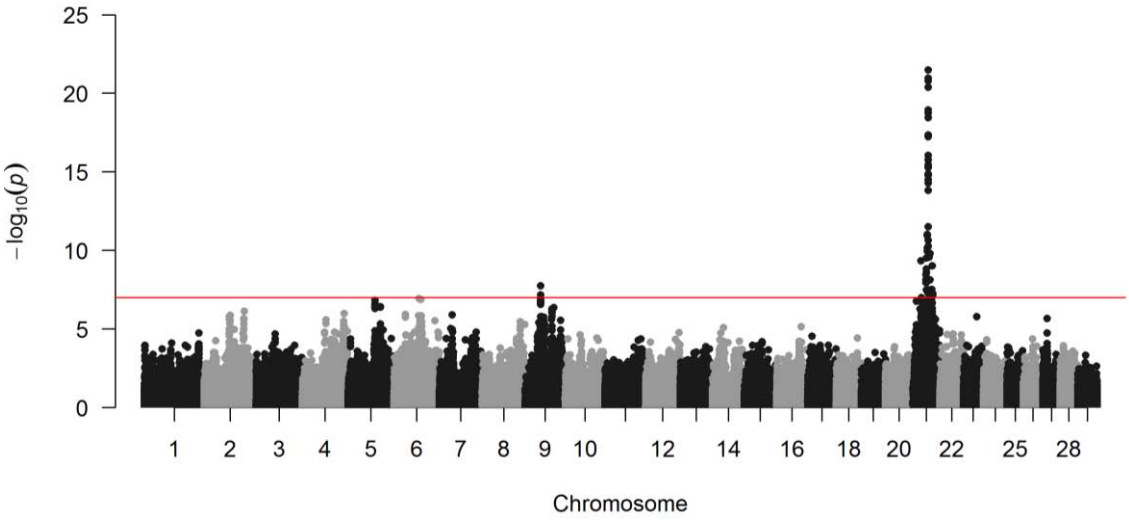

Figure 35: Amarela/ Vermelha Chitada Manhattan Plot

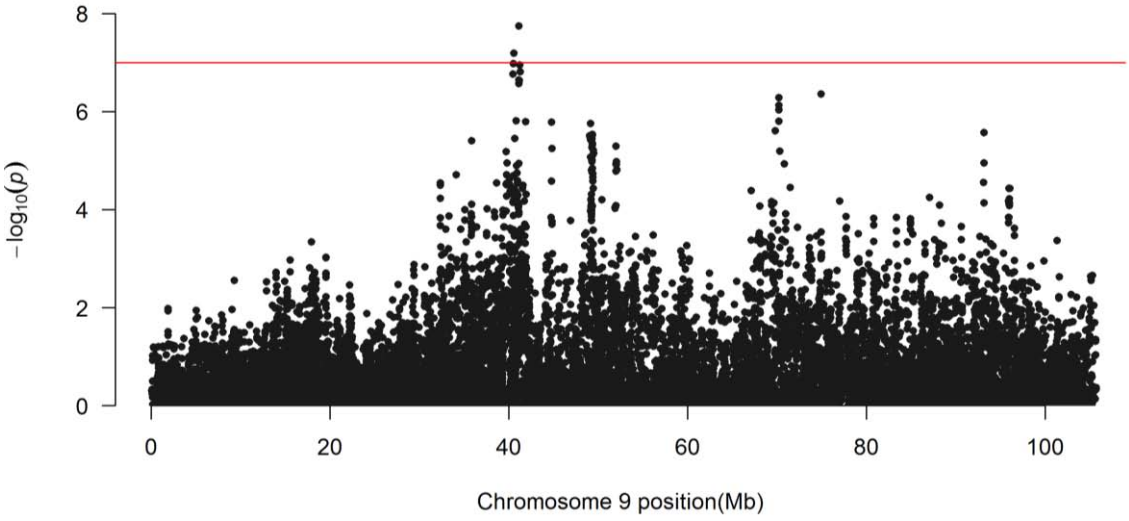

Figure 36: Amarela/ Vermelha Chitada Manhattan Plot Chr. 9

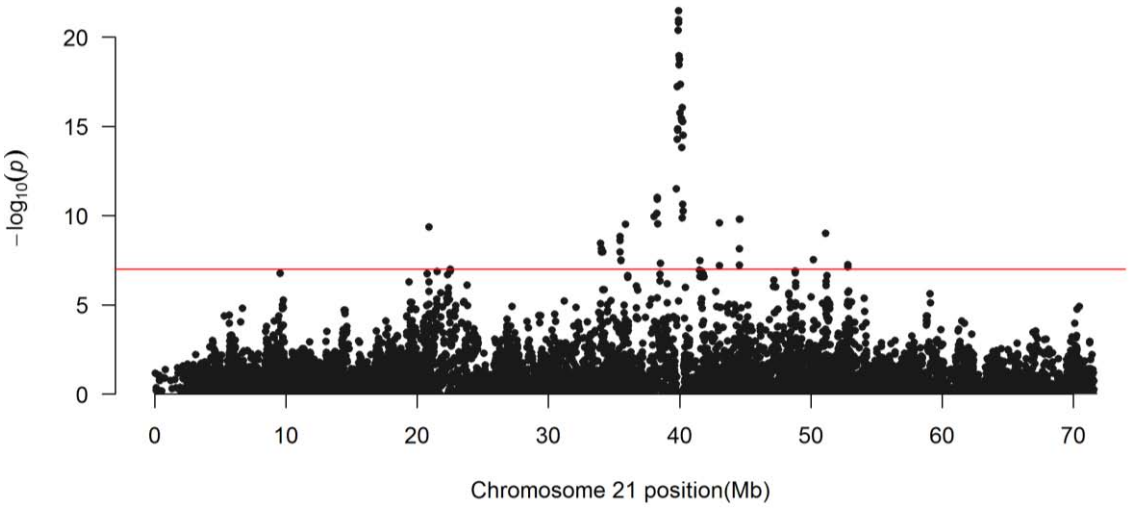

Figure 37: Amarela/ Vermelha Chita Manhattan Plot Chr. 21

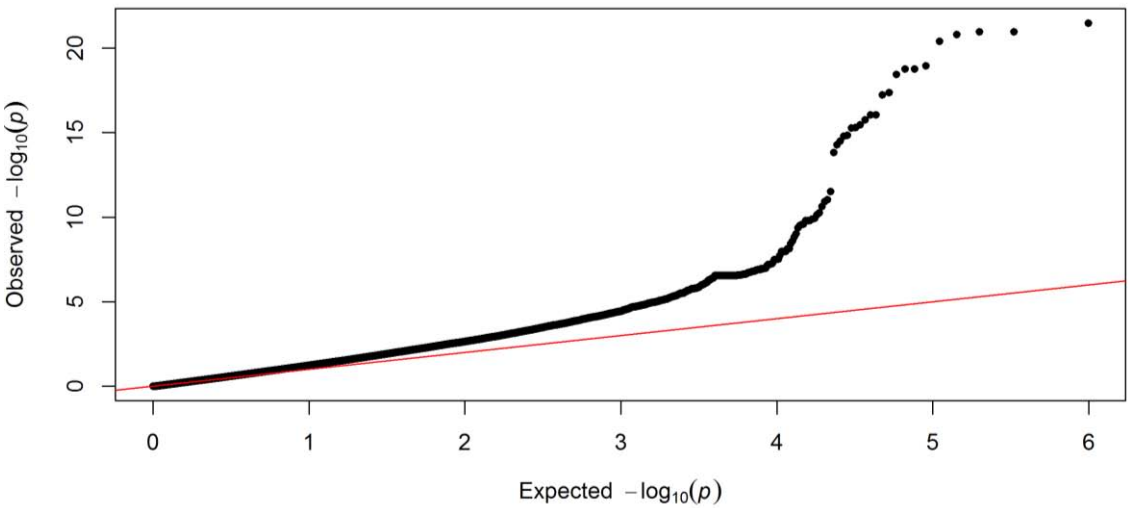

Figure 38: Amarela/ Vermelha Chitada QQ-Plot

Chita Clara vs. Chita de Amarela and Chita de Vermelha

The comparisons of the groups “Chita Clara” (á 267) and “Chita de Amarela/Vermelha” (á 452) showed the same peaks on Chromosome 21 and six. In “Chita de Amarela/Vermelha” additional small peaks in Chromosome six (~14.000.000 Bp), 18 (~15.000.000 Bp), 21 (~10.000.000) came up. As these groups had a similar genotype but differed in the phenotype there was a comparison done between them. In this comparison no SNP passed the significance line, check against Figure 39. Figure 40 shows the related QQ-plot, which shows that the groups were nearly evenly distributed.

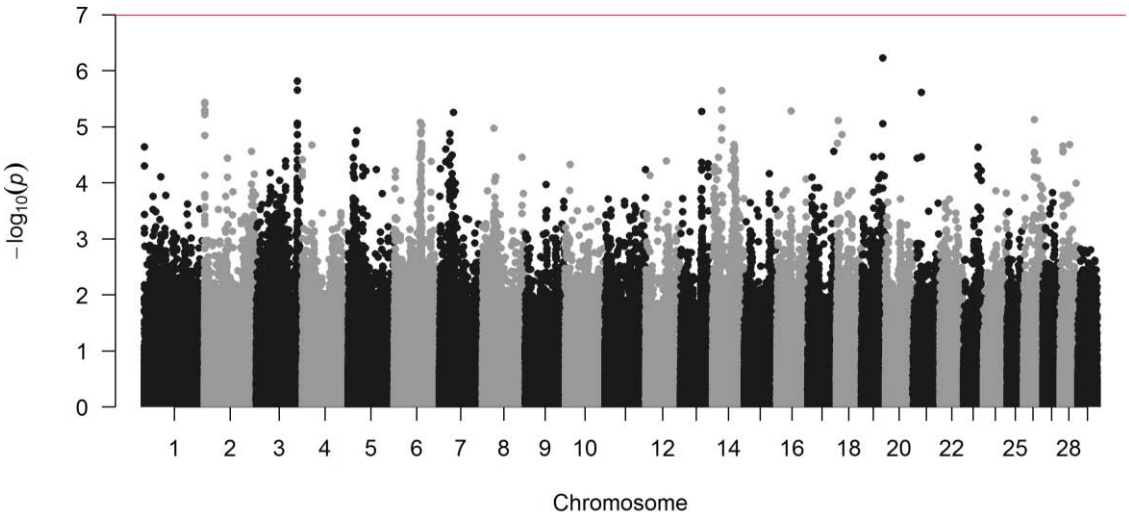

Figure 39: Chita Clara vs. Chita de Amarela/ Vermelha Manhattan Plot

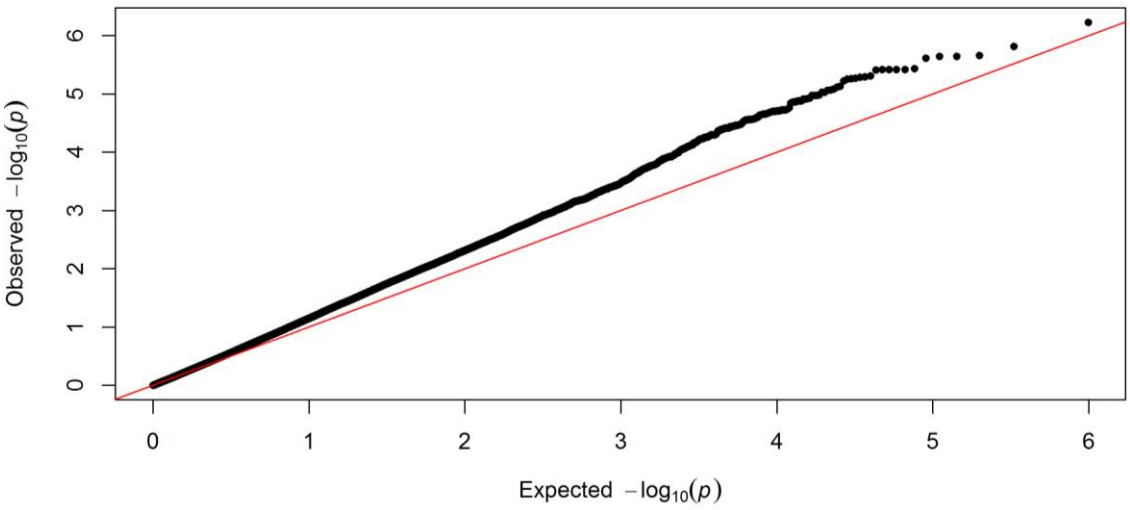

Figure 40: Chita Clara vs. Chita de Amarela/ Vermelha QQ-Plot

Chita de Amarela/ Vermelha/ Clara vs. Amarela/ Vermelha Chitada

The decision to classify an animal as “Chita de Amarela” or “Amarela Chitada” can sometimes be hard especially for people who are not trained for classifying. The same one is valid for “Chita de Vermelha” and “Vermelha Chitada”. That is why a comparison between these groups was calculated. As there was no difference between “Chita Clara” and “Chita de Amarela” and “Chita de Vermelha” observed, these three groups were put together (719 head of cattle) and compared to the groups “Amarela Chitada” and “Vermelha Chitada” (805 head of cattle).

Figure 41 display peaks on four Chromosomes. To visualize the peaks better Manhattan plots with zooms for the concerning Chromosomes were created.

The region found on figure 42 on Chromosome six around 70.000.000 Bp is highly significant, the most significant SNP has a p-value of $4{,}93*10^{-38}$. On Chromosome 13 and 16 one SNP passed the significance line at 43.000.000 and 12.000.000 Bp respectively. Figure 43 shows the Manhattan plot for Chromosome 13. The peak of Chromosome 16 can be observed on Figure 44. In the significant region on Chromosome 18 MC1R is located, but the region is broader than in the other results as one can see in Figure 45. There occur significant SNPs in the range from 8.000.000 to 15.000.000 Bp.

The big difference between calculated and expected p-values on the QQ-plot in Figure 46 underlines the strong peaks.

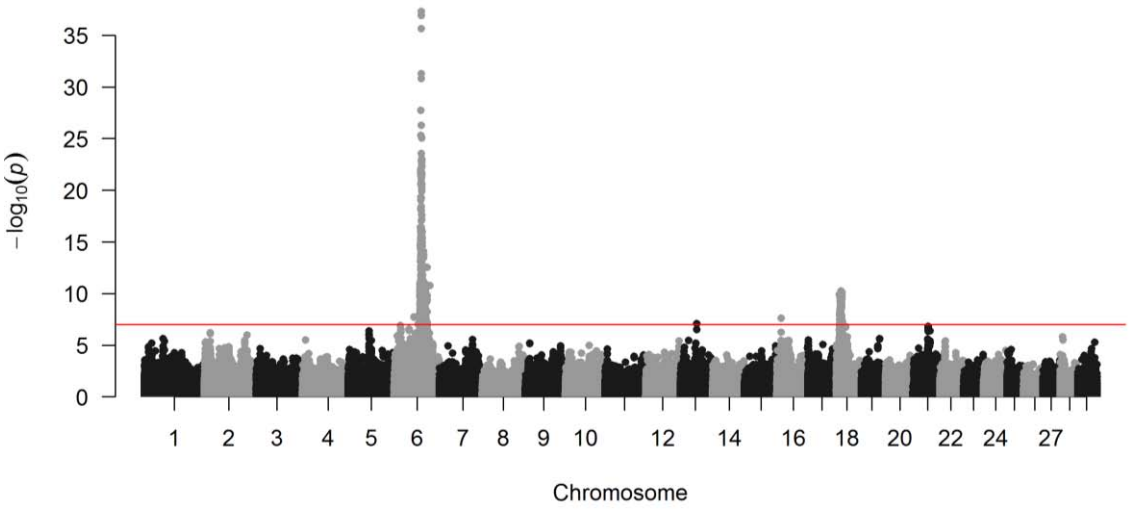

Figure 41: Chita de Amarela/ Vermelha/ Clara vs. Amarela/ Vermelha Chitada Manhattan Plot

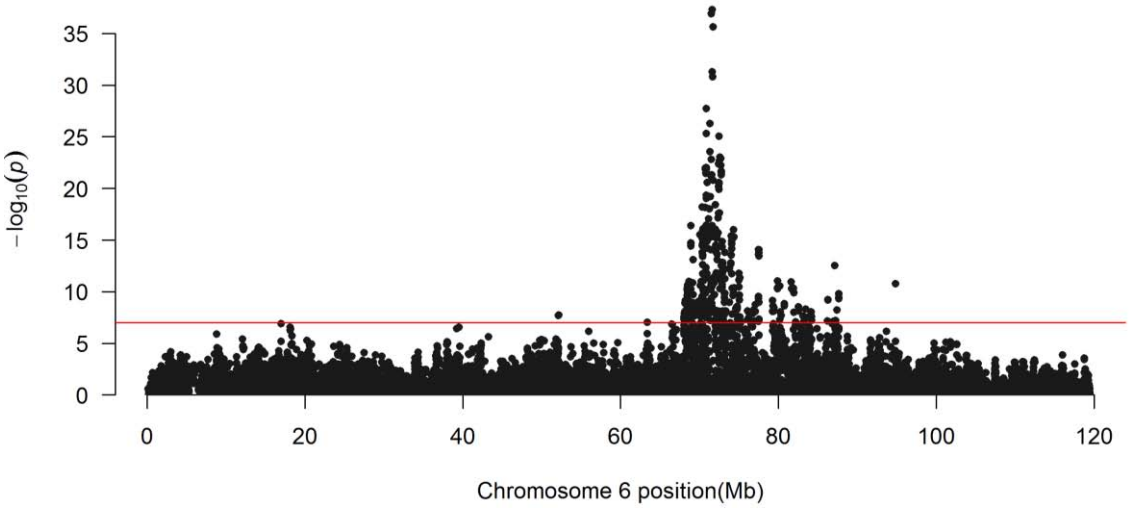

Figure 42: Chita de Amarela/ Vermelha/ Clara vs. Amarela/ Vermelha Chitada Manhattan Plot Chr. 6

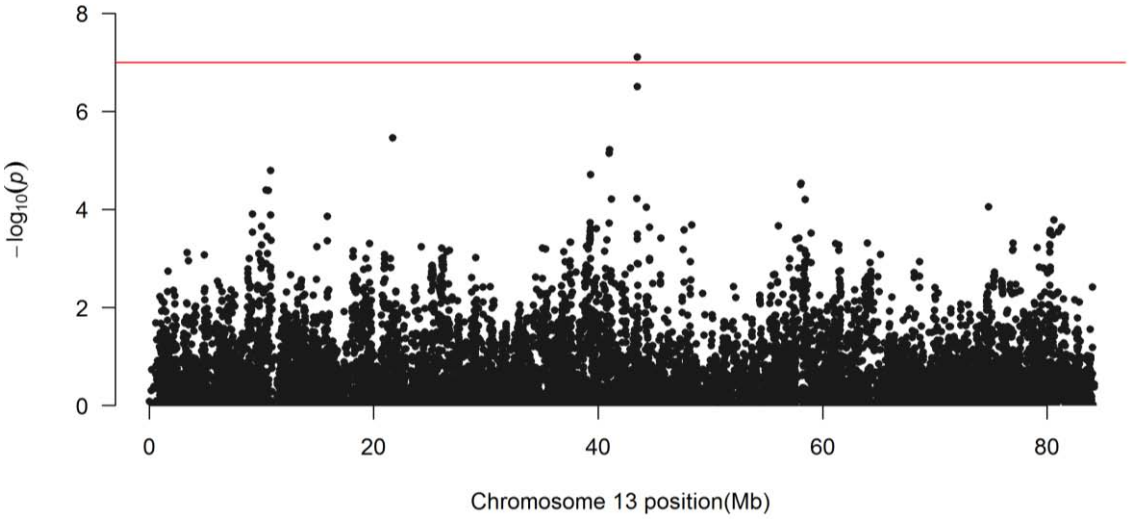

Figure 43: Chita de Amarela/ Vermelha/ Clara vs. Amarela/ Vermelha Chitada Manhattan Plot Chr. 13

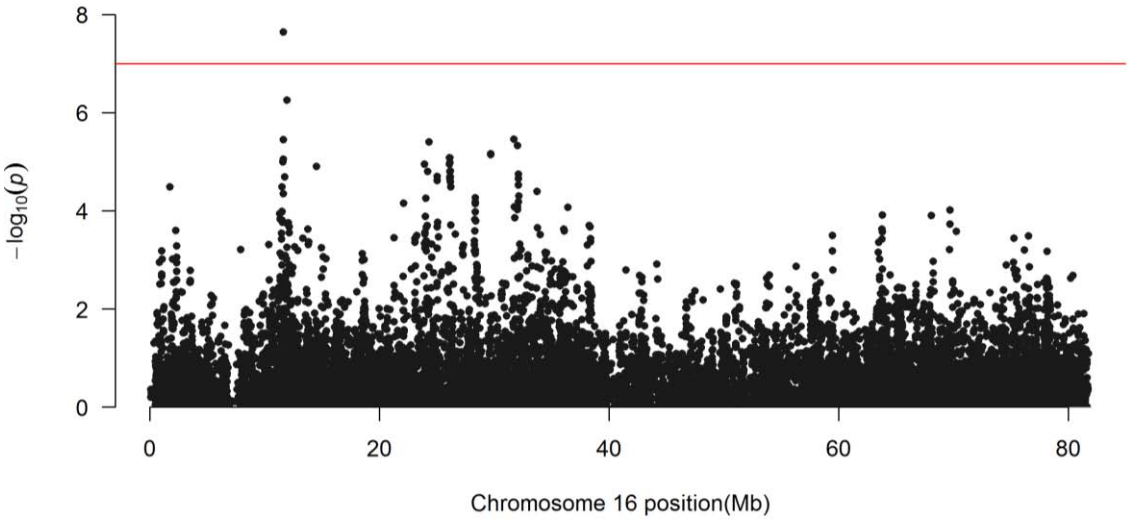

Figure 44: Chita de Amarela/ Vermelha/ Clara vs. Amarela/ Vermelha Chitada Manhattan Plot Chr. 16

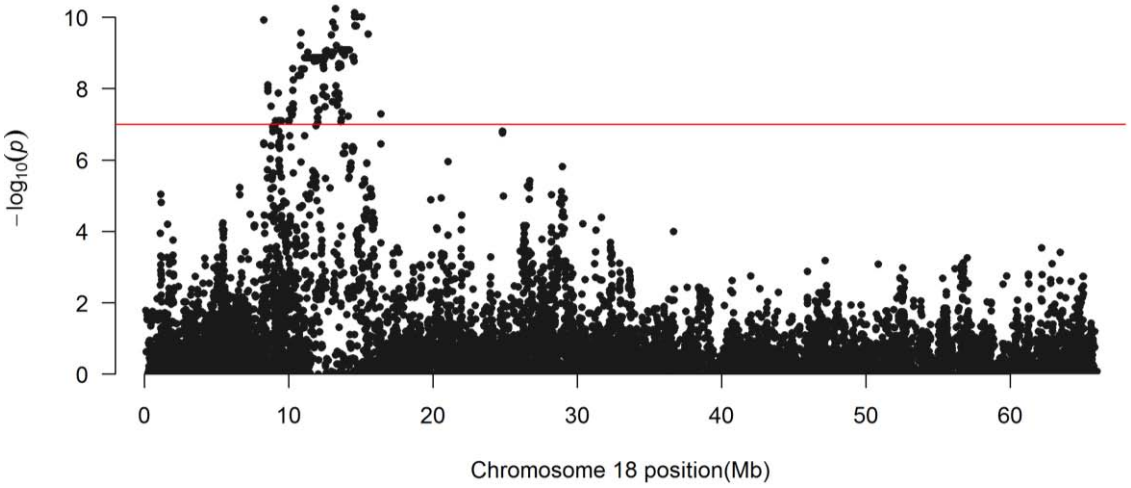

Figure 45: Chita de Amarela/ Vermelha/ Clara vs. Amarela/ Vermelha Chitada Manhattan Plot Chr. 18

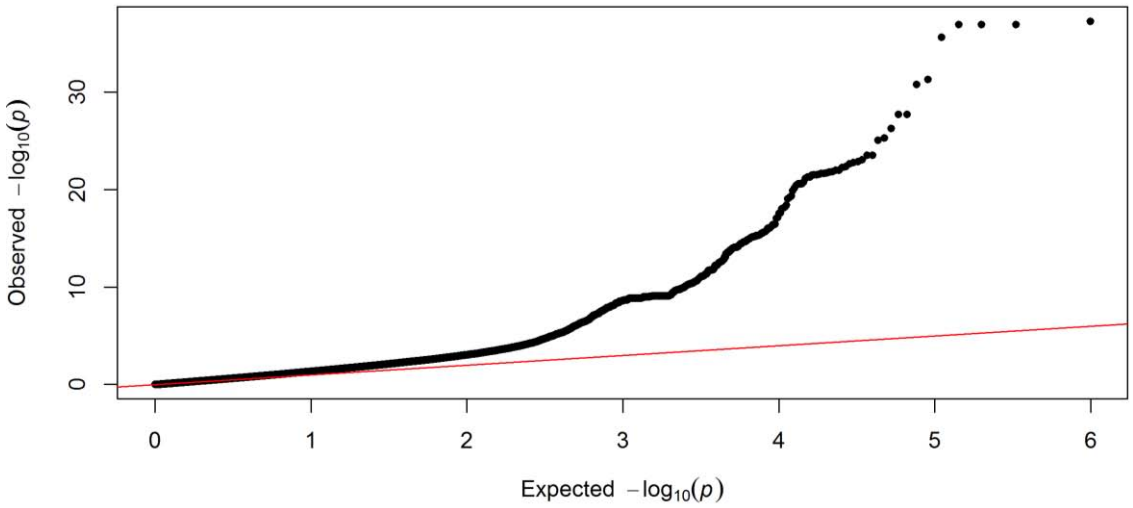

Figure 46: Chita de Amarela/ Vermelha/ Clara vs. Amarela/ Vermelha Chitada QQ-Plot

Animals with necklace

Gargantilha

There was only one animal classified as “Amarela Gargantilha” that built together with the 246 animals named “Vermelha Gargantilha” the group “Gargantilha”. The group was compared with the 220 animals out of the control-group. A Manhattan plot for the whole genome is shown in Figure 47. The peak at 40.000.000 Bp on Chromosome 21, seen in Figure 48 does not accommodate a known gene. Figure 49 provides the QQ-Plot for the “Gargantilha” group, which is characterised by big differences in calculated and observed p-values.

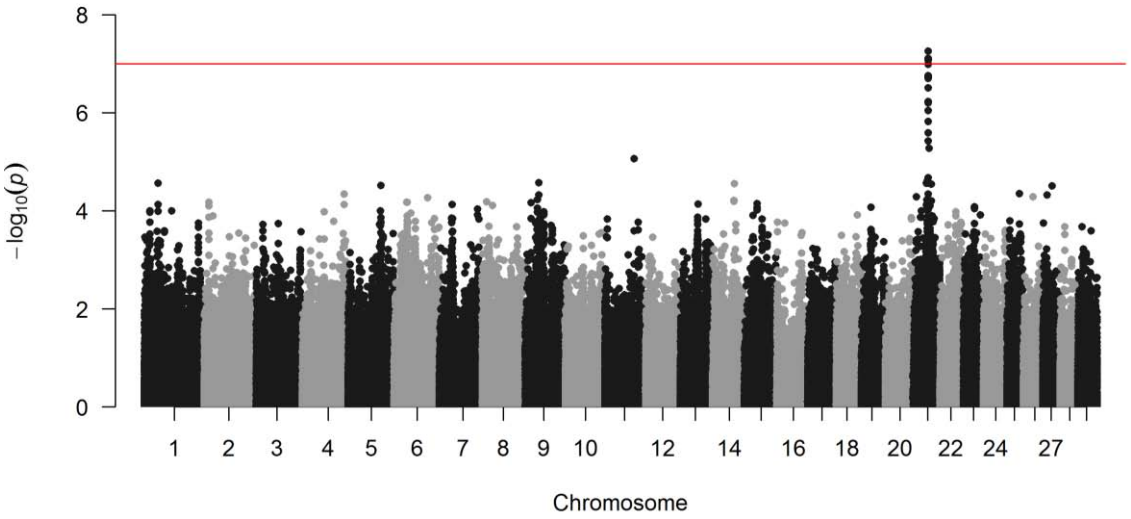

Figure 47: Gargantilha Manhattan Plot

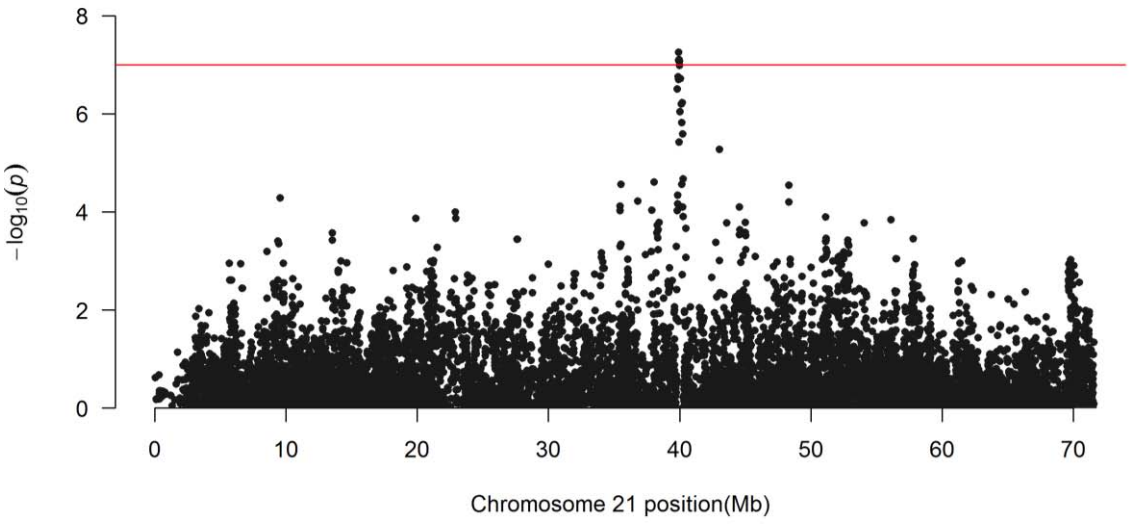

Figure 48: Gargantilha Manhattan Plot Chr. 21

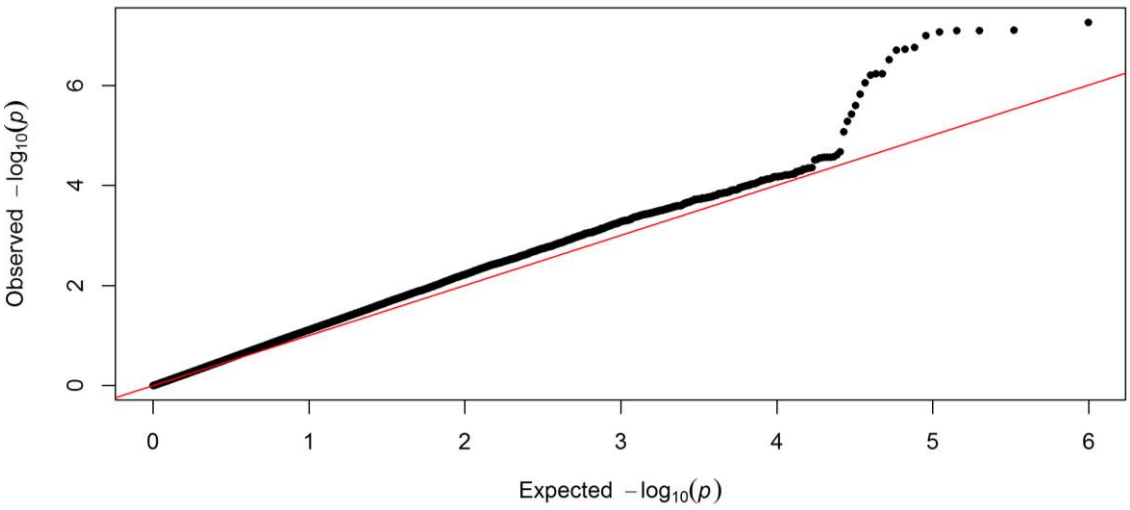

Figure 49: Gargantilha QQ-Plot

Discussion

Methodology

There were highly significant regions on the Chromosomes six, eighteen and twenty one. Also, some small peaks with only few significant SNPs were observed on Chromosomes six, nine, sixteen, eighteen and twenty one.

The small peaks did not show regions where genes related to pigmentation are located. The reasons for that could be the sometimes strongly differing sample sizes. When there are only few animals in one group and the other one is very big, a difference might not be observed or show a clear peak. That could have happened for instance in the comparison between “Moura Escura” and the control-group, as they differ strongly in the Phenotype, but in this study no difference in the genotype was found.

Maybe the small peaks with SNPs above the significance line but with comparatively high p-value occur due to the imputing and there would not be a SNP when 50K SNP Data was used.
It is also conceivable that there is no SNP in a significant region and therefore no difference between two groups found.

Still there were very clear and highly significant peaks found, suggesting that the genes in that regions are responsible for the different coat colour patterns of the Gir cattle breed in Brazil.

Animals with dark skin

The gene MC1R is located on Chromosome 18 it starts at 14757332 Bp and ends with Bp 14759082. In this region significant SNPs were found for the animals with the phenotype ”Moura”. According to the gene databank Ensembl.org the gene MC1R is reported to be responsible for coat colour extension. HANNA et al. did a study on cattle crossbreed between Bos Taurus (Angus) and Bos indicus (Nellore) and found indices that there might be genes (i.e. PDGFRA) interacting with MC1R and being responsible for different forms of black coat colour in cattle with Bos indicus influence (Hanna et al. 2014). Different haplotypes of MC1R also cause different coat colour expression in highland cattle (Schmutz and Dreger 2013). MC1R is also responsible for coat colour patterns in other mammals like rabbits (Fontanesi et al. 2010) and plays a role in the darkness of the skin in human (García-Borrón, Abdel-Malek, and Jiménez-Cervantes 2014).

The comparisons with the “Moura” animals also showed significant SNPs at 10.8 Mb where the gene COTL1 (Chr. 18; Start:10781076 Bp, End: 10825733 Bp) is located. This gene was not found to be related to coat colour, but it is discussed to play a role in autoimmune disorders in human (Jin et al. 2009). The reason why it occurred in this study can be one of the reasons discussed in the methodology chapter. Also, a gene nearby could play a role but there were no indicators for that found.

A problem especially in this group is the small sample size and due to that reason the strongly differing sample size compared to the control group. The Group “Moura Escura” was made up by twelve animals, 34 animals were classified as “Moura Clara” and only four animals showed the phenotype “Moura de Vermelha”. The control-group had 220 heads, more than four times as many animals.

Sprinkled animals

With altogether 1524 animals the sprinkled animals built up the biggest group and reliable results can be expected. The different categories of sprinkled animals were compared to the monochrome animals and ere also compered among each other.

There were highly significant peaks in the Manhattan plots on Chromosome six. The found regions differed in wideness, but they all appeared around 71.000.000 Bp where the most significant SNP for every group lied. That is where the gene KIT is located. KIT is a gene on Chromosome six, that starts at 71796318 Bp and ends with Bp 71917431, which is responsible for dominant white coat colour (Ensembl.org 01.10.2015). The description fits with the findings in this study as all the groups with a white base colour (“Chita Clara”, “Chita de Amarela” and “Chita de Vermelha”) show a clear peak at the aforesaid region.

Many mammals show colour patterns due to allelic diversity of KIT. Different haplotypes of this gene are reported to be responsible for white colour in pigs (Pielberg et al. 2002). Also the different spotting of the Checkered Giant rabbit is caused by variation of KIT. The role of the gene in coat colour is also known in cattle. FONTANESI et al. found that KIT is likely to play a role in the coat colour expression of Hereford cattle (L. Fontanesi et al. 2010). The role of this gene in the degree of spotting of Holstein and Simmental cattle was already discussed in 1999 (Reinsch et al. 1999). In Fleckvieh (dual-purpose Simmental) Kit was found to be strongly associated with different sorts of spots in the face (Mészáros et al. 2015).

A peak in Chromosome 21 was found in every comparison of sprinkled and monochrome animals. According to ensemble.org there is no gene around the most significant SNPs of each comparison located on Chromosome 21. In the

area around hosting the less significant SNPs there was no gene found, that is related to pigmentation. At the small peak in the group “Chita de Amarela” and “Chita de Vermelha” (9553854 Bp) as well as in the group “Vermelha Chitada” and “Amarela Chitada” (20881920 Bp) is no gene located. DURKIN et al. found that the translocation of a segment from Chromosome six including KIT to Chromosome 29 is responsible for colour sidedness in Belgian blue and Brown Swiss cattle (Durkin et al. 2012). The same effect could be conceivable in the sparkling in Gir.

One SNP on Chromosome 22 passed the significance line at Bp 54017830 in the comparison between “Chita Clara” and the control group. There the genes CCCR6 and CCR9 were found. These genes are currently not discussed to be responsible for colour but for immune traits such as resistance to nematodes (Araujo et al. 2009).

MC1R was found again in the comparison between “Chita de Amarela” and “Chita de Vermelha” against the control-group. But the region was broader and hosted more significant SNPs.

In the animals with a reddish or yellow base colour and white sprinkles a significant region on Chromosome nine was found. Ensembl.org provided information that the gene FIG4 located there is related to pigmentation. But there was no literature found that connects FIG4 and pigmentation in cattle.
No gene is located on the peak of Chromosome 16 that came up in the comparison of the sprinkled animals with different base colours.
On Chromosome 13 the gene ASB13 was found to be significant, according to ensemble.org it is responsible for intercellular mechanisms.

Animals with necklace

Five SNPs passed the significance line in the comparison between “Gargantilha” and the monochrome animals. They are located at nearly 40.000.000 Bp on Chromosome 21. It is the same region that was found for the sprinkled animals. These animals have sprinkles on their dewlap and a copy of KIT similar to the one in Brown Swiss and Belgian blue found by is DURKIN et al. is possible (Durkin et al. 2012).

The same phenotypic expression as “Gargantilha” and “Chita” in Gir is reported in Shorthorn and Belgian Blue cattle. Some animals, that were phenotyped as monochrome, were genotyped as roan. A closer look at those animals revealed that they had spots on the dewlap. The gene MGF from Chromosome 5 was found to be responsible for that case (Seitz et al. 1999). In this study no significant SNP was found on Chromosome five.

Conclusion

In this study, several very strong signals were found for genes or genomic regions responsible for particular coat colour and pattern phenotypes in Gir cattle.

A significant region for dark skin was found in this study on chromosome 18. Nine SNPs passed the significance line in the region of 11 Mb and 15Mb. The gene MC1R is located there. According to literature it is a strong candidate for being responsible for the dark skin as it is linked to pigmentation (ensemble.org). The group of animals that have dark skin contained the three subcategories “Moura Escura”, “Moura Clara” and “Moura de Vermelha”. There were no genetic differences between those clearly differing phenotypes observed. The reason for that might be the small sample size of 50 animals with dark skin.

Several peaks came up when the different groups of sprinkled animals where compared to the control-group or among each other. There were some lowly significant regions on Chromosome 13, 16 and 22 shown in the Manhattan plots, but these regions were not found to be connected to genes linked to coat colour in previous studies. There were also some very highly significant regions found for the sprinkled animals. The animals with white base colour and sprinkles of different degrees showed to highly significant peaks on Chromosome six and 21. The most significant SNPs on Chromosome six lie at 71Mb and lie in the region of the gene KIT which has already been reported to be responsible for spotted coat colour expression in mammals (Mészáros et al. 2015) and (Luca Fontanesi et al. 2014). There was a highly significant region observed on Chromosome 21, located around 40 Mb, but there was no gene related to colour found in that region. The animals with yellow or red base colour and white sprinkles do not show the peak related to KIT, but as well show the significant region on Chromosome 21.

The animals with sprinkled dewlap show a less significant peak than the spotted animals in the same region of Chromosome 21. A conceivable explanation for this could be a translocation of the KIT gene (Durkin et al. 2012).

Different comparisons with different grouping of the animals showed some variants of combinations of the found regions. It is conceivable that the genes interact and different combinations of them are responsible for variation in phenotypes.

To gain more knowledge about the coat colour patterns in Gir further studies on basis of whole genome sequences are recommended. That way the haplotypes responsible for the different expression of colour patterns can be found and the region on Chromosome 21 may be examined.

Acknowledgements

I want to thank my Supervisors Johann, Marcos, Adam and Solomon for supporting me, helping me and give me some of their time whenever I had a question.
Thanks for the nice time to everyone from the master student office, especially Daniel for always having an open ear for me.
Also I want to thank my family for giving me the possibility to study.

List of Figures

The pictures figure 1, 2 4-12 comes from the book Zavala, a vaca multicolorida (see reference list)

The picture figure 3 comes from the homepage of Ambiente de Apoio a Educação Científica e Tecnológica do IFC - Campus Camboriu visited on the 25.09.2014.

List of Tables

List of Abbreviations

Chr.	Chromosome
vs.	Versus
à	Quantifier
Bp	Basepair
Mb	Megabase

References

Araujo, Ricardo N, Terezinha Padilha, Dante Zarlenga, Tad Sonstegard, Erin E Connor, Curt Van Tassel, Walter S Lima, Evaldo Nascimento, and Louis C Gasbarre. 2009. “Use of a Candidate Gene Array to Delineate Gene Expression Patterns in Cattle Selected for Resistance or Susceptibility to Intestinal Nematodes.” *Veterinary Parasitology* 162 (1-2): 106–15. doi:10.1016/j.vetpar.2008.12.017.

Boison, S a, D J a Santos, a H T Utsunomiya, R Carvalheiro, H H R Neves, a M Perez O’Brien, J F Garcia, J Sölkner, and M V G B da Silva. 2015. “Strategies for Single Nucleotide Polymorphism (SNP) Genotyping to Enhance Genotype Imputation in Gyr (Bos Indicus) Dairy Cattle: Comparison of Commercially Available SNP Chips.” *Journal of Dairy Science* 98 (7): 4969–89. doi:10.3168/jds.2014-9213.

Brenig, B, J Beck, C Floren, K Bornemann-Kolatzki, I Wiedemann, S Hennecke, H Swalve, and E Schütz. 2013. “Molecular Genetics of Coat Colour Variations in White Galloway and White Park Cattle.” *Animal Genetics* 44 (4): 450–53. doi:10.1111/age.12029.

Dorshorst, Ben, Corneliu Henegar, Xiaoping Liao, Markus Sällman Almén, Carl-Johan Rubin, Shosuke Ito, Kazumasa Wakamatsu, et al. 2015. “Dominant Red Coat Color in Holstein Cattle Is Associated with a Missense Mutation in the Coatomer Protein Complex, Subunit Alpha (COPA) Gene.” *PloS One* 10 (6): e0128969. doi:10.1371/journal.pone.0128969.

Durkin, Keith, Wouter Coppieters, Cord Drögemüller, Naima Ahariz, Nadine Cambisano, Tom Druet, Corinne Fasquelle, et al. 2012. “Serial Translocation by Means of Circular Intermediates Underlies Colour Sidedness in Cattle.” *Nature* 482 (7383): 81–84. doi:10.1038/nature10757.

Fontanesi, L., M. Tazzoli, V. Russo, and J. Beever. 2010. “Genetic Heterogeneity at the Bovine KIT Gene in Cattle Breeds Carrying Different Putative Alleles at the Spotting Locus.” *Animal Genetics* 41 (3): 295–303. doi:10.1111/j.1365-2052.2009.02007.x.

Fontanesi, Luca, Emilio Scotti, Michela Colombo, Francesca Beretti, Lionel Forestier, Stefania Dall’Olio, Séverine Deretz, Vincenzo Russo, Daniel Allain, and Ahmad Oulmouden. 2010. “A Composite Six Bp in-Frame Deletion in the Melanocortin 1 Receptor (MC1R) Gene Is Associated with the Japanese Brindling Coat Colour in Rabbits (Oryctolagus Cuniculus).” *BMC Genetics* 11 (January): 59. doi:10.1186/1471-2156-11-59.

Fontanesi, Luca, Emilio Scotti, and Vincenzo Russo. 2010. “Analysis of SNPs in the KIT Gene of Cattle with Different Coat Colour Patterns and Perspectives to Use These Markers for Breed Traceability and Authentication of Beef and Dairy Products.” *Italian Journal of Animal Science* 9 (2): 217–21. doi:10.4081/ijas.2010.e42.

Fontanesi, Luca, Manuela Vargiolu, Emilio Scotti, Rocco Latorre, Maria Simonetta Faussone Pellegrini, Maurizio Mazzoni, Martina Asti, et al. 2014. “The KIT Gene Is Associated with the English Spotting Coat Color Locus and Congenital Megacolon in Checkered Giant Rabbits (Oryctolagus Cuniculus).” *PloS One* 9 (4): e93750. doi:10.1371/journal.pone.0093750.

García-Borrón, Jose C, Zalfa Abdel-Malek, and Celia Jiménez-Cervantes. 2014. “MC1R, the cAMP Pathway, and the Response to Solar UV: Extending the Horizon beyond Pigmentation.” *Pigment Cell & Melanoma Research* 27 (5): 699–720. doi:10.1111/pcmr.12257.

Hanna, Lauren L Hulsman, James O Sanders, David G Riley, Colette a Abbey, and Clare a Gill. 2014. “Identification of a Major Locus Interacting with MC1R and Modifying Black Coat Color in an F_2 Nellore-Angus Population.” *Genetics, Selection, Evolution : GSE* 46 (1). Genetics Selection Evolution: 4. doi:10.1186/1297-9686-46-4.

Hayes, Ben J, Jennie Pryce, Amanda J Chamberlain, Phil J Bowman, and Mike E Goddard. 2010. “Genetic Architecture of Complex Traits and Accuracy of Genomic Prediction: Coat Colour, Milk-Fat Percentage, and Type in Holstein Cattle as Contrasting Model Traits.” *PLoS Genetics* 6 (9): e1001139. doi:10.1371/journal.pgen.1001139.

Jin, Eun-Heui, Seung-Cheol Shim, Hwan-Gyu Kim, Soo-Cheon Chae, and Hun-Taeg Chung. 2009. “Polymorphisms of COTL1 Gene Identified by Proteomic Approach and Their Association with Autoimmune Disorders.” *Experimental and Molecular Medicine* 41 (5): 354. doi:10.3858/emm.2009.41.5.040.

Mészáros, Gábor, Elisabeth Petautschnig, Hermann Schwarzenbacher, and Johann Sölkner. 2015. “Genomic Regions Influencing Coat Color Saturation and Facial Markings in Fleckvieh Cattle.” *Animal Genetics* 46 (1): 65–68. doi:10.1111/age.12249.

Olson, T A. 1999. “Genetics of Colour Variation.” In *The Genetics of Cattle*, 33–54.

Otàvio Lemos, Jòse. 2012. *Zavala, a Vaca Multicolorida*. Edited by Jòse Otàvio Lemons. 1a Edicao. Rio de Janeiro: Jol Editora.

Pielberg, G, C Olsson, A Syva, and L Andersson. 2002. “Unexpectedly High Allelic Diversity at the KIT Locus Causing Dominant White Color in the Domestic Pig” 311 (January): 305–11.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel a R Ferreira, David Bender, Julian Maller, et al. 2007. “PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses.” *American Journal of Human Genetics* 81 (3): 559–75. doi:10.1086/519795.

Reinsch, N, H Thomsen, N Xu, M Brink, C Looft, E Kalm, G A Brockmann, et al. 1999. “A QTL for the Degree of Spotting in Cattle Shows Synteny with the KIT Locus on Chromosome 6”, 629–34.

Santana, M.L., R.J. Pereira, a.B. Bignardi, L. El Faro, H. Tonhati, and L.G. Albuquerque. 2014. “History, Structure, and Genetic Diversity of Brazilian Gir Cattle.” *Livestock Science* 163 (May). Elsevier: 26–33. doi:10.1016/j.livsci.2014.02.007.

Sargolzaei, Mehdi, Jacques P Chesnais, and Flavio S Schenkel. 2014. “A New Approach for Efficient Genotype Imputation Using Information from Relatives.” *BMC Genomics* 15 (1): 478. doi:10.1186/1471-2164-15-478.

Schmutz, Sheila M. 2012. “Genetics of Coat Color in Cattle.” In *Bovine Genomics*, edited by James Womack, 20–33. Iowa: Wiley-Blackwell.

Schmutz, Sheila M, and Dayna L Dreger. 2013. “Interaction of MC1R and PMEL Alleles on Solid Coat Colors in Highland Cattle.” *Animal Genetics* 44 (1): 9–13. doi:10.1111/j.1365-2052.2012.02361.x.

Seitz, Jeffrey J, Sheila M Schmutz, Tracey D Thue, and Fiona C Buchanan. 1999. “A Missense Mutation in the Bovine MGF Gene Is Associated with the Roan Phenotype in Belgian Blue and Shorthorn Cattle” 712: 710–12.

Turner, Stephen D. 2014. “Qqman:an R Package for Visualizing GWAS Results Using Q-Q and Manhattan Plots.” doi:10.1086/519795.