

Article

Reverse Genetic Screen for Deleterious Recessive Variants in the Local Simmental Cattle Population of Switzerland

Irene M. Häfliger ¹, Franz R. Seefried ² and Cord Drögemüller ^{1,*}

¹ Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch

² Qualitas AG, Chamerstrasse 56, 6300 Zug, Switzerland; franz.seefried@qualitasag.ch

* Correspondence: cord.droegemueller@vetsuisse.unibe.ch; Tel.: +41-316842529

Simple Summary: Today's Swiss Simmental represents a local dual-purpose breed of cattle. Within closed populations, deleterious variants can reach problematic frequencies, explaining substantial proportions of inbreeding depression. Depletions in homozygous genotypes for certain haplotypes among large cohorts of animals genotyped for the purpose of genomic selection is a widely used approach to pinpoint undesired recessive alleles. In the course of a reverse genetic screen, we aimed to identify single recessive Mendelian variants that potentially affect fertility and rearing success without any phenotypic information available. We detected eleven genome regions showing obvious depletion of homozygosity based on genome-wide SNP data. Furthermore, after performing whole-genome sequencing of selected animals, we propose three candidate causative variants affecting different genes with possibly detrimental effects for embryonic development. The established haplotypes, as well as the identified protein-changing variants, can be directly implemented into breeding practice to avoid the risk of mating carriers and thereby increase breeding success.



check for updates

Citation: Häfliger, I.M.; Seefried, F.R.; Drögemüller, C. Reverse Genetic Screen for Deleterious Recessive Variants in the Local Simmental Cattle Population of Switzerland. *Animals* **2021**, *11*, 3535. <https://doi.org/10.3390/ani11123535>

Academic Editor: Michael E. Davis

Received: 4 November 2021

Accepted: 8 December 2021

Published: 12 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: We herein report the result of a large-scale reverse genetic screen in the Swiss Simmental population, a local dual-purpose cattle breed. We aimed to detect possible recessively inherited variants affecting protein-coding genes, as such deleterious variants can impair fertility and rearing success significantly. We used 115,000 phased SNP data of almost 10 thousand cattle with pedigree data. This revealed evidence for 11 genomic regions of 1.17 Mb on average, with haplotypes (SH1 to SH11) showing a significant depletion in homozygosity and an allele frequency between 3.2 and 10.6%. For the proposed haplotypes, it was unfortunately not possible to evaluate associations with fertility traits as no corresponding data were available. For each haplotype region, possible candidate genes were listed based on their known function in development and disease. Subsequent mining of single-nucleotide variants and short indels in the genomes of 23 sequenced haplotype carriers allowed us to identify three perfectly linked candidate causative protein-changing variants: a SH5-related *DIS3*:p.Ile678fs loss-of-function variant, a SH8-related *CYP2B6*:p.Ile313Asn missense variant, and a SH9-related *NUBPL*:p.Ser143Tyr missense variant. None of these variants occurred in homozygous state in any of more than 5200 sequenced cattle of various breeds. Selection against these alleles in order to reduce reproductive failure and animal loss is recommended.

Keywords: cattle; *Bos taurus*; reproduction; breeding; fertility; embryonic lethality; loss-of-function variants; whole-genome sequencing; SNP genotyping



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Simmental is a globally recognized cattle breed, originating in the Simmental valley in the canton Bern in Switzerland. Autochthonous of Switzerland, today up to 40 million cattle worldwide are designated as members of the Simmental breed. However, Simmental are characterized by very different local breeding objectives (<http://wsff.info>; accessed on 24 October 2021). In the 19th century, crossbreeding of local cattle with Simmental cattle exported from Switzerland led to Fleckvieh populations in neighboring

countries, including Austrian and German Fleckvieh, French Montbeliarde, and Italian Pezzata Rossa cattle. In the 20th century, introduction of animals of the Holstein breed into the Swiss Simmental population led to today's Fleckvieh population in Switzerland (Swiss Fleckvieh). A recent analysis of the population structure of Swiss cattle showed a clear differentiation between today's Swiss Simmental cattle and all other Swiss cattle populations, including the modern Swiss Fleckvieh [1,2]. Various historically younger Central European Simmental populations that descended from the Swiss Simmental can also be clearly distinguished from today's Swiss Simmental animals, which showed the highest inbreeding level [3]. In contrast, recent studies found a comparatively low degree of genomic inbreeding in purebred Swiss Simmental that might be explained by the continued use of natural service sires, which is likely the major reason for their remarkably high level of genetic diversity although this population has been closed for a long time [1,2]. Therefore, the current so-called Original Simmental breed of Switzerland represents a unique purebred population. At the end of 2020, more than 23 thousand dual-purpose animals in the Simmental population were registered in the Swiss herdbook (https://www.swissherdbook.ch/fileadmin/Domain1/PDF_Dokumente/05-Statistiken-Formulare/53-Jahresstatistik/1_Wichtigste_Zahlen/D_sh_JS_2020_HBZahlen_web.pdf; accessed on 24 October 2021).

A universal problem in cattle breeding is reproductive failure. It was shown that the reproduction success is negatively associated with production traits [4,5]. These effects have been thoroughly studied and possible reasons are selection programs focusing on production traits coupled with negative correlation with reproductive traits [6]. Regarding early pregnancy loss in Simmental cattle, e.g., a Croatian study described the body condition score, parity, and milk yield as important influencing factors [7]. Lower reproductive success is an economic problem for farmers and a major reason for cattle slaughter [8,9]. To tackle fertility issues, one approach is to apply a so-called reverse genetic screen, where only genomic data is used rather than phenotypic information. A whole-genome sequencing (WGS)-based approach was suggested by Charlier et al. (2016), where sequencing data of entire genomes is mined for variants affecting fertility by causing embryonic lethality [10]. A similar approach proposes scanning the cumulative population-wide SNP genotyping data from genomic selection programs to identify haplotype regions indicative of a depletion in homozygosity [11]. The analysis includes the statistical evaluation of segregating haplotypes regarding the Hardy–Weinberg equilibrium (HWE) [11]. Such genomic regions harboring deficient homozygous haplotypes indicate potential causal variants for embryonic death or unwanted phenotypes in newborns, leading to the exclusion of homozygous animals from breeding programs [11]. This SNP-based reverse genetics approach has been used successfully in diverse breeds and species (e.g., [11–31]). Haplotypes never observed in homozygous state indicate the presence of recessive, predominantly embryonic lethal variants. Potential causative variants were detected by analyzing whole-genome or whole-exome sequencing data linked to the identified haplotypes (e.g., [10,19–30,32–35]). In general, these approaches are especially interesting for small populations where phenotypic information is sparse.

In Simmental-derived breeds in Central Europe, eight causal protein-changing variants for recessively inherited disorders are known so far. In German Fleckvieh cattle, a SNP-based reverse genetic study identified four deficient homozygous haplotypes located on chromosomes 1, 10, and 12 and proposed two candidate causal variants in two genes: the FH2-related frameshift variant in *SLC2A2* and the FH4-related missense variant in *SUGT1* [21]. In addition, a fifth haplotype is described, affecting the calf survival due to a congenital heart failure and severe liver damage (<https://www.lfl.bayern.de/itz/rind/12227/index.php>; accessed on 24 October 2021). Based on forward genetic studies, where affected individuals were examined, several causal variants for Mendelian disorders were found. A frameshift variant in *GON4L* was associated with the autosomal recessively inherited disorder dwarfism (OMIA 001985-9913) [36]. Furthermore, a nonsense variant in *PLD4* causing a recessive genodermatosis observed in German Fleckvieh (OMIA 001935-

9913) and a missense variant in *OPA3* associated with a form of dilated cardiomyopathy predominantly affecting Swiss Fleckvieh (OMIA 000162-9913) were described [37,38]. In German Fleckvieh, a *MOCS1*-related form of arachnomelia was identified to be due to a frameshift variant (OMIA 001541-9913) [39]. Furthermore, the BH2-related *TUBD1* missense variant, known to cause juvenile mortality in Braunvieh cattle (OMIA 001939-9913), was also observed in German and Austrian Fleckvieh [23]. To the best of our knowledge, so far, no carriers for any of these eight deleterious alleles have been found in the purebred Swiss Simmental population.

In purebred Swiss Simmental cattle, no genomic analysis has yet been carried out to systematically identify recessively inherited harmful variants that affect female reproduction or calf rearing. Based on a reverse genetic screen, the present study aimed at a comprehensive analysis of available SNP and WGS data to identify genomic regions containing deficient homozygous haplotypes as well as linked candidate causal variants for hidden phenotypes.

2. Materials and Methods

The national breeding association provided us with the available SNP data, including all genotyped purebred Swiss Simmental cattle born after 2009 and their ancestors. The genomic positions of the markers relate to the latest cattle reference sequence ARS-UCD1.2 [40,41]. Due to the application of several routinely available SNP arrays ranging from 9000 to 150,000 SNPs, the data had to be imputed. Therefore, the software Fimpute v2.2 [42] was used with default parameters to increase the number of markers and correct for wrongly called markers. In order to assure the quality, SNPs with a minor allele frequency <0.01 were excluded from the dataset. Furthermore, SNPs that could not pass the quality measures for the call rate per SNP >0.99 were excluded and animals with call rates <0.8 were excluded too. Quality control was applied before and after imputation. This resulted in a final SNP data set of 114,890 markers for 9965 animals.

As a first step, haplotypes showing a deviation from the HWE, indicated by depletion of homozygosity, were identified. The first subset of data analyzed included only fully genotyped trios where the complete trio (sire, dam, and offspring) were genotyped ($n = 2626$), further called “trio” approach. The second dataset analyzed included genotyped trios where an offspring and two paternal animals (sire and maternal grandfather) were genotyped ($n = 3969$), subsequently called parent–grandparent “pgp” approach. Both data sets, trio and pgp, were used in screening for window size of 50 markers within the software snp1101 [43]. The analyzed haplotypes overlap, as the windows were continuously moved marker by marker. The snp1101 software used the Fisher exact test of HWE [44] to analyze the resulting haplotypes. Furthermore, the p-values were corrected for a false discovery rate with the Benjamini–Yekutieli method and a significance level of 5% was applied [45]. For each significant haplotype region, we selected the most significant haplotype (lowest p-value). Regarding the identification of candidate variants, haplotype regions were defined by the selected haplotypes and increased on both sides by 2 Mb in order to make sure that the regions of overlapping significant haplotypes were included.

Within the 11 regions detected, between 14 and 262 genes were annotated, giving a total of 669 genes. To these genes, information regarding associated phenotypes were collected from the various online databases: HUGO Gene Nomenclature Committee (<https://www.genenames.org>; accessed on 7 October 2021), Mouse Genome Informatics (<http://www.informatics.jax.org>; accessed on 7 October 2021), International Mouse Phenotyping Consortium (<https://www.mousephenotype.org>; accessed on 7 October 2021), Online Mendelian Inheritance in Man (<https://www.omim.org>; accessed on 7 October 2021), Online Mendelian Inheritance in Animals (<https://omia.org/home/>; accessed on 7 October 2021), and Decipher (<https://www.deciphergenomics.org>; accessed on 7 October 2021).

With the selected haplotypes, we predicted individual diplotypes that represent if an animal carries one, two, or no copies of the haplotype. Based on these diplotypes, we selected three carrier animals for whole-genome sequencing (WGS) for each haplotype

region. Therefore, 23 Simmental cattle (1 female and 22 male) carrying one or more significant haplotypes were selected for this study. WGS data were prepared as previously described [46]. However, recalibration was performed with the variant catalogue of the 1000 Bull Genomes Project run 7 (BQSR file version 2) [47,48]. The 23 genomes sequenced in this study were submitted to the 1000 Bull Genomes project [47,48] and are therefore part of the 5116 animals in the recent variant catalogue (run9). This international dataset was used to evaluate candidate causative variants in a larger cohort and across breeds, to evaluate breed specificity as well as reduced homozygosity. Furthermore, we had access to an additional 115 publicly available genomes from the Swiss Comparative Bovine Resequencing project, deposited in the European Nucleotide Archive under project accession PRJEB18113, that were yet not added to the 1000 Bull Genomes project. The combined WGS dataset of 5231 genomes includes 62 purebred Swiss Simmental cattle.

In order to identify haplotype-associated candidate causal variants, linkage disequilibrium (LD) analysis was performed using plink v1.9 [49]. LD (r^2) was calculated between the diplotype state and the WGS-derived genotypes for the 23 animals that were selected for WGS. Analysis was restricted to protein-changing variants, including variants annotated by SnpEff (version 4.3; [50]) with the following sequence ontology terms: missense_variant, start_lost, stop_gained (nonsense), stop_lost, stop_retained_variant, splice_acceptor_variant, splice_donor_variant, conservative_inframe_deletion, conservative_inframe_insertion, disruptive_inframe_deletion, disruptive_inframe_insertion, exon_loss_variant, frameshift_variant, and gene_fusion.

Furthermore, to improve the understanding of candidate causal variants, their base conservation scores from UCSC database called PhyloP and PhastCons were applied [51,52]. To apply these values, firstly, the variants of the bovine genome needed to be mapped to the human genome 38 [53] with the tool LiftOver of the UCSC tools. Secondly, the conservation scores of 99 vertebrates for these human positions were obtained. Additionally, the effects of protein-changing variant were estimated using PROVEAN [54] and PredictSNP [55].

3. Results

In this reverse genetic study, we used SNP data to identify haplotypes showing a depletion in homozygosity and applied WGS data to pinpoint candidate causal variants in purebred Swiss Simmental animals. Reduced homozygosity due to hidden recessive variants in cattle could, so far, only be causally explained by coding variants. Therefore, we focused on variants having a moderate impact, such as missense variants, conservative in-frame insertions and deletions up to a size of 50 bp, as well as on all other protein-changing with high impact including loss-of-function variants, such as stop-gains (nonsense), splice site-disrupting SNVs, frameshift indels in a coding sequence, or deletions that remove coding exons.

3.1. Identification of Deficient Homozygous Haplotypes in Swiss Simmental Cattle

We detected seven haplotype regions with the trio approach and nine haplotype regions applying the pgg approach (Table 1, Figure 1, Table S1). Five of these haplotype regions appear in both analyses. We named the haplotype regions, in accordance to previous studies performed in other cattle breeds, as Simmental Haplotypes (SH) 1 to 11 [22] (Table 1, Table S1). All 11 identified haplotypes show a deficiency of at least 85 percent of the expected homozygous animals within the studied Swiss Simmental population (Table 1, Table S1). Four selected haplotypes (SH5, SH7, SH8, and SH10) presented a complete deficit of observed homozygous animals, whereas the others showed a partial deficiency ranging from 85 to 96% of the expected homozygotes (Table 1; Table S1). The average length of the eleven haplotypes is 1.17 Mb and ranges from 0.73 to 1.94 Mb (Table 1).

Table 1. List of deficient homozygous haplotypes in Swiss Simmental cattle.

Haplotype ^a	Analysis	Chr	Position (Mb) ^b	Length (Mb)	Number of Homozygotes			Allele Frequency
					Observed	Expected	Deficiency (%)	
SH1	trio and pgp	1	65.793–66.891	1.10	3	17	85	0.042
SH2	trio	2	1.191–1.920	0.73	1	16	94	0.040
SH3	pgp	7	104.784–105.640	0.86	1	17	94	0.041
SH4	pgp	11	29.389–30.353	0.96	4	22	85	0.047
SH5	trio	12	46.828–47.837	1.01	0	38	100	0.062
	pgp		46.831–47.843	1.01				
SH6	pgp	14	11.304–12.373	1.07	2	15	88	0.039
SH7	trio	16	6.120–8.011	1.89	0	30	100	0.055
	pgp		6.122–8.060	1.94				
SH8	pgp	18	48.763–50.005	1.24	0	11	100	0.0337
	trio		48.806–50.017	1.21				
SH9	trio and pgp	21	41.855–42.851	1.00	2	19	96	0.043
SH10	pgp	24	41.945–43.140	1.20	0	10	100	0.032
SH11	trio	28	38.937–40.069	1.13	1	12	92	0.035

^a SH meaning Simmental Haplotype. ^b in accordance with the reference sequence ARS-UCD1.2.

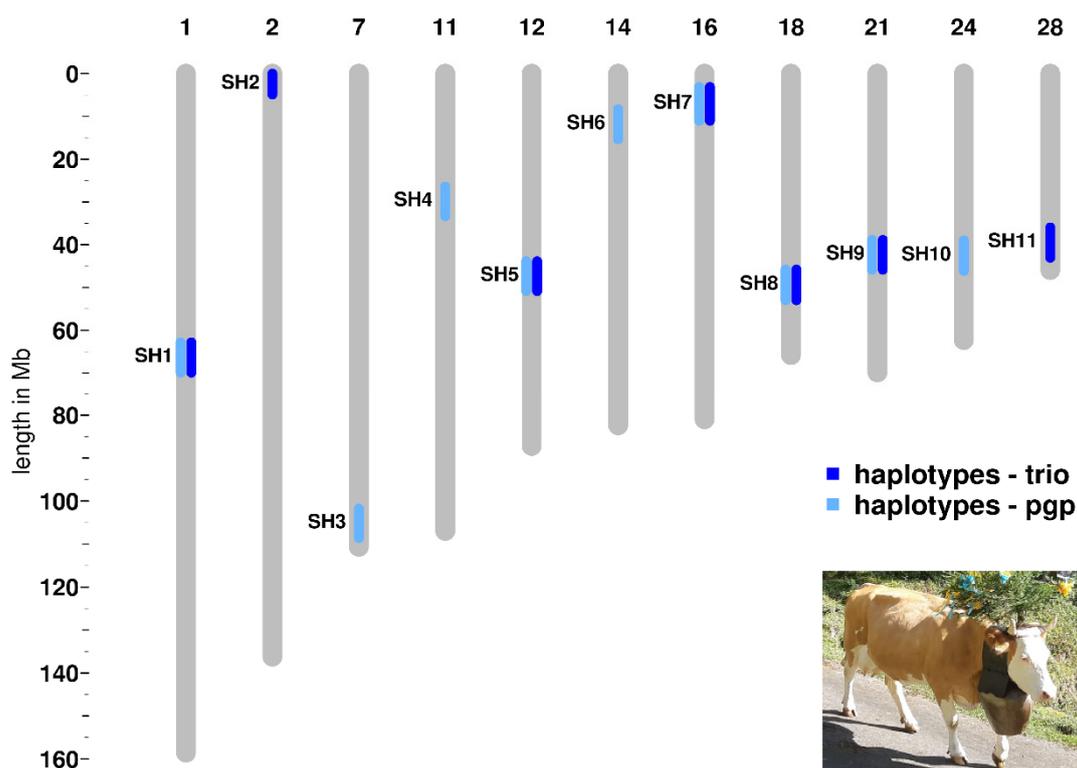


Figure 1. Genomic co-localization of the deficient homozygous haplotype regions detected in Swiss Simmental cattle using two different approaches. A typical breed-specific cow is shown in the lower right corner.

3.2. Identification of Candidate Genes in Haplotype Regions

The intensive analysis of all annotated protein-coding genes in the defined haplotype regions, extended by 2 Mb on each side, led to a comprehensive list of candidate genes possibly affecting either prenatal or postnatal lethality or associated sub-lethal phenotypes. We extracted 145 positional candidate genes of special interest, as they are associated with mammalian autosomal recessive disorders in human, mice, or other animals and listed in the consulted databases (Table S2). Loss-of-function mouse models of many of these

genes have revealed defects that affect embryonic or perinatal to pre-weaning survival and therefore represent suitable functional candidates for this study. The presented short list of the 43 most probable candidate genes includes all genes that are associated with sub-lethal or lethal phenotypes (Table 2).

Table 2. List of candidate genes known to cause recessive disorders with lethal and sub-lethal phenotypes in human.

Haplotype	Gene	Gene Description	Recessive Disorder
SH1	<i>IQCB1</i> §	IQ motif containing B1	Senior-Loken syndrome 5
	<i>MYLK</i> §	myosin light chain kinase	Megacystis-microcolon-intestinal hypoperistalsis syndrome 1
	<i>CASR</i> *	calcium sensing receptor	Hyperparathyroidism, neonatal
	<i>POGLUT1</i> *	protein O-glucosyltransferase 1	Limb-girdle muscular dystrophy-dystroglycanopathy
	<i>TIMMDC1</i> *	translocase of inner mitochondrial membrane domain containing 1	Mitochondrial complex I deficiency, nuclear type 31
SH2	<i>HERC2</i> *§	HECT and RLD domain containing E3 ubiquitin protein ligase 2	Mental retardation, MRT 38
	<i>OCA2</i> *§	OCA2 melanosomal transmembrane protein	Albinism, brown oculocutaneous
SH4	<i>CRIP1</i> *	CXXC repeat containing interactor of PDZ3 domain	Short stature with microcephaly and distinctive facies
	<i>EPCAM</i> *	epithelial cell adhesion molecule	Congenital diarrhea with tufting enteropathy
	<i>FSHR</i> *§	follicle stimulating hormone receptor	Ovarian dysgenesis 1
	<i>MSH2</i> §	mutS homolog 2	Mismatch repair cancer syndrome 2
	<i>MSH6</i>	mutS homolog 6	Mismatch repair cancer syndrome 3
	<i>PIGF</i> *	phosphatidylinositol glycan anchor biosynthesis class F	Onychodystrophy, osteodystrophy, impaired intellectual development, and seizures syndrome
	<i>PPP1R21</i> *	protein phosphatase 1 regulatory subunit 21	Neurodevelopmental disorder with hypotonia, facial dysmorphism, and brain abnormalities
	<i>TTC7A</i> *	tetratricopeptide repeat domain 7A	Gastrointestinal defects and immunodeficiency syndrome
SH5	<i>PIBF1</i> *	progesterone immunomodulatory binding factor 1	Joubert syndrome 33
SH6	<i>MYC</i> *	MYC proto-oncogene, bHLH transcription factor	Burkitt lymphoma, somatic
SH7	<i>CFH</i>	complement factor H	Basal laminar drusen, complement factor H deficiency
	<i>CR2</i> §	complement C3d receptor 2	Immunodeficiency, CVID7
	<i>IL10</i> §	interleukin 10	Critical role in the control of immune responses
SH8	<i>ARHGEF1</i> §	Rho guanine nucleotide exchange factor 1	Immunodeficiency 62
	<i>BCKDHA</i> *	branched chain keto acid dehydrogenase E1 subunit alpha	Maple syrup urine disease, type Ia @
	<i>B9D2</i> *	B9 domain containing 2	Meckel syndrome 10, Joubert syndrome 34
	<i>COQ8B</i> *	coenzyme Q8B	Nephrotic syndrome, type 9
	<i>DLL3</i> §	delta-like canonical Notch ligand 3	Spondylocostal dysostosis 1
	<i>ERF</i> *	ETS2 repressor factor	Spondylocostal dysostosis 1
	<i>ETHE1</i> §	ETHE1 persulfide dioxygenase	Ethylmalonic encephalopathy

Table 2. Cont.

Haplotype	Gene	Gene Description	Recessive Disorder
	<i>LTBP4</i> §	latent transforming growth factor beta binding protein 4	Cutis laxa, autosomal recessive, type IC
	<i>MEGF8</i> *	multiple EGF-like domains 8	Carpenter syndrome 2
	<i>PLEKHG2</i>	pleckstrin homology and RhoGEF domain containing G2	Leukodystrophy and acquired microcephaly
	<i>RYR1</i> *	ryanodine receptor 1	Neuromuscular disease
	<i>SMG9</i> *	SMG9 nonsense mediated mRNA decay factor	Heart and brain malformation syndrome
	<i>SPINT2</i> *	serine peptidase inhibitor, Kunitz type 2	Diarrhea 3, secretory sodium, congenital, syndromic
	<i>SPTBN4</i>	spectrin beta, non-erythrocytic 4	Neurodevelopmental disorder with hypotonia, neuropathy, and deafness
	<i>TGFB1</i> *	transforming growth factor beta 1	Inflammatory bowel disease, immunodeficiency, and encephalopathy
	<i>TIMM50</i> *	translocase of inner mitochondrial membrane 50	3-methylglutaconic aciduria, type IX
	<i>XRCC1</i> *	X-ray repair cross complementing 1	Spinocerebellar ataxia, SCAR26
SH9	<i>NUBPL</i> *	nucleotide binding protein-like	Mitochondrial complex I deficiency, nuclear type 21
	<i>AFG3L2</i> *	AFG3-like matrix AAA peptidase subunit 2	Spastic ataxia 5
	<i>LAMA1</i> *	laminin subunit alpha 1	Poretti–Boltshauser syndrome
SH10	<i>MC2R</i> §	melanocortin 2 receptor	Glucocorticoid deficiency due to ACTH unresponsiveness
	<i>NDUFV2</i> *	NADH:ubiquinone oxidoreductase core subunit V2	Mitochondrial complex I deficiency, nuclear type 7
	<i>PIEZO2</i> *	piezo type mechanosensitive ion channel component 2	Arthrogryposis, distal, with impaired proprioception and touch

* Homozygous lethal with complete penetrance, § homozygous lethal with incomplete penetrance, ® disorder described in cattle (OMIA-ID: 000627-9913).

3.3. Identification of Candidate Causal Variants

For three deficient homozygous haplotypes (SH5, SH8, and SH9), by linkage disequilibrium analysis, we found perfectly linked ($r^2 = 1$) candidate causal variants. These three haplotypes were detected with both the pgp and the trio approach. For each of these haplotypes, we propose a protein changing SNV (Table 3; Table S3). These three variants never occur in homozygous state in the analyzed 5231 bovine genomes of various cattle breeds (Table S3). Interestingly, the SH8-associated variant is apparently specific for Swiss Simmental; however, the variants associated with SH5 and SH9 occur sporadically in single animals of some other breeds (Table S3).

Table 3. Candidate causal variants for deficiency of homozygotes in Swiss Simmental cattle.

Haplotype	Gene	OMIM	Variant	Transcript ^a	Coding DNA Change	Predicted Protein Change
SH5	<i>DIS3</i>	607533	1 bp insertion (frameshift)	XM_025000110.1	c.2032dupA	p.Ile678AsnTer2
SH8	<i>CYP2B6</i>	123930	SNV (missense)	NM_001075173.1	c.938T > A	p.Ile313Asn
SH9	<i>NUBPL</i>	613621	SNV (missense)	NM_001193042.1	c.428C > A	p.Ser143Tyr

^a According to the NCBI Annotation Release 106 (National Center for Biotechnology Information, 2018b).

Among the three proposed non-synonymous variants are two missense variants altering evolutionary conserved residues and a frameshift variant that significantly truncates the encoded protein. The SH8-related SNV in exon 4 of the bovine *cytochrome P450 family 2 subfamily B member 6 (CYP2B6)* gene on chromosome 18 at position 50296371 is a missense variant (NM_001075173.1: p.Ile313Asn) that was predicted by PROVEAN to have a deleterious effect (Table 3, Table S3). The SH9-related SNV located in exon 6 of the bovine *nucleotide binding protein-like (NUBPL)* gene on chromosome 21 at position 42154344 represents a missense variant (NM_001193042.1: p.Ser143Tyr) predicted to be deleterious by PROVEAN and PredictSNP (Figure 2, Table 3, Table S3). Both presented missense mutations altering evolutionary conserved amino acids (Figure 2). The SH5-associated 1 bp insertion located in exon 16 of the bovine *DIS3 homolog, exosome endoribonuclease and 3'-5' exoribonuclease (DIS3)* gene on chromosome 12 at position 47511687 represents a loss-of-function variant. It was predicted to result in a frameshift after isoleucine 678 with a premature stop codon (NP_025000110.1: p.Ile678AsnTer2), resulting in a significantly shortened amino acid sequence, if expressed, when compared with the wild-type protein (Table 3; Table S3). In addition, the comparative DNA sequence approach (PhyloP and phastCons) showed a high conservation across species for all three variant positions (Table S3).

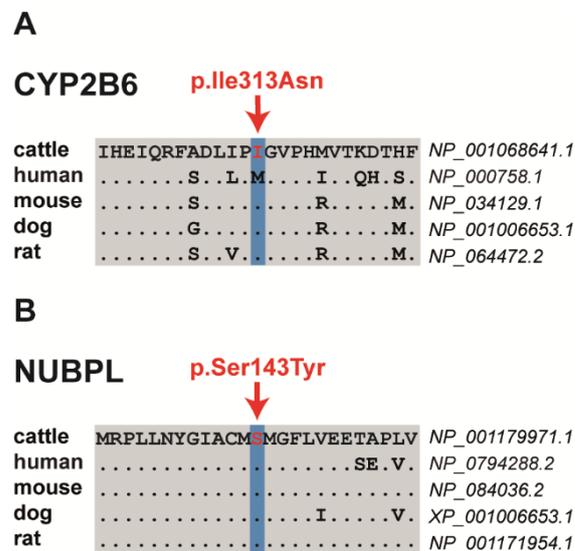


Figure 2. Multiple sequence alignment of the amino acid sequences of CYP2B6 (A) and NUBPL (B) around the position of the missense variants (shown in red), indicating evolutionary conservation across all species.

4. Discussion

For the first time, the genomic data of the current Swiss Simmental dual-purpose cattle population were analyzed for reduced homozygosity due to hidden recessive monogenic variants and validated using the international variant catalogue of the 1000 Bull Genomes Project. In addition to environmental factors, inherited deleterious variants lead to natural or artificial selection against homozygous individuals, which also explains embryonic lethality, reduce rearing success, or the exclusion from the breeding population due to poor development. Unfortunately, these phenomena are not systematically monitored and are therefore difficult phenotypes to study. Although reduced reproductive success can theoretically be detected in the sires estimated breeding values for certain fertility traits, these effects are only noticeable for deleterious alleles that have reached a high frequency in the population [56]. To overcome this issue, we performed a genome-wide missing homozygosity scan, revealing eleven haplotype regions with considerable homozygous depletion. After subsequent mining of genome sequence data for candidate causal variants, we propose three non-synonymous variants that probably cause the obvious deficiency of homozygous animals.

We applied two different approaches, a trio and pgp approach, that include genotyped trios and applies a Fisher exact test. Therefore, we were able to detect haplotypes that segregate at a lower allele frequency. As expected, we found more haplotype regions with the pgp approach in comparison to the trio approach, most likely because more genotyped groups were available. Nevertheless, the trio approach, detecting two regions not found with pgp, appears very powerful, probably because it directly traces the inheritance of the haplotypes. Previous approaches used the assumption of random mating and the deviation from HWE based on allele frequencies or used the deviation from expected number of homozygous offspring based on the haplotype state of the sire and maternal grandsire [14,20]. In contrast, the herein applied trio approach allows performing such an analysis with a small population such as Simmental. The five haplotype regions that were found in both analyses are the most probable genome regions harboring hidden harmful variants. Especially the haplotypes that never occur in homozygous state led to the suspicion of embryonic lethal variants segregating in the population. It is suspected that some haplotypes arose due to imputation errors introduced due to genotyping bias, SNP density/panel, sample size, and a bias introduced by the chosen software [57]. This would explain the haplotype regions SH3 and SH11 that arose in regions for which we could not identify any plausible functional candidate genes. Otherwise, we were able to identify candidate genes within all missing homozygous regions. Unfortunately, effect estimations of the haplotypes towards traits of female fertility and rearing success were not reasonable, as the available phenotypic data is currently limited. To avoid the detection of mostly sporadic associations rather than actual effects, it is planned to conduct such haplotype association studies in the future.

Candidate causal variants are proposed for the haplotypes SH5, SH8, and SH9 in the genes *DIS3*, *CYP2B6*, and *NUBPL*, respectively. These variants all show complete depletion in homozygosity, perfect LD to the associated haplotype, and high conservation scores when compared across 99 genomes, indicating their importance in basic biological functions. As in other mammals, it is expected that ~100 harmful recessive variants will be found per individual in cattle, of which up to five of these impact essential genes and cause embryonic lethality or severe disease when homozygous [10]. Nevertheless, it is recognized that it is very difficult to clarify the actual deleterious functions of these variants, although given the genes involved, it is assumed that these variants influence fitness. Reverse genetic screens to identify genes with major effects, as used in the current study, are therefore helpful to assign function to variants in candidate genes and/or so far less characterized genes such as *DIS3* and *CYP2B6*.

The SH5-related loss-of-function variant found in bovine *DIS3*, most likely leading to missing homozygosity, represents the first time a pathogenic variant that most likely causes embryonic lethality has been identified. If the mutant mRNA transcript were to escape nonsense-mediated decay, even if this truncated protein was expressed, it would lack roughly 30 percent of the C-terminal part, and therefore it is not expected to contribute any function. *DIS3*, also known as ribosomal RNA-processing protein 44 (RRP44), is a RNase II/R-like enzyme located primarily in the nucleus (Table S3) [58]. The protein has catalytic function in the RNA exosome complex, which is responsible for 3'-end processing and RNA degradation of a broad variety of RNAs [58,59]. Biological functions are associated with RNA metabolism, mitotic control, spindle-fiber formation, antibody diversification, microtubule production, and growth and development [60]. Recently, pathogenic variants in genes encoding both structural and catalytic subunits of the RNA exosome have been linked to human disease, such as *EXOSC3* and *EXOSC8* related forms of pontocerebellar hypoplasia, representing recessive neurodegenerative diseases [61]. To our knowledge, the *DIS3* gene has not yet been associated with Mendelian diseases, but variants are reported to be associated with various types of cancers and multiple myeloma [60–62]. Interestingly, the book of Fasken et al. (2020) provides a summary of the most common variants in *DIS3* which all occur in heterozygous state, are associated with multiple myeloma, and seem to have mild effects only, while Tomecki et al. (2014) suggests the potential lethality of

mutations in the PIN domain of *DIS3*. In *Drosophila*, a knock-down model led to wingless animals implicating an important role in development [63]. Nevertheless, the inactivation of *DIS3* in B cells was shown to lead to an increase in unbalanced DNA translocations [64]. Lastly, public databases for mouse phenotypes indicate a complete pre-weaning lethality of *DIS3* knock-out mice (<https://www.mousephenotype.org/data/genes/MGI:1919912>; <http://www.informatics.jax.org/marker/MGI:1919912>; accessed on 7 October 2021).

Despite a long list of candidate genes located within the SH8-region, we propose a missense variant in the *CYP2B6* gene as a candidate causal variant. The main reason for this is the perfect LD to the haplotype, the complete absence of homozygous animals, and the prediction of the DNA position to be highly conserved and the amino acid exchange to be deleterious. Regarding the gene function, CYP2B6 is a protein of the cytochromes P450 subfamily 2B (HGNC: 20604). This enzyme is known to be of importance for drug metabolism, as well as endogenous compounds, environmental toxins, and other substances [65,66]. For example, the susceptibility to Efavirenz depends on the individual *CYP2B6* genotype (OMIM: 123930). Several SNV were detected in human that are associated with the expression level and activity (increased and decreased) of *CYP2B6* with a population-wide importance [65] (<https://www.pharmvar.org/gene/CYP2B6>; accessed on 26 October 2021). In monkeys and human, it was shown that *CYP2B6* is expressed in the brain and affected by nicotine and alcohol consumption; however, its neurological function remains unclear [65–67]. The protein is also expressed in the placenta [68] and it was shown that the pregnancy hormone estradiol induces the expression of *CYP2B6* [69]. Nevertheless, what the function and importance of *CYP2B6* is in maintaining pregnancy is unclear. As the herein identified bovine variant is predicted to be deleterious, we speculate that function of *CYP2B6* might be impaired during development.

Lastly, we propose the missense variant in the bovine *NUBPL* gene, exchanging a strongly conserved residue predicted to be deleterious and affecting a highly conserved nucleotide, to be causal for the deficit of homozygosity of SH9. As we observe few haplotype carriers in our data, we speculate that the effect of the variant is not fully penetrant or that signs of poor development appear later in life, after initial genotyping of young animals. *NUBPL*, also known as *iron-sulfur protein required for NADH dehydrogenase (IND1)*, is a protein that is vital for the assembly of the respiratory complex I [70]. More precisely, the *NUBPL* supplies Fe/S clusters to the respiratory complex I and thereby ensures that important subunits are delivered to build the whole complex [70]. In human, pathogenic variants affecting *NUBPL* are associated with the autosomal recessive mitochondrial complex I deficiency disorder (OMIM: 613621, 618242) [71–73]. These variants include missense, frameshift and splice site variants, as well as small and large insertions and deletions. Clinical symptoms of mitochondrial complex I deficiency include, among others, ataxia, dysarthria, hypotonia, nystagmus, spasticity, and tremor [73]. Furthermore, variants in *NUBPL* were hypothesized as risk factors for Parkinson’s disease [74]. A mouse model identified the necessity of the protein as knock-out alleles led to homozygous lethality (MGI: 1924076) [75].

However, if the proposed variants are indeed depleted in a number of homozygous animals, it needs to be evaluated by further genotyping of larger cohorts before implementation into selection schemes. This is planned by adding the variants to a custom array for genotyping larger cohorts of further animals, both Swiss Simmental as well as other local cattle populations. An alternative approach to confirm the absence of homozygous animals, particularly in the offspring of carrier-by-carrier mating’s, would strongly support the deleterious nature of the variants as shown before [10]. In particular, for the *NUBPL*-associated variant, the observed homozygous animals should be examined in detail.

The chosen approach requires that the haplotype and the causative variant are in near perfect linkage disequilibrium and, obviously, this is not always the case. This could be an explanation for the fact that we could not identify any potentially causal variants in the other haplotype regions. An alternative approach, therefore, is to mine the genome sequence data for candidate variants, e.g., loss-of-function in the listed essential candidate

genes, and to genotype these directly in large cohorts of Swiss Simmental cattle. Using this approach, nine causal variants were uncovered in cattle that would not have been detected using SNP-based haplotype approaches [10].

Finally, another drawback of our study was the restriction to protein-changing variants. Moreover, we had no evidence for perfectly linked non-coding regulatory variants, as well as the limitation to consider only SNV and short indels, overlooking possible larger structural variants.

5. Conclusions

In the presented project, we mined SNP and WGS data by applying a reverse genetic approach. Without any phenotypic evidence but mining the data of almost 10 thousand SNP genotyped Swiss Simmental cattle and more than 5200 WGS animals from a variety of breeds, we propose three candidate variants in the genes *DIS3*, *CYP2B6*, and *NUBPL* causing embryonic lethality and/or yet unknown recessive developmental disorders. After phenotypic validation of these variants, selection against these variants is recommended.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ani11123535/s1>, Table S1: Haplotype information of SH1 to SH11, Table S2: Comprehensive list of candidate genes located in the haplotype regions, including gene information from MGI, IMPC, OMIM, and OMIA, Table S3: Comprehensive list of potential candidate causal variants.

Author Contributions: Conceptualization, C.D. and F.R.S.; formal analysis, I.M.H. and F.R.S.; investigation, I.M.H. and F.R.S.; data curation, I.M.H.; writing—original draft preparation, I.M.H.; writing—review and editing, C.D., F.R.S.; visualization, I.M.H.; supervision, C.D. and F.R.S.; project administration and funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Swiss National Science Foundation, grant number 172911.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not an experimental study, but an analysis of existing genotype data.

Data Availability Statement: The SNP data of Swiss Simmental cattle are owned by the breeding association swissherdbook. Therefore, we ask interested people to contact the authors or the breeding association directly in order to gain access to the SNP data. The WGS data are publicly available at the European Nucleotide Archive under project accession PRJEB18113.

Acknowledgments: We are very thankful for the genomic data provided by the Swiss cattle breeding association swissherdbook. The authors also wish to thank Swissgenetics and Selectstar for providing semen. The authors are thankful to the University of Bern institutions, namely the Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit, for performing the WGS experiments and providing high performance computing infrastructure, respectively. Finally, we acknowledge the 1000 Bull Genomes Project consortium, led by Christy Vander Jagt, Amanda Chamberlain, and Hans Daetwyler, for providing the variant catalogue.

Conflicts of Interest: The authors declare no conflict of interest.

Ethics Approval and Consent to Participate: According to the Swiss laws, no ethics approval was needed. In this study, available data were reused and mined using novel approaches.

References

1. Signer-Hasler, H.; Burren, A.; Neuditschko, M.; Frischknecht, M.; Garrick, D.; Stricker, C.; Gredler, B.; Bapst, B.; Flury, C. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genet. Sel. Evol.* **2017**, *49*, 83. [CrossRef]
2. Stranzinger, G.F.; Steiger, D.; Kneubühler, J.; Hagger, C. Y chromosome polymorphism in various breeds of cattle (*Bos taurus*) in Switzerland. *J. Appl. Genet.* **2007**, *48*, 241–245. [CrossRef]
3. Cesarani, A.; Gaspa, G.; Pauciullo, A.; Degano, L.; Vicario, D.; Macciotta, N.P.P. Genome-wide analysis of homozygosity regions in European Simmental bulls. *J. Anim. Breed. Genet.* **2021**, *138*, 69–79. [CrossRef]
4. Pryce, J.E.; Woolaston, R.; Berry, D.P.; Wall, E.; Winters, M.; Butler, R.; Shaffer, M. World Trends in Dairy Cow Fertility. In Proceedings of the 10th World Congress on Genetics Applied to Livestock Production, Vancouver, BC, Canada, 17–22 August 2014; Volume 10, p. 6.

5. Walsh, S.W.; Williams, E.J.; Evans, A.C.O. A review of the causes of poor fertility in high milk producing dairy cows. *Anim. Reprod. Sci.* **2011**, *123*, 127–138. [[CrossRef](#)] [[PubMed](#)]
6. Ma, L.; Sonstegard, T.S.; Cole, J.B.; Vantassell, C.P.; Wiggans, G.R.; Crooker, B.A.; Tan, C.; Prakapenka, D.; Liu, G.E.; Da, Y. Genome changes due to artificial selection in U.S. Holstein cattle. *BMC Genom.* **2019**, *20*, 128. [[CrossRef](#)] [[PubMed](#)]
7. Zobel, R.; Tkalčić, S.; Pipal, I.; Buić, V. Incidence and factors associated with early pregnancy losses in Simmental dairy cows. *Anim. Reprod. Sci.* **2011**, *127*, 121–125. [[CrossRef](#)]
8. Seegers, H.; Beaudeau, F.; Fourichon, C.; Bareille, N. Reasons for culling in French Holstein cows. *Prev. Vet. Med.* **1998**, *36*, 257–271. [[CrossRef](#)]
9. Rilanto, T.; Reimus, K.; Orro, T.; Emanuelson, U.; Viltrop, A.; Mõtus, K. Culling reasons and risk factors in Estonian dairy cows. *BMC Vet. Res.* **2020**, *16*, 173. [[CrossRef](#)]
10. Charlier, C.; Li, W.; Harland, C.; Littlejohn, M.; Coppieters, W.; Creagh, F.; Davis, S.; Druet, T.; Faux, P.; Guillaume, F.; et al. NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. *Genome Res.* **2016**, *26*, 1333–1341. [[CrossRef](#)] [[PubMed](#)]
11. VanRaden, P.M.; Olson, K.M.; Null, D.J.; Hutchison, J.L. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J. Dairy Sci.* **2011**, *94*, 6153–6161. [[CrossRef](#)]
12. Casellas, J.; Canas-Alvarez, J.J.; Gonzalez-Rodriguez, A.; Puig-Oliveras, A.; Fina, M.; Piedrafita, J.; Molina, A.; Diaz, C.; Baro, J.A.; Varona, L. Bayesian analysis of parent-specific transmission ratio distortion in seven Spanish beef cattle breeds. *Anim. Genet.* **2017**, *48*, 93–96. [[CrossRef](#)] [[PubMed](#)]
13. Casellas, J.; Gualarte, R.J.; Farber, C.R.; Varona, L.; Mehrabian, M.; Schadt, E.E.; Lusia, A.J.; Attie, A.D.; Yandell, B.S.; Medrano, J.F. Genome Scans for Transmission Ratio Distortion Regions in Mice. *Genetics* **2012**, *191*, 247–259. [[CrossRef](#)] [[PubMed](#)]
14. Derks, M.F.L.; Megens, H.J.; Bosse, M.; Lopes, M.S.; Harlizius, B.; Groenen, M.A.M. A systematic survey to identify lethal recessive variation in highly managed pig populations. *BMC Genom.* **2017**, *18*, 858. [[CrossRef](#)]
15. Derks, M.F.L.; Megens, H.J.; Bosse, M.; Visscher, J.; Peeters, K.; Bink, M.; Vereijken, A.; Gross, C.; de Ridder, D.; Reinders, M.J.T.; et al. A survey of functional genomic variation in domesticated chickens. *Genet. Sel. Evol.* **2018**, *50*, 17. [[CrossRef](#)]
16. Georges, M.; Charlier, C.; Hayes, B. Harnessing genomic information for livestock improvement. *Nat. Rev. Genet.* **2019**, *20*, 135–156. [[CrossRef](#)]
17. Abdalla, E.A.; Id-Lahoucine, S.; Cánovas, A.; Casellas, J.; Schenkel, F.S.; Wood, B.J.; Baes, C.F. Discovering lethal alleles across the turkey genome using a transmission ratio distortion approach. *Anim. Genet.* **2020**, *51*, 876–889. [[CrossRef](#)] [[PubMed](#)]
18. Hoff, J.L.; Decker, J.E.; Schnabel, R.D.; Taylor, J.F. Candidate lethal haplotypes and causal mutations in Angus cattle. *BMC Genom.* **2017**, *18*, 799. [[CrossRef](#)]
19. Mesbah-Uddin, M.; Hoze, C.; Michot, P.; Barbat, A.; Lefebvre, R.; Boussaha, M.; Sahana, G.; Fritz, S.; Boichard, D.; Capitan, A. A missense mutation (p.Tyr452Cys) in the CAD gene compromises reproductive success in French Normande cattle. *J. Dairy Sci.* **2019**, *102*, 6340–6356. [[CrossRef](#)] [[PubMed](#)]
20. Fritz, S.; Capitan, A.; Djari, A.; Rodriguez, S.C.; Barbat, A.; Baur, A.; Grohs, C.; Weiss, B.; Boussaha, M.; Esquerre, D.; et al. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. *PLoS ONE* **2013**, *8*, e65550. [[CrossRef](#)]
21. Pausch, H.; Schwarzenbacher, H.; Burgstaller, J.; Flisikowski, K.; Wurmser, C.; Jansen, S.; Jung, S.; Schnieke, A.; Wittek, T.; Fries, R. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. *BMC Genom.* **2015**, *16*, 312. [[CrossRef](#)]
22. Cole, J.B.; VanRaden, P.M.; Null, D.J.; Hutchison, J.L.; Cooper, T.A.; Hubbard, S.M. Haplotype Tests for Recessive Disorders that Affect Fertility and Other Traits. 2021. Available online: https://www.aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html (accessed on 15 April 2021).
23. Schwarzenbacher, H.; Burgstaller, J.; Seefried, F.R.; Wurmser, C.; Hilbe, M.; Jung, S.; Fuerst, C.; Dinhopf, N.; Weissenböck, H.; Fuerst-Waltl, B.; et al. A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. *BMC Genom.* **2016**, *17*, 400. [[CrossRef](#)] [[PubMed](#)]
24. Sonstegard, T.S.; Cole, J.B.; VanRaden, P.M.; Van Tassell, C.P.; Null, D.J.; Schroeder, S.G.; Bickhart, D.; McClure, M.C. Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. *PLoS ONE* **2013**, *8*, e54872. [[CrossRef](#)] [[PubMed](#)]
25. Charlier, C.; Agerholm, J.S.; Coppieters, W.; Karlskov-Mortensen, P.; Li, W.; de Jong, G.; Fasquelle, C.; Karim, L.; Cirera, S.; Cambisano, N.; et al. A Deletion in the Bovine FANCI Gene Compromises Fertility by Causing Fetal Death and Brachyspina. *PLoS ONE* **2012**, *7*, e43085. [[CrossRef](#)]
26. McClure, M.; Kim, E.; Bickhart, D.; Null, D.; Cooper, T.; Cole, J.; Wiggans, G.; Ajmone-Marsan, P.; Colli, L.; Santus, E.; et al. Fine Mapping for Weaver Syndrome in Brown Swiss Cattle and the Identification of 41 Concordant Mutations across NRCAM, PNPLA8 and CTTNBP2. *PLoS ONE* **2013**, *8*, e59251. [[CrossRef](#)]
27. Schütz, E.; Wehrhahn, C.; Wanjek, M.; Bortfeld, R.; Wemheuer, W.E.; Beck, J.; Brenig, B. The Holstein Friesian lethal haplotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. *PLoS ONE* **2016**, *11*, e0154602. [[CrossRef](#)]

28. Null, D.J.; Huthcins, J.L.; Bickhart, D.M.; VanRaden, P.M.; Cole, J.B. Discovery of a haplotype affecting fertility in Ayrshire dairy cattle and identification of a putative causal variant. In Proceedings of the American Dairy Science Association Annual Meeting, Pittsburgh, PA, USA, 25–29 June 2017; Volume 100, p. 199.
29. Guarini, A.R.; Sargolzaei, M.; Brito, L.F.; Kroezen, V.; Lourenco, D.A.L.; Baes, C.F.; Miglior, F.; Cole, J.B.; Schenkel, F.S. Estimating the effect of the deleterious recessive haplotypes AH1 and AH2 on reproduction performance of Ayrshire cattle. *J. Dairy Sci.* **2019**, *102*, 5315–5322. [[CrossRef](#)]
30. Cooper, T.A.; Wiggans, G.R.; Null, D.J.; Hutchison, J.L.; Cole, J.B. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. *J. Dairy Sci.* **2014**, *97*, 3878–3882. [[CrossRef](#)]
31. Ben Braiek, M.; Fabre, S.; Hozé, C.; Astruc, J.-M.; Moreno-Romieux, C. Identification of homozygous haplotypes carrying putative recessive lethal mutations that compromise fertility traits in French Lacaune dairy sheep. *Genet. Sel. Evol.* **2021**, *53*, 41. [[CrossRef](#)]
32. Michot, P.; Fritz, S.; Barbat, A.; Boussaha, M.; Deloche, M.C.; Grohs, C.; Hoze, C.; Le Berre, L.; Le Bourhis, D.; Desnoes, O.; et al. A missense mutation in PFAS (phosphoribosylformylglycinamide synthase) is likely causal for embryonic lethality associated with the MH1 haplotype in Montbeliarde dairy cattle. *J. Dairy Sci.* **2017**, *100*, 8176–8187. [[CrossRef](#)]
33. Fritz, S.; Hoze, C.; Rebours, E.; Barbat, A.; Bizard, M.; Chamberlain, A.; Escoufflaire, C.; Vander Jagt, C.; Boussaha, M.; Grohs, C.; et al. An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. *J. Dairy Sci.* **2018**, *101*, 6220–6231. [[CrossRef](#)]
34. Hozé, C.; Escoufflaire, C.; Mesbah-Uddin, M.; Barbat, A.; Boussaha, M.; Deloche, M.C.; Boichard, D.; Fritz, S.; Capitan, A. Short communication: A splice site mutation in CENPU is associated with recessive embryonic lethality in Holstein cattle. *J. Dairy Sci.* **2020**, *103*, 607–612. [[CrossRef](#)]
35. Kunz, E.; Rothammer, S.; Pausch, H.; Schwarzenbacher, H.; Seefried, F.R.; Matiasek, K.; Seichter, D.; Russ, I.; Fries, R.; Medugorac, I. Confirmation of a non-synonymous SNP in PNPLA8 as a candidate causal mutation for Weaver syndrome in Brown Swiss cattle. *Genet. Sel. Evol.* **2016**, *48*, 21. [[CrossRef](#)] [[PubMed](#)]
36. Schwarzenbacher, H.; Wurmser, C.; Flisikowski, K.; Misurova, L.; Jung, S.; Langenmayer, M.C.; Schnieke, A.; Knubben-Schweizer, G.; Fries, R.; Pausch, H. A frameshift mutation in GON4L is associated with proportionate dwarfism in Fleckvieh cattle. *Genet. Sel. Evol.* **2016**, *48*, 25. [[CrossRef](#)] [[PubMed](#)]
37. Jung, S.; Pausch, H.; Langenmayer, M.C.; Schwarzenbacher, H.; Majzoub-Altweck, M.; Gollnick, N.S.; Fries, R. A nonsense mutation in PLD4 is associated with a zinc deficiency-like syndrome in Fleckvieh cattle. *BMC Genom.* **2014**, *15*, 623. [[CrossRef](#)] [[PubMed](#)]
38. Owczarek-Lipska, M.; Plattet, P.; Zipperle, L.; Drögemüller, C.; Posthaus, H.; Dolf, G.; Braunschweig, M.H. A nonsense mutation in the optic atrophy 3 gene (OPA3) causes dilated cardiomyopathy in Red Holstein cattle. *Genomics* **2011**, *97*, 51–57. [[CrossRef](#)]
39. Seichter, D.; Russ, I.; Förster, M.; Medugorac, I. SNP-based association mapping of Arachnomelia in Fleckvieh cattle. *Anim. Genet.* **2011**, *42*, 544–547. [[CrossRef](#)]
40. National Center for Biotechnology Information ARS-UCD1.2 Assembly. 2018. Available online: https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1 (accessed on 31 May 2018).
41. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*, gaaa021. [[CrossRef](#)]
42. Sargolzaei, M.; Chesnais, J.P.; Schenkel, F.S. A new approach for efficient genotype imputation using information from relatives. *BMC Genom.* **2014**, *15*, 478. [[CrossRef](#)]
43. Sargolzaei, M. *SNP1101 User's Guide. Version 1.0*; HiggsGene Solut. Inc.: Guelph, ON, Canada, 2014.
44. Wigginton, J.E.; Cutler, D.J.; Abecasis, G.R. A note on exact tests of Hardy-Weinberg equilibrium. *Am. J. Hum. Genet.* **2005**, *76*, 887–893. [[CrossRef](#)]
45. Benjamini, Y.; Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **2001**, *29*, 1165–1188. [[CrossRef](#)]
46. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [[CrossRef](#)]
47. Bovine Research Community 1000 Bull Genomes Project. 2018. Available online: <http://www.1000bullgenomes.com/> (accessed on 15 June 2018).
48. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [[CrossRef](#)] [[PubMed](#)]
49. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
50. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)] [[PubMed](#)]
51. Pollard, K.S.; Hubisz, M.J.; Rosenbloom, K.R.; Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **2010**, *20*, 110–121. [[CrossRef](#)]

52. Siepel, A.; Bejerano, G.; Pedersen, J.S.; Hinrichs, A.S.; Hou, M.; Rosenbloom, K.; Clawson, H.; Spieth, J.; Hillier, L.D.W.; Richards, S.; et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **2005**, *15*, 1034–1050. [[CrossRef](#)] [[PubMed](#)]
53. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; Fitzhugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [[CrossRef](#)]
54. Choi, Y.; Sims, G.E.; Murphy, S.; Miller, J.R.; Chan, A.P. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE* **2012**, *7*, e46688. [[CrossRef](#)]
55. Bendl, J.; Stourac, J.; Salanda, O.; Pavelka, A.; Wieben, E.D.; Zendulka, J.; Brezovsky, J.; Damborsky, J. PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. *PLoS Comput. Biol.* **2014**, *10*, e1003440. [[CrossRef](#)] [[PubMed](#)]
56. Kadri, N.K.; Sahana, G.; Charlier, C.; Iso-Touru, T.; Guldbandsen, B.; Karim, L.; Nielsen, U.S.; Panitz, F.; Aamand, G.P.; Schulman, N.; et al. A 660-Kb Deletion with Antagonistic Effects on Fertility and Milk Production Segregates at High Frequency in Nordic Red Cattle: Additional Evidence for the Common Occurrence of Balancing Selection in Livestock. *PLoS Genet.* **2014**, *10*, e1004049. [[CrossRef](#)]
57. Shi, S.; Yuan, N.; Yang, M.; Du, Z.; Wang, J.; Sheng, X.; Wu, J.; Xiao, J. Comprehensive Assessment of Genotype Imputation Performance. *Hum. Hered.* **2018**, *83*, 107–116. [[CrossRef](#)]
58. Tomecki, R.; Kristiansen, M.S.; Lykke-Andersen, S.; Chlebowska, A.; Larsen, K.M.; Szczesny, R.J.; Drazkowska, K.; Pastula, A.; Andersen, J.S.; Stepien, P.P.; et al. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.* **2010**, *29*, 2342–2357. [[CrossRef](#)]
59. Schmid, M.; Jensen, T.H. *The Nuclear RNA Exosome and Its Cofactors BT—The Biology of mRNA: Structure and Function*; Oeffinger, M., Zenklusen, D., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 113–132. ISBN 978-3-030-31434-7.
60. Robinson, S.R.; Oliver, A.W.; Chevassut, T.J.; Newbury, S.F. The 3' to 5' Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease. *Biomolecules* **2015**, *5*, 1515–1539. [[CrossRef](#)]
61. Fasken, M.B.; Morton, D.J.; Kuiper, E.G.; Jones, S.K.; Leung, S.W.; Corbett, A.H. *The RNA Exosome and Human Disease BT—The Eukaryotic RNA Exosome: Methods and Protocols*; LaCava, J., Vaňáčová, Š., Eds.; Springer New York: New York, NY, USA, 2020; pp. 3–33. ISBN 978-1-4939-9822-7.
62. Tomecki, R.; Drazkowska, K.; Kucinski, I.; Stodus, K.; Szczesny, R.J.; Gruchota, J.; Owczarek, E.P.; Kalisiak, K.; Dziembowski, A. Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target. *Nucleic Acids Res.* **2014**, *42*, 1270–1290. [[CrossRef](#)]
63. Towler, B.P.; Jones, C.I.; Viegas, S.C.; Apura, P.; Waldron, J.A.; Smalley, S.K.; Arraiano, C.M.; Newbury, S.F. The 3'-5' exoribonuclease Dis3 regulates the expression of specific microRNAs in *Drosophila* wing imaginal discs. *RNA Biol.* **2015**, *12*, 728–741. [[CrossRef](#)] [[PubMed](#)]
64. Laffleur, B.; Lim, J.; Zhang, W.; Chen, Y.; Pefanis, E.; Bizarro, J.; Batista, C.R.; Wu, L.; Economides, A.N.; Wang, J.; et al. Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in B cells. *Nat. Genet.* **2021**, *53*, 230–242. [[CrossRef](#)]
65. Turpeinen, M.; Zanger, U.M. Cytochrome P450 2B6: Function, genetics, and clinical relevance. *Drug Metabol. Drug Interact.* **2012**, *27*, 185–197. [[CrossRef](#)] [[PubMed](#)]
66. Lee, A.M.; Miksys, S.; Palmour, R.; Tyndale, R.F. CYP2B6 is expressed in African Green monkey brain and is induced by chronic nicotine treatment. *Neuropharmacology* **2006**, *50*, 441–450. [[CrossRef](#)]
67. Miksys, S.; Lerman, C.; Shields, P.G.; Mash, D.C.; Tyndale, R.F. Smoking, alcoholism and genetic polymorphisms alter CYP2B6 levels in human brain. *Neuropharmacology* **2003**, *45*, 122–132. [[CrossRef](#)]
68. Li, L.; Zhang, Q.-Y.; Ding, X. A CYP2B6-humanized mouse model and its potential applications. *Drug Metab. Pharmacokinet.* **2018**, *33*, 2–8. [[CrossRef](#)]
69. Koh, K.H.; Jurkovic, S.; Yang, K.; Choi, S.-Y.; Jung, J.W.; Kim, K.P.; Zhang, W.; Jeong, H. Estradiol induces cytochrome P450 2B6 expression at high concentrations: Implication in estrogen-mediated gene regulation in pregnancy. *Biochem. Pharmacol.* **2012**, *84*, 93–103. [[CrossRef](#)] [[PubMed](#)]
70. Sheftel, A.D.; Stehling, O.; Pierik, A.J.; Netz, D.J.A.; Kerscher, S.; Elsässer, H.-P.; Wittig, I.; Balk, J.; Brandt, U.; Lill, R. Human ind1, an iron-sulfur cluster assembly factor for respiratory complex I. *Mol. Cell. Biol.* **2009**, *29*, 6059–6073. [[CrossRef](#)] [[PubMed](#)]
71. Calvo, S.E.; Tucker, E.J.; Compton, A.G.; Kirby, D.M.; Crawford, G.; Burt, N.P.; Rivas, M.; Guiducci, C.; Bruno, D.L.; Goldberger, O.A.; et al. High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nat. Genet.* **2010**, *42*, 851–858. [[CrossRef](#)]
72. Kevelam, S.H.; Rodenburg, R.J.; Wolf, N.I.; Ferreira, P.; Lunsing, R.J.; Nijtmans, L.G.; Mitchell, A.; Arroyo, H.A.; Rating, D.; Vanderver, A.; et al. NUBPL mutations in patients with complex I deficiency and a distinct MRI pattern. *Neurology* **2013**, *80*, 1577–1583. [[CrossRef](#)] [[PubMed](#)]
73. Kimonis, V.; Al Dubaisi, R.; Maclean, A.E.; Hall, K.; Weiss, L.; Stover, A.E.; Schwartz, P.H.; Berg, B.; Cheng, C.; Parikh, S. NUBPL mitochondrial disease: New patients and review of the genetic and clinical spectrum. *J. Med. Genet.* **2021**, *58*, 314–325. [[CrossRef](#)]

-
74. Eis, P.S.; Huang, N.; Langston, J.W.; Hatchwell, E.; Schüle, B. Loss-of-Function NUBPL Mutation May Link Parkinson's Disease to Recessive Complex I Deficiency. *Front. Neurol.* **2020**, *11*, 1332. [[CrossRef](#)]
 75. Dickinson, M.E.; Flenniken, A.M.; Ji, X.; Teboul, L.; Wong, M.D.; White, J.K.; Meehan, T.F.; Weninger, W.J.; Westerberg, H.; Adissu, H.; et al. High-throughput discovery of novel developmental phenotypes. *Nature* **2016**, *537*, 508–514. [[CrossRef](#)]