

Article

# Mapping QTL Associated with Resistance to Avian Oncogenic Marek's Disease Virus (MDV) Reveals Major Candidate Genes and Variants

Jacqueline Smith <sup>1,\*</sup> , Ehud Lipkin <sup>2</sup> , Morris Soller <sup>2</sup>, Janet E. Fulton <sup>3</sup> and David W. Burt <sup>1,†</sup> 

<sup>1</sup> The Roslin Institute and Royal (Dick) School of Veterinary Studies R(D)SVS, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK; d.burt@uq.edu.au

<sup>2</sup> Department of Genetics, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel; ehud.lipkin@mail.huji.ac.il (E.L.); soller@mail.huji.ac.il (M.S.)

<sup>3</sup> Hy-Line International, P.O. Box 310, 2583 240th St., Dallas Center, IA 50063, USA; jfulton@hyline.com

\* Correspondence: jacqueline.smith@roslin.ed.ac.uk

† Current address: Office of Deputy Vice-Chancellor (Research), University of Queensland, 306 Carmody Road, St Lucia QLD 4072, Australia.

Received: 2 August 2020; Accepted: 26 August 2020; Published: 30 August 2020



**Abstract:** Marek's disease (MD) represents a significant global economic and animal welfare issue. Marek's disease virus (MDV) is a highly contagious oncogenic and highly immune-suppressive  $\alpha$ -herpes virus, which infects chickens, causing neurological effects and tumour formation. Though partially controlled by vaccination, MD continues to have a profound impact on animal health and on the poultry industry. Genetic selection provides an alternative and complementary method to vaccination. However, even after years of study, the genetic mechanisms underlying resistance to MDV remain poorly understood. The Major Histocompatibility Complex (MHC) is known to play a role in disease resistance, along with a handful of other non-MHC genes. In this study, one of the largest to date, we used a multi-faceted approach to identify quantitative trait locus regions (QTLR) influencing resistance to MDV, including an F<sub>6</sub> population from a full-sib advanced intercross line (FSIL) between two elite commercial layer lines differing in resistance to MDV, RNA-seq information from virus challenged chicks, and genome wide association study (GWAS) from multiple commercial lines. Candidate genomic elements residing in the QTLR were further tested for association with offspring mortality in the face of MDV challenge in eight pure lines of elite egg-layer birds. Thirty-eight QTLR were found on 19 chicken chromosomes. Candidate genes, microRNAs, long non-coding RNAs and potentially functional mutations were identified in these regions. Association tests were carried out in 26 of the QTLR, using eight pure lines of elite egg-layer birds. Numerous candidate genomic elements were strongly associated with MD resistance. Genomic regions significantly associated with resistance to MDV were mapped and candidate genes identified. Various QTLR elements were shown to have a strong genetic association with resistance. These results provide a large number of significant targets for mitigating the effects of MDV infection on both poultry health and the economy, whether by means of selective breeding, improved vaccine design, or gene-editing technologies.

**Keywords:** Marek's disease virus; QTL; GWAS; transcriptomics; FSIL; disease resistance; candidate gene

## 1. Introduction

Marek's disease (MD) represents a significant global economic and animal welfare issue. This immunosuppressive disease is responsible for an estimated 2 billion USD annual economic

loss to the global poultry industry [1], through bird mortality, lost egg production, and vaccination costs. The virus responsible, Marek's disease virus (MDV), is an  $\alpha$ -herpes virus that initially infects B-cells, experiences a latency period, and can then proceed to develop as an oncogenic disease after the infection of T-cells [2]. Furthermore, if birds do not succumb to MDV itself, they are often left severely compromised with secondary infections such as *Escherichia coli* (*E. coli*) [3]. The commercially utilized vaccines are not sterilizing vaccines. They prevent the formation of tumours, but do not prevent infection by MDV or shedding of the pathogenic virus [4]. Both vaccine and pathogenic MDVs are found in vaccinated flocks, resulting in the emergence of increasingly more virulent strains [5]. As more virulent strains emerge, vaccine treatments are becoming less and less effective [6].

In the race to keep ahead of this viral evolution, genetic selection is a possible tool to aid in breeding for viral resistance. Indeed, genetic selection following multiple generations of progeny challenge has been shown to improve survival in commercial chicken populations [7]. Selection in the poultry industry works relatively well, but knowing the causative elements for resistance would provide a route for much more precise selection. However, the underlying genetic changes, and the genes and specific variants impacting this genetic resistance are still largely unknown. For decades, researchers have sought to identify the genes responsible for MDV resistance, with limited success. It has become clear that many quantitative trait locus (QTL)/genes are involved in the resistance phenotype, with resistance not being a simple trait with one or a few major genes but a classic quantitative trait with many genes of small effect, thus making it difficult to identify causal variants [8].

MD is also of interest from a clinical perspective, as it can serve as a model for human lymphoma. MDV-induced lymphomas are found to over-express the Hodgkin's disease antigen CD30, with expression correlating with the viral Meq oncogene [9]. Indeed, after infection, the chicken CD30 promoter has also been shown to be hypo-methylated [10]. Identifying the genetic mechanisms underlying MDV resistance is therefore not only of great importance to the poultry industry but will also have implications for increasing our understanding of human cancers.

Different Major Histocompatibility Complex (MHC) haplotypes are known to confer different susceptibilities to the virus [11–13]. Studies have also reported the influence of non-MHC regions on resistance to MDV [14]. Several studies have implicated specific genes in MDV resistance/susceptibility. These include *GH1* [15], *SCYC1* [16], *SCA2* [17], *IRG1* [18], *CD79B* [19], *SMOC1*, and *PTPN3* [20]. These studies are usually done, not with relevant commercial lines, but with experimental or inbred lines and examine whole tissues, although recent work has investigated the host response to MDV in specific cells such as macrophages, which are an early target for the virus [21,22]. Recent studies on the role of long non-coding RNAs [23,24] and microRNAs (in both the host and virus) have also been carried out [25–27], including a study of serum exosomes from lymphoma-bearing birds [28]. In addition, the role of epigenetics in resistance to MDV has been studied, with regions of differential methylation between susceptible and resistant lines of birds highlighted [29,30].

Here, the availability of large-scale, phenotyped commercial populations, genome wide analysis technologies and an  $F_6$  advanced intercross line [31], has given us the opportunity to carry out, for the first time, a high-resolution analysis of genes underlying MDV resistance in commercially relevant populations. We use multiple genetic resources at our disposal, including the  $F_6$  population of an advanced full-sib inter-cross line (FSIL) previously analysed in a low-resolution study for MD resistance using microsatellite markers to identify genomic regions associated with survival following MD challenge [31]. Genomic DNA of the original 10 founder individuals and the subsequently produced  $F_6$  was available for fine mapping through genome sequencing, and/or genotyping using a genome-wide 600K SNP chip [32]. Furthermore, an extensive multi-generation (15) and multi-line (8) collection of DNAs from progeny-challenged males was available to further examine candidate genes and related variants associated with survival in the face of MDV infection.

In this report, we reveal for the first time MD as a true complex trait, controlled by many QTL. Integration of multiple lines of evidence ( $F_6$ , multi-generation/multi-line collection, host gene expression responses to viral infection, genome annotations, etc.) on a large scale enabled a high-resolution

analysis that predicted mutations within genes, miRNAs, and lncRNAs highly associated with MDV response in commercial egg production lines. This analysis not only provides new markers for MD resistance but also reveals more about the biology behind the mechanism of MDV susceptibility, information that should lead to more precise selection strategies in the future.

## 2. Materials and Methods

### 2.1. Experimental Animals

All procedures carried out on the birds involved in this study were conducted in compliance with Hy-Line International Institutional Animal Care and Use Committee guidelines.

#### 2.1.1. Full Sib Advanced Inter-Cross Line (FSIL)

F<sub>6</sub> birds from the FSIL were used to map QTL affecting MD resistance. The development of the FSIL F<sub>6</sub> challenge population has been previously described [31,33,34]. It was initiated from a cross of two partially inbred commercially utilized elite White Leghorn lines, known to differ in their resistance to MDV. Five independent FSIL families were developed and expanded over five generations. In all five families, the male parent was from the less resistant line and the female parent was from the more resistant line.

At the F<sub>6</sub> generation, 1615 chicks were challenged with vv+ MDV strain 686 following the protocol of Fulton et al. [7]. The experiment was carried out in two hatches. It was terminated at 152 days of age in the first hatch, and after 145 days of age in the second. Resistance was measured as survival time (age at death). For birds that survived to the end of the experiment, survival time was taken as age at end of the experiment. To standardize the two hatches, survival for all birds that survived to end of experiment was set to 149 days. This measure of resistance was used in all association tests.

This population was segregating for two MHC haplotypes (B<sup>2</sup> and B<sup>15</sup>). Given the known strong association of MHC-type with MD resistance, all analyses were done within MHC-type within each family.

#### 2.1.2. Pure line Pedigreed Sire Populations with Daughter Progeny Records for MDV Mortality

As part of the routine selection process within the Hy-Line (HL) breeding program, individual males were mated to multiple females to produce 30 half-sib female progeny per sire. The dams were MD vaccinated following normal industry practices. Progeny were vaccinated at 1 day of age with HVT/SB1 and at 7 days of age inoculated subcutaneously with 500 PFU of vv+ MDV (provided by Avian Disease and Oncology Lab, East Lansing MI). Mortality was recorded from 3 to 17 weeks of age (termination of experiment), as described in Fulton et al. [7]. The sire MD tolerance phenotype is the proportion of survivors among the daughters upon MD challenge.

Data were available from 15 generations of 8 elite lines (not the same generations in all lines), with 1081 to 1393 males per line having MD progeny mortality data, for a total of 9391 males. The lines included three different egg production breeds, namely five White Leghorn lines, two White Plymouth Rock lines and one Rhode Island Red line. The sires and their progeny from these 8 pure lines (all with pedigree data) were used to test functional genomic elements (genes, miRNAs, lncRNAs in the mapped QTL regions (QTLR)), for association with MD progeny mortality. Also tested were a group of singleton coding SNPs predicted to have deleterious effects on protein structure and function (henceforth called potentially functional mutations, see below for details).

### 2.2. Genomic Sequencing of Founder Birds of the FSIL

Genome sequence information was produced for the 10 F<sub>0</sub> founder individuals of the F<sub>6</sub> population. This was used to identify variants that differed in the genomes of the parental birds and may have functional impact. Sequencing the DNA from the 10 F<sub>0</sub> founder birds was carried out by the Edinburgh Genomics sequencing facility (Edinburgh, UK). Samples were prepared for sequencing

using 1 ug of genomic DNA following the TruSeq PCR free kit protocol (FC-121-3001, Illumina, Little Chesterford, UK). Resulting libraries were quality checked on an Agilent DNA 1000 bioanalyzer (Agilent Technologies, South Queensferry, UK) and then clustered onto HiSeq Rapid V2 flow cell at a concentration of 15 pM. Sequencing was carried out on an Illumina HiSeq 2500 using the HiSeq Rapid v2 SBS reagents (Illumina, Little Chesterford, UK), for 150 cycle paired end reads. Each of the 10 samples was sequenced to around 15× coverage. Quality of sequences was determined using FastQC [35], and mapping to the chicken reference genome (Gallgal4) was carried out using BWA (v0.7.0) [36]. The resulting bam file was sorted with Samtools (v0.1.18) [37]. Picard tools (v1.95) was then used to add read groups and mark duplicates (<http://broadinstitute.github.io/picard/>). The mpileup program within Samtools carried out SNP calling (with options: -q20 -Q20 -AB -ugf) and these variants were then filtered using the bcftools package within Samtools (v0.1.18).

### 2.3. Annotation of Genome Variants

Variants distinguishing the genomes of the F<sub>6</sub> parental birds were classified as exonic—synonymous or non-synonymous, intronic, 5'-upstream, 3'-downstream or intergenic, as determined by the SNPEff program (v3.6c) [38]. Non-synonymous coding variants were further designated as predicted to be highly deleterious to protein function (henceforth, “potentially functional mutations”), moderately deleterious or having a low likelihood of being deleterious to protein function. In all, a total of 5,718,725–6,154,628 variants (in all 5 males v all 5 females) were annotated within each sample. Some of the identified variants were used to test the QTLR, as described below.

### 2.4. 600 K High Density SNP Genotyping of F<sub>6</sub> Birds

F<sub>6</sub> birds were genotyped genome-wide with the high density 600k Affymetrix SNP array [32] for GWAS. DNA was available from 1615 MDV challenged birds. After quality control (QC), 1192 F<sub>6</sub> birds provided high quality genotypes (178, 234, 357, 221, and 202 birds from Families 1–5, respectively). Genotyping was performed using 200 ng of gDNA with the standard protocol for the Axiom Affymetrix platform (Thermo Fisher Scientific, Inchinnan, UK). Samples were amplified using the Axiom 2.0 Reagent Kit (#901758, Thermo Fisher Scientific), and the resultant product checked for both quantity and quality of fragmentation. QC was performed using absorbance assay for quantity and running a sample on a 4% agarose E-Gel (#G800804, Thermo Fisher Scientific) to check for fragmentation. The samples were then hybridised to the Axiom Chicken Genotyping array (#902148, Thermo Fisher Scientific). Hybridisation, wash, stain and scanning were carried out within the GeneTitan MC Instrument. The resulting .CEL files were loaded into Axiom Analysis Suite (v2.0) for first stage analysis.

### 2.5. GWAS of the F<sub>6</sub> Population to Identify MD Resistance QTL

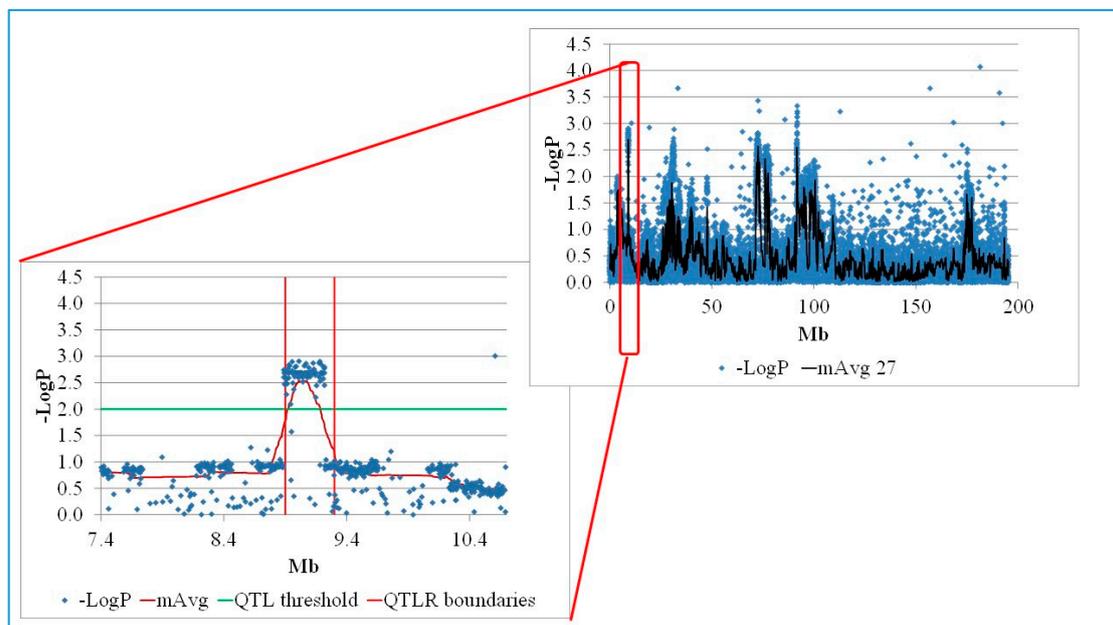
The F<sub>6</sub> genotypes were used to map QTL that impacted survival following MDV challenge. After applying Affymetrix standard QC the remaining markers were further filtered to remove markers with minor allele frequency  $\leq 0.01$  and markers with significant deviation from Hardy–Weinberg Equilibrium ( $p \leq 0.001$ ). An independent GWAS was then carried out within each of the 5 families of this study, using JMP Genomics SNP-Trait association Trend test (JMP Genomics, Version 9, SAS Institute Inc., Cary, NC, USA, 1989–2019). Survival to the end of the experiment was taken as the Censor Variable; age at death was taken as a survival trait and MHC as a class variable and fixed effect. To obtain experimental significance thresholds, the proportion of false positives (PFP) [39,40] was used to correct for multiple tests.

### 2.6. Identification of QTL and Their Confidence Intervals

Although many marker  $\times$  family combinations were nominally significant to highly significant (comparison wise  $p \leq 0.05$  to  $p \leq 0.001$ ), very few remained significant after PFP corrections for multiple tests. Nevertheless, visual inspection of the chromosomal Manhattan plots by families showed distinct clusters of markers with high  $-\text{LogP}$  values intermixed with markers with low  $-\text{LogP}$  values (Figure 1).

Therefore, following Lipkin et al. [41], we identified QTL by using a moving average of  $-\text{LogP}$  (mAvg) to smooth the Manhattan plots. We used a window size of  $\sim 0.1$  Mb (27 markers) with a step size of 1 marker and a critical threshold mAvg of  $-\text{LogP} = 2.0$  ( $p = 0.01$ ) to declare significance, and Log drop 1 [42] to define QTLR boundaries.

Many QTLR had overlapping boundaries across two or more families, thus providing replication and increased assurance of significance. Conservatively assuming such overlaps represents the same underlying genetic element, we combined these QTLR within and across families, taking the start and end log-drop boundaries of the QTLR as the first upstream and last downstream marker across all families in the QTLR. Table 1 shows final list of 38 QTLR after consolidation within and across families. Also shown are the number of genes (Ensembl database v79) and non-coding RNAs [43] found in each QTLR.



**Figure 1.** Manhattan plot of a significant  $F_6$  QTL (Family 3: QTLR 1, Chromosome 1).  $\text{LogP}$ ,  $-\log_{10}P$  of a single marker; mAvg, moving average of  $-\text{LogP}$  values of a window of 27 markers; QTLR boundaries, boundaries of the quantitative trait locus region obtained by Log Drop 1 (see Section 2.6).

**Table 1.** Mapped QTLRs.

QTLR	Chr	Start	End	Length	Genes	lncRNAs
1	1	8,854,589	9,261,317	406,729	0	1
2	1	13,268,810	14,294,877	1,026,068	12	34
3	1	52,012,381	52,517,953	505,573	2	4
4	1	71,738,977	73,085,956	1,346,980	16	28
5	1	75,288,912	77,765,499	2,476,588	86	79
6	1	91,482,501	91,803,475	320,975	10	9
7	1	101,761,945	104,477,713	2,715,769	31	37
8	1	109,395,871	110,913,451	1,517,581	20	32
9	1	169,656,958	172,228,091	2,571,134	32	44
10	1	174,394,951	175,634,386	1,239,436	23	33
11	1	194,193,118	194,788,548	595,431	18	23
12	2	15,960	883,257	867,298	45	20
13	2	46,423,161	46,868,789	445,629	11	12
14	2	105,493,833	108,988,920	3,495,088	34	47
15	2	125,168,753	127,089,304	1,920,552	25	34
16	2	138,767,830	139,701,774	933,945	4	7
17	3	108,220,206	109,260,649	1,040,444	13	9

Table 1. Cont.

QTLR	Chr	Start	End	Length	Genes	lncRNAs
18	4	8,308,498	11,268,107	2,959,610	54	99
19	4	84,393,173	88,579,400	4,186,228	68	73
20	5	7,568,851	8,147,837	578,987	4	3
21	5	18,806,924	19,673,354	866,431	4	3
22	6	2,077,640	2,709,412	631,773	19	13
23	6	29,536,109	29,817,337	281,229	10	16
24	6	31,006,769	31,448,342	441,574	7	6
25	7	13,062,871	16,436,053	3,373,183	55	82
26	10	22,643	1,713,384	1,690,742	63	28
27	11	7,397,790	8,440,259	1,042,470	9	28
28	12	8,996,686	9,432,693	436,008	15	16
29	13	10,363,430	12,176,727	1,813,298	27	40
30	14	8,085,563	9,335,685	1,250,123	38	50
31	14	13,138,194	15,087,518	1,949,325	82	41
32	16	1630	490,907	489,278	47	14
33	17	3,442,598	5,634,042	2,191,445	64	84
34	18	3,196,488	4,093,129	896,642	22	32
35	24	4,489,675	5,514,833	1,025,159	42	32
36	26	4,378,168	5,036,699	658,532	22	17
37	27	1,540,112	2,270,461	730,350	23	11
38	28	1,282,726	1,571,011	288,286	15	12

QTLR—ordinal number of the QTLR after consolidation within and across families (see Section 2.6); Chr—chromosome; Start/End—Galgal4 QTLR coordinates of the first and last SNP in the QTLR; Length—Size of the QTLR (bp); Genes and lncRNA—number of genes (Ensembl database v79) and non-coding RNAs [43] found in the QTLR.

## 2.7. RNA-Seq Analysis

### 2.7.1. MDV Challenge Experiment

An experiment measuring host response to MDV viral strain 686 challenge (based on gene expression measured by a whole genome RNA-seq assay) was carried out with 10 male and 10 female birds from HL W36 commercial production hybrids. The parental lines used to produce the W36 commercial hybrid are closely related to the lines used to develop the F<sub>6</sub> experimental cross. Previously it was reported that males are more resistant to MD than females [44–46]. On this basis, we compared differentially expressed genes (DEGs) of males and females as surrogate for comparison of chicks from more- and less-resistant lines. Hence, gender balanced groups were used for the RNA-seq challenge experiment, with 5 males + 5 females in each of the challenged and control groups. This allowed us to identify viral response genes in each sex. Thus, this experiment included 4 comparisons: Females—challenged v control; Males—challenged v control; Controls—males v females; Challenged—males v females. Although some of the DEGs in the cross-sex comparisons will be due to sex rather than response to MDV, others can be expected to reveal some of the genes behind the host immune responses and potential differential susceptibility.

To reflect infection in the field, ten 6-day old HL W36 commercial birds (5 males + 5 females) were infected with 500 pfu of the very virulent plus (vv+) MDV strain 686, by subcutaneous injection in the neck (virus kindly provided by the USDA/Avian Disease and Oncology Lab, East Lansing, MI, USA). All birds were vaccinated at 1 day of age with HVT/SB1 vaccine. Spleen tissues from these 10 challenged and from 10 unchallenged control birds were harvested at 4 dpi, flash frozen in liquid nitrogen and stored at −80 °C for subsequent RNA preparation and RNA-Seq analysis. All birds were housed together, with challenged and control chicks separated by 2.44 m.

### 2.7.2. RNA Preparation

RNA was prepared from the 4-dpi flash-frozen spleen tissues of each of the above 20 chicks after homogenization in Qiazol reagent (Qiagen, Manchester, UK), and subsequent preparation using RNeasy RNA isolation kit (Qiagen, Manchester, UK) as per the manufacturer's guidelines. RNA was resuspended in dH<sub>2</sub>O.

### 2.7.3. Transcriptomic Sequencing and Analysis

Strand-specific 100 bp paired-end RNA-Seq was carried out by GATC (Konstanz, Germany) using an Illumina HiSeq2500 genome sequencer. 19,510,400–26,596,800 reads representing at least 15× coverage were produced for each sample. Sequencing quality was assessed by means of FastQC [35] and reads mapped to the reference genome (Galgal4) with Tophat2 (v2.0.14) [47] using the Bowtie2 aligner (v2.2.3) [48]. Untrimmed reads with stated insert sizes were mapped using the Ensembl 79 genome annotation. 78.1% of reads mapped to the Galgal4 reference genome. The Cuffdiff package within Cufflinks (v2.2.1) was used to quantify transcripts and determine differential expression between experimental groups [49].

## 2.8. QTLR Tests

### 2.8.1. Selection of Candidate Genomic Elements Underlying the 38 QTLR for Further Testing and Validation in the 8 HL Elite Pure Lines

In principal, a random selection of markers from a QTLR region could be used for testing in the 8 elite-sire validation populations. However, we decided to take a step forward, and search the QTLR for attractive candidate elements that could be tested in the validation populations. The assumption being that for a marker to show association across a number of populations, it must be in high LD with the causative mutation in all of these populations. For this to hold, the marker must be very close to the causative mutation, which is what we would expect for markers taken from the quantitative trait element itself.

### 2.8.2. Identification of Candidate Genes within QTLR

The BioMart data mining tool within the Ensembl database (release 79) was used to identify genes lying under the 38 QTLR identified from the F<sub>6</sub> data (Table 1). This information was then compared with the gene expression data (from RNA-Seq analysis of the MDV challenge experiment) in order to identify potential candidate genes for MD resistance. The DAVID analysis tool (v6.8) (<https://david.ncicrf.gov/home.jsp>) identified gene ontology (GO) terms associated with DEGs under each QTLR. Following this, Ingenuity Pathway Analysis software (IPA v2.3) (<https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/>) revealed which canonical pathways were perturbed by MDV infection in the host, and allowed us to analyze the gene interaction networks involved in the host response. Likely upstream regulators of differentially regulated genes were identified and DEGs in male and female birds were compared.

### 2.8.3. Selection of Genetic Markers within QTLR for Genotyping

Markers located in genomic elements underlying each QTLR (genes, miRNAs, lncRNAs, and potentially functional mutations) were identified from genome sequences aligned to the reference chicken genome for each of the 8 elite lines utilized in this study. SNPs predicted to alter amino acids were preferentially chosen, but variants were also selected to encompass the entire gene where possible. All SNPs selected were validated using 200 samples from cohorts separated by 13–15 generations from each line to confirm that the marker was polymorphic (i.e., identified both homozygote and heterozygote individuals), and to define the minimum number of markers required to define a haplotype within each genomic element (gene, miRNA, or lncRNA) in each line. This minimum number of SNPs was then used to identify the haplotype status of each individual bird. The elite

lines were selected for MD resistance, but the selection was at a low level, as selection emphasis was for production-related traits including egg numbers, shell quality and feed efficiency. If SNPs were homozygous in both generations, then the SNP composition was assumed monomorphic for the intervening and subsequent generations.

#### 2.8.4. Selection of Candidate Coding Genes

Twenty candidate coding genes (Table 2) were selected for further study in the 8 elite sire lines based on (i) position under a QTLR defined by this or previous genetic studies, (ii) known biology of the gene (e.g., involvement in innate immunity, cell death, or cancer), (iii) polymorphisms (presence of variants), (iv) differential expression between challenged and control, or more and less resistant birds—either in this or previous studies (e.g., Smith et al. 2011 [18]), and (v) allele specific expression [50]. Specific details for the detection assays are provided in Supplementary Table S1.

**Table 2.** Candidate genes in the QTLRs, by location (Galgal4).

QTLR	Gene	Description	Chr	Start	End	Length
5	CSTA	Cystatin A (Stefin A)	1	76,253,583	76,258,549	4967
5	LAG3	Lymphocyte-activation gene 3	1	76,676,057	76,678,087	2031
5	C1S	Complement C1s subcomponent	1	76,859,359	76,868,459	9101
7	ADAMTS5	ADAM metalloproteinase with thrombospondin type 1 motif, 5	1	102,403,453	102,442,746	39,294
8	SUPT20H	Spt20 homolog, saga complex component	1	171,574,831	171,612,194	37,364
10	FLT3	Fms related tyrosine kinase 3	1	175,281,894	175,309,466	27,573
11	RELT	RELT tumor necrosis factor receptor	1	194,551,973	194,555,607	3635
13	TRANK1	Tetratricopeptide repeat and ankyrin repeat containing 1	2	46,537,531	46,576,714	39,184
19	CTNNA2	Gallus gallus catenin (cadherin-associated protein), $\alpha$ 2	4	86,948,459	87,057,215	108,757
29	HAVCR1	Hepatitis A virus cellular receptor 1 precursor	13	10,778,915	10,785,137	6223
29	TIMD4	T-cell immunoglobulin and mucin domain-containing protein 4	13	10,789,088	10,803,683	14,596
30	SOCS1	Suppressor of cytokine signaling 1	14	8,935,374	8,936,082	709
32	TAP1	Antigen peptide transporter 1	16	67,623	76,124	8502
32	IL4I1	Interleukin 4 induced 1	16	191,143	194,187	3045
33	TLR4	Toll-like receptor 4 precursor	17	3,566,454	3,571,907	5454
33	BRINP1	Bmp/retinoic acid inducible neural specific 1	17	3,985,882	4,064,290	78,409
33	CD7	T-cell antigen CD7	18	3,681,786	3,686,083	4298
35	TAGLN	Transgelin	24	5,229,236	5,232,028	2793
36	TREML2	Triggering receptor expressed on myeloid cells isoform 1	26	4,693,285	4,699,278	5994
37	CD79B	CD79b molecule	27	1,622,587	1,627,022	4436

QTLR—ordinal number of the QTLR (Table 1); Chr—chromosome; Start/End—Galgal4 gene coordinates (bp); Length—size of the gene (bp).

#### 2.8.5. Selection of Candidate Non-Coding RNAs

With many causal variants for quantitative traits known to reside within regulatory elements [51], we deemed it important to examine non-coding transcripts as well. In contrast to the candidate genes, these were selected more or less on the basis of position underlying QTLR alone, with no information on potential biological function.

**Micro RNAs (miRNAs):** A total of 46 miRNAs were found underlying the QTLR. Sequence information was used to identify markers segregating in HL lines. For some miRNAs, polymorphisms were not found, while for others multiple polymorphic sites were observed. Since most of the miRNAs are very short, haplotype information could be derived only for some of them. Ultimately, 17 miRNAs were chosen for further association analysis (Table 3). Detection assays for 30 SNPs lying within these 17 transcripts were developed (Supplementary Table S1).

**Long non-coding RNAs (lncRNAs):** With a view to testing some lncRNAs potentially associated with MDV resistance, five small QTLR were chosen which harboured only a few (1–6) lncRNAs. From these RNAs, up to 5 lncRNAs per QTLR were tested for association with MD. Thus, there was a relatively high likelihood that one of these would be the genomic element housing the causative mutation. From among these, a further 17 transcripts were chosen for further association analysis (Table 4). Two lncRNAs overlapped (labelled with negative distance). Detection assays for 80 SNPs across these 17 transcripts were developed (Supplementary Table S1).

**Table 3.** Candidate miRNAs in the QTLRs, by location (Galgal4).

QTLR	Gene Name	Ensembl Gene ID	Chr	Start	End	Length
5	gga-mir-6553	ENSGALG00000026406	1	76,381,420	76,381,519	100
7	gga-mir-1738	ENSGALG00000025522	1	104,085,133	104,085,223	91
12	gga-mir-1749	ENSGALG00000025364	2	569,326	569,423	98
12	gga-mir-6710	ENSGALG00000027384	2	724,388	724,497	110
18	gga-mir-1462	ENSGALG00000025207	4	8,488,386	8,488,495	110
25	gga-mir-1659	ENSGALG00000025154	7	13,177,026	13,177,126	101
25	gga-mir-6624	ENSGALG00000026751	7	15,758,919	15,759,028	110
25	gga-mir-1570	ENSGALG00000025238	7	16,274,218	16,274,317	100
26	gga-mir-6660	ENSGALG00000028738	10	1,076,742	1,076,851	110
30	gga-mir-1644	ENSGALG00000025220	14	8,124,516	8,124,601	86
31	gga-mir-6643	ENSGALG00000027935	14	13,143,450	13,143,559	110
31	gga-mir-1588	ENSGALG00000025419	14	14,125,510	14,125,611	102
31	gga-mir-1715	ENSGALG00000025576	14	14,285,835	14,285,936	102
35	gga-mir-6612	ENSGALG00000026265	24	4,779,927	4,780,036	110
35	gga-mir-6671	ENSGALG00000027111	24	5,035,783	5,035,892	110
35	gga-mir-1745	ENSGALG00000026550	24	5,163,387	5,163,473	87
37	gga-mir-6573	ENSGALG00000028331	27	1,532,553	1,532,671	119

QTLR—ordinal number of the QTLR (Table 1); Chr—chromosome; Start/End—miRNA bp coordinates, Length—size of the miRNA (bp).

**Table 4.** Candidate lncRNA in the QTLRs, by location (Galgal4).

QTLR	Gene ID; Transcript ID; Stable Transcript ID	Chr	Start	End	Length	Distance	Strand	Exons
1	R192;R192.1;RI_1_8950428_8950983_1_1	1	8,950,428	8,950,983	556		+	1
3	R1291;R1291.1;RI_1_52122457_52122949_1_1	1	52,122,457	52,122,949	493	43,171,475	-	1
3	R1292;R1292.1;RI_1_52183105_52322961_6_1	1	52,183,105	52,322,961	139,857	60,157	-	6
3	R1293;R1293.2;RI_1_52481252_52482976_1_1	1	52,481,252	52,482,976	1725	158,292	-	1
3	R1293;R1293.3;RI_1_52481282_52488508_2_1	1	52,481,282	52,488,508	7227	-1693	-	2
20	R12773;R12773.1;RI_5_7604020_7604920_1_1	5	7,604,020	7,604,920	901		+	1
20	R12779;R12779.1;RI_5_7727117_7728010_1_1	5	7,727,117	7,728,010	894	122,198	-	1
20	R12784;R12784.1;RI_5_7960049_7961460_1_1	5	7,960,049	7,961,460	1412	232,040	+	1
21	R13147;R13147.6;RI_5_18814123_18815319_2_1	5	18,814,123	18,815,319	1197	10,852,664	+	2
21	R13148;R13148.1;RI_5_18861679_18862530_1_1	5	18,861,679	18,862,530	852	46,361	+	1
21	R13149;R13149.1;RI_5_19632618_19793846_5_1	5	19,632,618	19,793,846	161,229	770,089	-	5
24	R15316;R15316.1;RI_6_31138889_31157672_3_1	6	31,138,889	31,157,672	18,784		-	3
24	R15317;R15317.1;RI_6_31199730_31200561_1_1	6	31,199,730	31,200,561	832	42,059	-	1
24	R15318;R15318.1;RI_6_31204510_31210163_7_1	6	31,204,510	31,210,163	5654	3950	+	7
24	R15319;R15319.1;RI_6_31263768_31264516_1_1	6	31,263,768	31,264,516	749	53,606	+	1
24	R15320;R15320.2;RI_6_31277294_31279639_1_1	6	31,277,294	31,279,639	2346	12,779	-	1
24	R15321;R15321.1;RI_6_31368941_31369400_1_1	6	31,368,941	31,369,400	460	89,303	-	1

QTLR—ordinal number of the QTLR (Table 1); gene id; transcript id; stable transcript id—unique lncRNA identifier; Chr—chromosome; Start/End—lncRNA bp coordinates; Length—size of the lncRNA (bp); Distance—bp between lncRNAs on the same chromosome; negative labelled value—overlap between this lncRNA and the previous one. Data from [43].

### 2.8.6. Selection of Potentially Functional Variants in Coding Genes

Analysis of the complete sequences of the 10 F<sub>0</sub> birds showed multiple SNPs distinguishing the founder birds that could impact gene function. Alternative variants fixed in one parental line compared to the other and which were predicted to cause highly deleterious changes in the gene protein product, were chosen as candidates for further study. Across the genome as a whole, a total of 252 variants of this nature were identified in 217 genes, with 22 of these variants lying within genes underlying the 38 QTLR. Twelve of these variants (representing 9 genes) were segregating in one or more of the 8 elite lines and were chosen for further studies in these lines. These 12 SNPs are summarized in Table 5. Specific details on the detection assays are provided in Supplementary Table S1.

**Table 5.** Markers in the QTLRs with genetic variations predicted to be functionally deleterious to the gene product.

QTLR	Gene	Marker	Chr	bp	Consequence	Effect
14	ENSGALG00000015219	MD-07	2	107,385,545	S/*	Stop gained
25	gga-mir-7474	MD-16	7	14,743,048	actgat	Frame shift
26	ENSGALG00000001269	MD-20	10	48,846	P/X	Frame shift
29	ENSGALG00000004078	MD-23	13	11,619,922	gtt/gGTTTt	Frame shift
30	IGSF6	MD-28	14	8,738,347	L/X	Frame shift
32	BG1	MD-29	16	117,493		Splice site acceptor
32	BG1	MD-30	16	118,808	Q/*	Stop gained
32	ENSGALG00000028367	MD-32	16	206,100		Splice site acceptor
32	ENSGALG00000028367	MD-33	16	206,241		Splice site acceptor
36	ENSGALG00000003188	MD-44	26	4,629,538	*C	Stop lost
37	SCN4A	MD-45	27	1,646,318	aga/agaCAGAAAT	Frame shift
37	SCN4A	MD-46	27	1,646,601	Q/*	Stop gained

QTLR—ordinal number of the QTLR (Table 1); Chr—chromosome; bp—Galgal4 location of deleterious variant; Consequence—amino acid change; \*—stop codon.

### 2.8.7. Selecting QTLR Markers from the Candidate Genomic Elements, for Genotyping in the 8 Elite HL Pure Lines

Markers located in the genomic elements underlying the QTLR (genes, miRNAs, lncRNAs, and singleton variants with potentially functional effects), were identified from genome sequences aligned to the reference chicken genome for each of the lines utilized in this study. For the candidate genes underlying the QTLR, SNPs predicted to alter amino acids were preferentially chosen. For all genomic elements (not including the singleton functional mutations), variants were also selected to encompass the entire gene when possible, so as to define the haplotype status of each bird for subsequent association tests. All SNPs selected were validated using the same approach as defined above.

### 2.9. Genotyping of the Eight Elite Lines

Genomic DNA was extracted from whole blood using salt/ethanol precipitation and stored at  $-20\text{ }^{\circ}\text{C}$  until use. DNA was diluted in  $\text{dH}_2\text{O}$  to  $25\text{ ng}/\mu\text{L}$ . Genotyping was done as single-plex assays on the SNPLINE system (LGC, Hoddesdon, UK). All assays used KASP chemistry, which is a fluorescent-based competitive allele-specific detection method [52]. Genotyping was done in 1536 well plates, and results analysed with Kraken software (LGC, Hoddesdon, UK) as previously described [53]. Specific details on primers for each assay are provided in Table S1.

### 2.10. QTLR Association Analysis

Association of markers and marker-haplotypes with sires progeny-tested for post-challenge survival in the 8 HL elite lines was tested using JMP Genomics SNP-Trait association Trend test (JMP Genomics, Version 9, SAS Institute Inc., Cary, NC, USA, 1989–201). When analysing a line-marker combination, the generation and MHC background were taken as class variables and fixed effects, and when analysing a marker combining all lines together, the line was added to the fixed effects. To obtain experimental significance thresholds, the proportion of false positives (FP) [39,40] was used to correct for multiple tests.

## 3. Results

### 3.1. High Resolution MD QTL Mapping in Advanced Full-Sib Inter-Cross $F_6$ Families

Following genotyping of the  $F_6$  families by the Affymetrix 600K HD SNP chip, 40–43% of SNPs (depending on family) were found to be informative and passed QC (total, 238,777 to 259,098

informative SNPs per family; average of 246,400 SNPs in a family). The distribution of the marker  $p$ -values was fairly uniform over the range of allele frequencies. A clear excess of small  $p$ -values (an indication of presence of true QTL effects) was not seen. In accordance with this, only in families 2, 4 and 5, were markers significant at  $\text{PFP} \leq 0.2$ . However,  $n_1$ , the estimated number of falsified null hypotheses at  $\text{PFP} \leq 0.2$  [39,40], ranged from 1718 in Family 1 (0.7% of all markers tested in the family) to 17,385 in Family 2 (6.7% of markers tested), indicating the actual presence of true QTL effects.

As described in Methods, many markers were nominally significant, and distinct clusters of nominally significant markers intermixed with non-significant markers were clearly seen. Hence, to identify QTL we used a moving average of marker  $-\text{Log}_{10} p$  values (mAvg) across a window size of 0.1 Mb (27 SNPs) to smooth the plots. Examining Manhattan plots (Figure 1 and Supplementary Figure S1), showed that the smoothing did yield a good monotonic plot, allowing identification of QTL and the use of the drop method [42] to define their boundaries.

Using a threshold of  $\text{mAvg} \geq 2.0$  and QTLR boundaries defined by Log Drop 1, a total of 57 regions meeting our criteria for significance were identified, ranging from 4 to 16 in a family. In some families, several regions clustered within 1 Mb of one another. Nine such chromosome-by-family clusters were identified, containing 2–4 significant sub-regions per cluster. Conservatively, we counted each cluster as representing a single within-family QTLR. Similarly, aligning QTLR maps of the five families, showed that six significant regions overlapped or were within 1 Mb of one another across two families. Here too, counting the same QTLR across different families only once, resulted in a final total of 38 QTLR being identified (Table 1). These 38 QTLR were the focus of further examination of variants as defined by multiple methodologies employed below.

### 3.2. Annotation of QTLR

Transcripts of genes located under the 38 mapped QTLR were identified by using the Biomart tool within the Ensembl database (v79) (Supplementary Table S2). The number of known annotated genes in each QTLR ranged from 0 to 86, for a total of 1072 genes. Using the annotation developed and described by Kuo et al. [43], which was based on characterisation of full-length transcripts using long-read sequencing and a bioinformatics pipeline to define coding and non-coding RNAs, we identified 1–99 lncRNAs within each defined QTLR, with a total 1153 lncRNAs (Table 1).

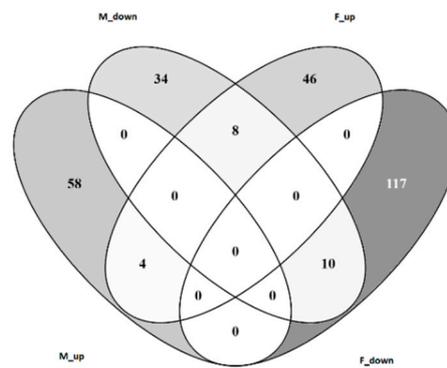
### 3.3. Identification and Analysis of SNPs in $F_0$ Parental Birds

To investigate variants residing in regions of the genome potentially associated with MDV resistance, the 10 founder parents that gave rise to the  $F_6$  population used in this study were sequenced and variants determined between males and females. Combined across both lines 8,273,112 SNPs were identified. Examining the 38 QTLR within the genome highlighted 133,394 SNPs in these regions, which were fixed for alternative SNP variants with one parental line compared to the other (Supplementary Table S3). With a view to testing for causality, any of these variants could then be tested for genetic association with the progeny-test for survival in the eight elite HL lines.

### 3.4. Transcriptomic Analysis of Male and Female Commercial Birds Challenged with MDV

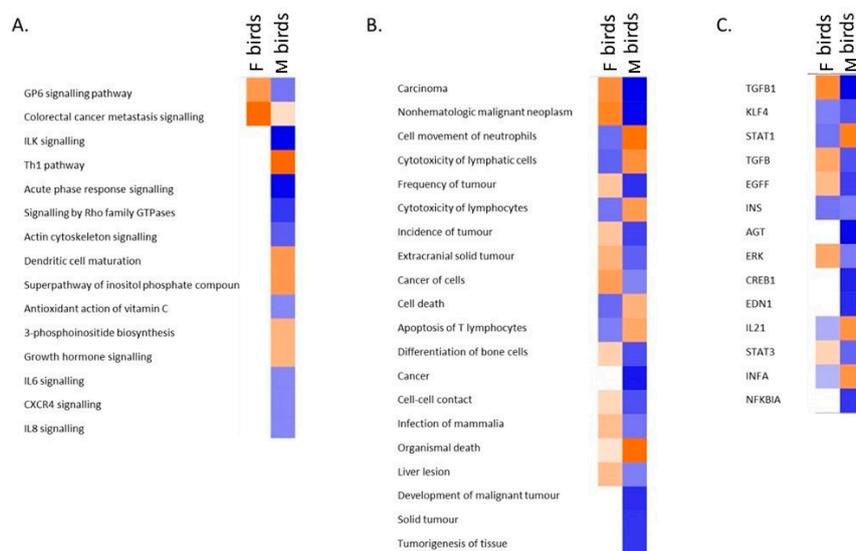
Based on the observation noted above, that male and female birds tend to show differential resistance, we examined the expression profiles of five male and five female W36 birds challenged with  $\text{vv}^+$ MDV, and five male and five female W36 non-challenged control birds (see Section 2.7). Gene expression results from this transcriptomic analysis are presented in Table S4.

First, we compared gene expression in challenged and control chicks within each sex of W36 birds to find that 185 genes (58 up- and 127 down-regulated) were differentially expressed in male chicks, while there were 114 genes (62 up- and 52 down-regulated) differentially expressed upon challenge of the female birds. When these genes are compared across sex, the individual responses are seen to be quite different, with minimal overlap of DEGs between the male and female groups (Figure 2).



**Figure 2.** Venn diagram showing the intersection of host responses in male and female W36 birds. M-up—number of genes up regulated in male birds; M-down—number of genes down regulated in male birds; F-up—number of genes up regulated in female birds; F-down—number of genes down regulated in female birds.

When the different responses were compared, it is seen that biological pathways predominantly involved in the host response in female chicks include activation of cancer signalling, whereas the DEGs identified in the male birds have roles in the Th1 immune response and dendritic cell maturation amongst others (Figure 3A). The diseases and biofunctions associated with the genes involved in each response are depicted in Figure 3B. Strikingly, with MD being an oncogenic disease, functions related to cancer and tumorigenesis are seen to be up-regulated in female birds, but are repressed in the male chicks. Investigation of genes acting as upstream regulators of identified DEGs again indicates differential regulatory mechanisms in each sex. *TGFB* appears to be a key regulator of mechanisms in females, whereas *STAT1* (an interferon response gene) is suggested to have a more prominent role in the males (Figure 3C).



**Figure 3.** Comparison of the gene expression underlying the host response in W36 birds. Orange colours represent up-regulation while blue indicates down-regulation. Differential activation/repression is indicated as shown: (A). Biological pathways (B). Diseases and biofunctions (C). Upstream regulators.

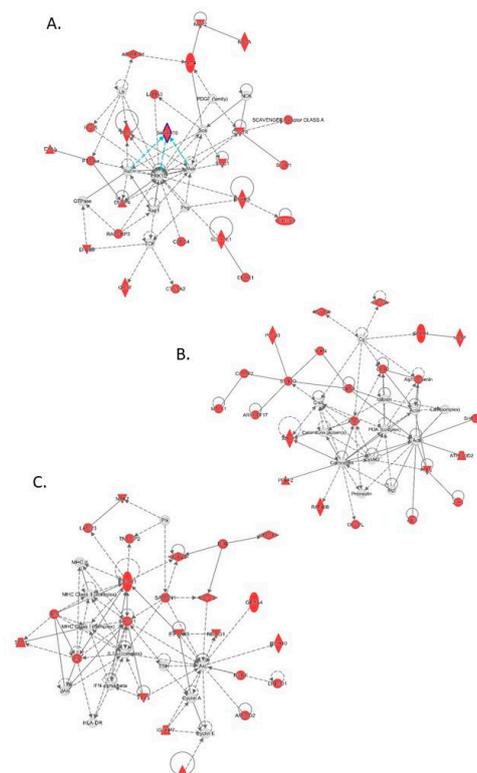
Examination of the unchallenged control birds from each group (male and female W36 birds) identified 302 genes which were expressed in an inherently different manner between the sexes. So, these DEGs, being independent of MD challenge, are due to gender differences and/or intrinsic immune differences, manifesting in differing resistance to infection. Significant DEGs include the obvious

immune related genes, but also various serine proteases and carboxypeptidases, which may suggest a role in resistance to MDV. Genes previously implicated in MDV resistance including MHC genes, *IRG1* [18] and *SCYC1* [16], were also highlighted.

### 3.5. Analysis of Differentially Expressed Genes Located in QTLR

When DEGs are located within QTLR affecting MD resistance, it is plausible that these genes may be the causative quantitative trait gene (QTG) responsible for the QTLR effect. Based on this hypothesis, the list of genes located in the 38 QTLR identified in this study was compared to the list of DEGs. Genes that were both expressed under challenge and that were also located under one of the 38 mapped QTL, were considered likely candidate QTG (Supplementary Table S5).

To determine if any particular types of genes were represented by the DEGs underlying significant genomic regions (identified either from this or previous studies [31,33,54–56]), the gene ontologies associated with these DEGs were studied. Table S6 shows that immunoglobulins, stimulus response genes and genes involved in regulating the immune response are all highly represented. Analysis of the biological pathways controlled by these genes was also examined. To look for possible common pathways, network analysis was performed. Three networks stood out significantly: cell movement, cell signalling, and cancer (Figure 4A); cell signalling and nervous system development (Figure 4B); antimicrobial resistance, inflammation, and cell death (Figure 4C). When candidate upstream regulators of these genes were investigated, the cytokines *TNF*, *EDN1*, *IL1B* and *IFNG* were all indicated as potentially regulating many of the genes under study. Interestingly, the *TNF* gene is located in QTLR 18 on Chr 4,387,071 bp downstream of the top window in this QTLR.



**Figure 4.** Gene networks represented by differentially expressed genes found under the 38 significant QTLR. Genes coloured in red are genes within the network which are in the dataset under study. (A) Cell movement, cell signalling and cancer; (B) Cell signalling and nervous system development; and (C) Antimicrobial resistance, inflammation, and cell death.

### 3.6. Validating QTLR for Association with MDV Resistance

#### 3.6.1. Populations, Genomic Elements and Markers

In a unique aspect of this study, QTLR identified in the F<sub>6</sub> mapping population were re-tested for association with MD resistance in 8 pure lines maintained under selection for multiple commercial production traits at HL. Markers for QTLR validation were chosen from among the various classes of genomic elements identified underlying the QTLR, namely genes, miRNAs and lncRNAs. Also included were a group of singleton coding SNPs predicted to have deleterious effects on protein structure and function (the functional mutations).

#### 3.6.2. Genetic Association Tests

When possible, a number of markers were genotyped in each candidate gene, so that association could be based on marker haplotypes rather than individual markers. When this is done, all of the information is in the 'Haplotypes'. The individual markers are usually in complete LD and hence do not add anything to the analysis. Thus, for statistical analyses what is important are the number of element x line tests, rather than number of marker x line tests. We counted an element x line test as significant if the haplotype was significant, or if one or more markers in the element was significant.

A total of 355 markers located in 66 genetic elements (genes, miRNA, lncRNA, and functional mutations) underlying 26 of the 38 QTLR were tested for association with MD mortality in the 8 elite lines (Supplementary Table S7). In 46 of the elements, multi-marker haplotypes were also tested. The association tests were carried out separately within each of the 8 elite lines, and also in data combined across the 8 lines, yielding 2032 *p*-values. A total of 387 element by line tests were performed.

Significance of individual candidate genes: Twenty genes located in the QTLR were tested by a total of 127 gene x line tests of markers encompassing the entire gene. A test was considered significant if the haplotype x line test was significant, or if any one of the markers encompassing the haplotype were significant. All but *IL411*, *TAGLN* and *CD79B* (in QTLR 32, 35 and 37), were significant in two or more marker x line tests. Thus, 17 genes may be taken as strong candidates to be a QTG (Table S8).

Association of markers located in miRNA: A total of 17 miRNAs were chosen (Table 3), and 28 markers tested for association (Supplementary Table S7) for a total of 96 element x line tests. Only 5 significant to highly significant *p*-values were obtained. (Supplementary Table S8).

Significance of individual miRNAs: Significant results were found in only five out of the 17 miRNA tested. Most interesting was gg-mir-6553 with the highest proportion of significant tests among the miRNAs (Supplementary Table S8). This miRNA is located in QTLR 5, along with three significant genes (*CSTA*, *LAG3*, and *C15*). The results indicate two putative causative quantitative elements in this QTLR: one upstream in the region of *CSTA*, the other downstream in the region of *LAG3* and *C15*.

Associations of markers located in lncRNAs: A total 17 lncRNAs were chosen (Table 4), and 80 markers were tested for association (Table S7) in a total of 114 element x line tests, of which 9 were significant (Supplementary Table S8).

Significance of individual lncRNAs: Significant results were found with 9 of the 17 lncRNAs tested (Table S8). QTLR worth mentioning are 20 and 21 on Chr5, both tested by lncRNAs only, are detailed in Supplementary File S1.

Associations of QTLR functional mutations in coding genes: Twelve potentially functional mutations were tested in a total of 50 element x line tests, of which six were significant (Table 5).

Significance of genes harbouring potentially functional mutations: Significant results were found in 5 of the 9 QTLR genes harbouring potentially functional mutations. In Line WPR1, the 2 markers in the gene *BG1* (markers MD-29 and MD-30 in QTLR 32 on Chr 16) had practically the same *p*-values as the markers of the candidate gene *TAP1* (Supplementary Table S8), thus strengthening the results of that gene. It should be noted that *BG1* and *TAP1* lie within the chicken MHC, as noted above, a locus known to impact MDV resistance.

### 3.6.3. Summary of Element x Line Tests

Supplementary Table S9 summarises the element x line tests by type of element. Tests of the coding genes are outstanding in the high proportion of significant element x line tests; more than double that of the other genomic elements. Since the candidate genes were chosen on the basis of much more information than the other classes of elements, the high proportion of significant tests can be taken to support the proposition that the process by which the candidate genes were selected was effective and that an appreciable fraction of the chosen candidate genes are the actual quantitative trait genes. This must be qualified, however, as the distribution of the different classes of genomic elements among the QTLs was not random, and thus may be biasing the results.

### 3.7. Validation of QTL

A total of 26 QTLs were tested for significance of element x line tests. Of these, eight did not have any significant element x line tests, five had one significant test and 14 had 2–6 significant element x line tests. Thus, 18 QTLs can be considered as validated. Of the unconfirmed QTLs, QTLs 1, 3, and 24 were tested by lncRNAs only, QTLs 12 and 18 by miRNAs only, QTL 14 by functional mutations only, and QTL 25 by both miRNA and functional mutations. More markers need to be tested in these QTLs to decide if the lack of significance is a result of lack of informativity (Type 2 error), or if indeed these QTLs are false positives (Type I error). More detail on the association found within each individual QTL is presented in Supplementary File S1.

## 4. Discussion

QTL mapping in an FSIL F<sub>6</sub> population phenotyped for survival in the face of MDV challenge, identified 38 QTLs distributed over 19 autosomes. Use of such a resource allowed for the identification of QTLs at a higher resolution than have been mapped previously, thus allowing for easier identification of potential candidate genes. The mapped QTLs, along with genomic sequences of the F<sub>0</sub> founder individuals, and transcriptomic information from challenged and control birds, has allowed us to identify genes, miRNAs, lncRNAs, and potentially functional mutations located under these QTLs as candidates for association with progeny mortality from Marek's disease. Genetic association studies in multiple elite lines have confirmed the significant effects of most of these candidates on MD. Here we will discuss the potential role of some of the most significant candidates.

Many of the genes we have associated with MD response in this study have biological roles clearly relevant to the pathogenesis of MDV infection. One of the primary targets of the virus are B-cells and genes known to be associated with B-cells, include two of our candidates, *CD7* and *TLR4*. *CD7* is a surface antigen found on naive and memory T-cells, as well as defined NK cell progenitors, and provides survival signaling through its tyrosine kinase activity in association with T-cell activation and is involved in T-cell/B-cell interactions. The Toll-like receptor, *TLR4* is found on the surface of B-lymphocytes and macrophages and drives the induction of IL-1  $\beta$ , TNF $\alpha$ , and other pro-inflammatory cytokines implicated as being important in this study.

After initial infection and a period of latency, T-cells become infected. Genes related to T-cell signalling pathways include our MD-associated *ADAMTS5*, *CD7*, *HAVCR1*, *LAG3*, *RELT*, *TIMD4* and *TREML2*. *ADAMTS5* (ADAM Metalloproteinase with Thrombospondin Type 1 Motif 5) encodes a metalloproteinase that plays an important role in inflammation and cell migration. It also has a critical role in T-lymphocyte migration from draining lymph nodes following viral infection. *HAVCR1*, Hepatitis A Virus Cellular Receptor 1 (T-Cell Immunoglobulin Mucin Receptor 1), is a receptor for TIMD4. *HAVCR1* plays a critical role in regulating immune cell activity, particularly regarding the host response to viral infection, while *TIMD4* is a T-cell immunoglobulin involved in regulating T-cell proliferation and lymphotoxin signalling. *LAG3* (Lymphocyte-Activation Gene 3) belongs to the immunoglobulin superfamily and acts as an inhibitory receptor on activated T-cells. It negatively regulates the activation, proliferation, and effector function of both CD8<sup>+</sup> and CD4<sup>+</sup> T-cells as well as

mediating immune tolerance. RELT is a member of the TNF-receptor superfamily. It can activate the NF-kappaB pathway and selectively bind TNF receptor-associated factor 1 (TRAF1). This receptor acts via CD3 signalling to stimulate T-cell proliferation, suggesting its regulatory role in the immune response. TREML2 (Triggering Receptor Expressed on Myeloid Cells Like 2) is a cell surface receptor that may play a role in both the innate and adaptive immune responses. It interacts with CD276 on T-cells, enhancing T-cell activation.

Once infection of T-cells has occurred, the disease can then proceed to become oncogenic. Once again, we see that many of our MD associated genes have functions that have been implicated in cancer, including *BG1* which encodes an Ig-superfamily type I transmembrane receptor-like protein that contains an immuno-receptor tyrosine-based inhibition motif (ITIM). *BG1* has previously been documented as conferring MHC-associated resistance to MDV-induced lymphoma [57]. Other candidates include *BRINP1* (silenced in some bladder cancers), *CD7* (associated with leukaemia), *CSTA* (encodes a stefin that functions as a cysteine protease inhibitor, suggested as a prognostic tool for cancer), and *SUPT20H* (a known tumour rejection antigen). *FLT3* is another significant candidate, with mutations in this gene being common in acute myeloid leukaemia. *FLT3* is involved in activation of various pathways including apoptosis, and proliferation and differentiation of hematopoietic cells.

One of the main pathologies of MD is its effect on the nervous system, and so it is interesting to see that some of our MD associated genes are involved in the function/growth of neurons (*SCN4A* and *CTNNA2*). *SCN4A* (Sodium Voltage-Gated Channel  $\alpha$  Subunit 4) encodes one member of the sodium channel  $\alpha$  subunit gene family involved in generation and propagation of action potentials in neurons and muscle. *CTNNA2* (Catenin  $\alpha$  2) is thought to be involved in the regulation of cell–cell adhesion and differentiation in the nervous system. It is required for proper regulation of cortical neuronal migration and neurite growth.

The remaining MD associated genes are seen to have general roles as immune system genes: *C1S*, *TAP1* and *SOCS1*. *C1S* (Complement component 1S) encodes a serine protease component of the complement system which enhances the host antibody immune response. *TAP1* (Transporter 1, ATP Binding Cassette Subfamily B Member) is involved in the transport of antigens from the cytoplasm to the endoplasmic reticulum for association with MHC class I molecules. *SOCS1* (Suppressor of Cytokine Signalling 1) encodes a protein which functions downstream of cytokine receptors, and takes part in a negative feedback loop to attenuate cytokine signaling via the JAK/STAT3 pathway. All of these candidate genes had more than one test significant at  $p \leq 0.05$ .

We would hypothesize that many of our candidate genes are working in conjunction with each other to confer the resistance phenotype, whether that be through their proximity in the genome (e.g., *CSTA*, *LAG3* and *C1S* on Chr5 and *HAVCR1* and *TIMD4* on Chr13) or their interaction in biological networks (e.g., *SOCS1*, *TLR4* and *TREML2* or *RELT*, *LAG3*, *HAVCR1* and *TIMD4*). This would support the idea that some traits may be associated with pathway-level interactions as opposed to discrete gene functions [58,59].

Examination of these genes and their significance of association with MDV resistance across the elite lines indicates a few top candidates, namely: the cluster of genes in QTLR5 (*CSTA*, *C1S* and *LAG3*), *FLT3* in QTLR10, *CTNNA2* in QTLR19 and *TAP1* in QTLR32. It is also interesting to note the distribution of associated genes across the different lines of birds. Examination of the coding candidate genes shows that only two genes, *FLT3* and *CTNNA2*, are significant in all three varieties of birds. The genes *ADAMTS5* and *TAP1* only showed significant association in Plymouth Rock birds, whereas *CSTA*, *LAG3*, *C1S*, *SUPT20H* and the functional variant in *SCN4A* identified as significant candidates in White Leghorn and Rhode Island Red breeds. Finally, several genes are only significant in White Leghorns, namely *RELT*, *TRANK1*, *HAVCR1*, *TIMD4*, *SOCS1*, *TLR4*, *CD7*, *TREML2*, along with the functional variant in ENSGALG00000003188.

Genes identified in this analysis include many novel candidates for resistance as well as highlighting genes proposed in previous studies. For example *CD8B* (T-cell glycoprotein), *CTLA4* (immunoglobulin) and *CD72* (B-cell associated) are postulated as important lncRNA target genes by You et al. [24] and

are found under our QTLs (*CD8B*—QTL19), and differentially expressed in our transcriptomic work (*CTLA4* and *CD72*). Similarly, *ATF2* (involved in carcinogenesis is found in QTL25) was proposed as an important target for the miRNA gga-mir15b during MDV infection [25]. Also in QTL25 we find gga-mir-10b, previously seen to be upregulated in the spleen during MDV infection [27]. Other potentially interesting miRNA targets include *PBEF1* (pre-B-cell enhancing factor) and *FCHSD2* (involved in endocytosis) [26] that lie under QTL2 and 11, respectively. Further genes previously linked with MDV resistance include *GHI* (growth hormone) and *CD79B* (B-cell antigen), both of which lie under our QTL37.

One of the significant aspects of this research is that it utilized large, commercial production relevant lines, and the challenge virus is a very virulent ++ strain, frequently encountered by production birds in the field. In contrast, most previously published MDV resistance research utilizes specialized research lines, many of which are inbred, and selected for differential response to MDV. These studies utilized laboratory strains of the virus, for which commercial production birds now appear to be resistant. Furthermore, this study investigated MD resistance genes in three distinct breeds of chickens, namely White Leghorn, White Plymouth Rock, and Rhode Island Red, not just one laboratory line. Moreover, these MD association studies replicated the results from the FSIL study increasing our confidence in the causal nature of the QTL, and possibly the genes and variants in MDV resistance. The response to the virus was measured as mortality in a large progeny group (approximately 30 daughters) for over 9000 sires, using pre-existing information that had been developed within a commercially relevant production trait breeding program. This unique approach increases the relevance of the results to application into a commercial breeding program, while simultaneously provides information on the underlying mechanism of general viral resistance applicable to not only birds, but also other species. This information can provide insights into mechanisms for improving resistance or lead to the development of improved commercial vaccines.

## 5. Conclusions

Utilizing an FSIL F<sub>6</sub> population of birds phenotyped for response to Marek's disease virus infection has allowed us to map QTL for disease resistance at high-resolution. Combining this with expression data from challenged and control birds, we have identified candidate genes, miRNAs, lncRNAs and potentially functional mutations which have been validated in genetic association tests with MD mortality in diverse, commercially relevant, elite lines of poultry. This most comprehensive genetic study to date supplies us with variants in candidate genes that can now go on to be functionally tested for their utility in marker assisted selection, improved vaccine development, and potential future gene editing strategies.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4425/11/9/1019/s1>. Table S1: Primers. Figure S1: Manhattan plots of significant QTL. Table S2: Ensembl genes under QTL. Table S3: Variants under QTL. Table S4: Differentially expressed genes. Table S5: QTL genes also differentially expressed. Table S6: Gene ontology terms. Table S7: All markers used in association tests. Table S8: Association test results (*p*-values). Table S9: Summary of element x line tests. File S1: Details of association by individual QTL.

**Author Contributions:** Conceptualization, D.W.B. and J.E.F.; methodology, E.L., M.S., J.E.F. and J.S.; software, E.L., M.S. and J.S.; formal analysis, E.L., M.S. and J.S.; resources, J.E.F.; data curation, J.S.; writing—original draft preparation, J.S.; writing—review and editing, J.S., J.E.F., E.L., M.S. and D.W.B.; supervision, D.W.B.; project administration, D.W.B. and J.S.; funding acquisition, D.W.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Biotechnology and Biological Sciences Research Council (grant number BB/K006916/1).

**Acknowledgments:** The authors would like to thank Choon-Kiat Khoo, Richard Kuo and Lel Eory (Roslin) for invaluable bioinformatics assistance. Also requiring thanks are Ashlee R. Lund, Amy M. McCarron, Kara Pinegar-Maxon and Grant Liebe (Hy-Line) for their excellent technical assistance with genotyping. We are also extremely grateful to Edinburgh Genomics (Edinburgh) for carrying out the 600K SNP genotyping and the whole genome sequencing of the parental birds. We would also like to thank the GATC sequencing facility (Konstanz, Germany) for generating the transcriptomic data used in this study.

**Conflicts of Interest:** The authors declare no conflict of interest but would like to state that J.S. is a member of the Editorial Board of *Genes*.

**Ethics Approval:** All procedures carried out on the birds involved in this study were conducted in compliance with Hy-Line International Institutional Animal Care and Use Committee guidelines.

**Availability of Data:** Data have been submitted to the European Nucleotide Archive (ENA) at EMBL-EBI under study accession numbers PRJEB39142 (WGS) and PRJEB39361 (RNAseq).

## References

1. Morrow, C.; Fehler, F. Marek's disease: A worldwide problem. In *Marek's Disease, An Evolving Problem*; Davison, F., Nair, V., Eds.; Elsevier Ltd.: Oxford, UK, 2004; pp. 49–61.
2. Nair, V. Evolution of Marek's disease—A paradigm for incessant race between the pathogen and the host. *Vet. J.* **2005**, *170*, 175–183. [[CrossRef](#)] [[PubMed](#)]
3. Islam, A.F.; Wong, C.W.; Walkden-Brown, S.W.; Colditz, I.G.; Arzey, K.E.; Groves, P.J. Immunosuppressive effects of Marek's disease virus (MDV) and herpesvirus of turkeys (HVT) in broiler chickens and the protective effect of HVT vaccination against MDV challenge. *Avian Pathol.* **2002**, *31*, 449–461. [[CrossRef](#)] [[PubMed](#)]
4. Witter, R.L. Protective efficacy of Marek's disease vaccines. *Curr. Top. Microbiol. Immunol.* **2001**, *255*, 57–90. [[CrossRef](#)] [[PubMed](#)]
5. Read, A.F.; Baigent, S.J.; Powers, C.; Kgosana, L.B.; Blackwell, L.; Smith, L.P.; Kennedy, D.A.; Walkden-Brown, S.W.; Nair, V.K. Imperfect Vaccination Can Enhance the Transmission of Highly Virulent Pathogens. *PLoS Biol.* **2015**, *13*, e1002198. [[CrossRef](#)] [[PubMed](#)]
6. Boodhoo, N.; Gurung, A.; Sharif, S.; Behboudi, S. Marek's disease in chickens: A review with focus on immunology. *Vet. Res.* **2016**, *47*, 119. [[CrossRef](#)]
7. Fulton, J.E.; Arango, J.; Arthur, J.A.; Settar, P.; Kreager, K.S.; O'Sullivan, N.P. Improving the outcome of a Marek's disease challenge in multiple lines of egg type chickens. *Avian Dis.* **2013**, *57* (Suppl. S2), 519–522. [[CrossRef](#)]
8. Wolc, A.; Arango, J.; Jankowski, T.; Settar, P.; Fulton, J.E.; O'Sullivan, N.P.; Fernando, R.; Garrick, D.J.; Dekkers, J.C. Genome-wide association study for Marek's disease mortality in layer chickens. *Avian Dis.* **2013**, *57* (Suppl. S2), 395–400. [[CrossRef](#)]
9. Burgess, S.C.; Young, J.R.; Baaten, B.J.; Hunt, L.; Ross, L.N.; Parcels, M.S.; Kumar, P.M.; Tregaskes, C.A.; Lee, L.F.; Davison, T.F. Marek's disease is a natural model for lymphomas overexpressing Hodgkin's disease antigen (CD30). *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13879–13884. [[CrossRef](#)]
10. Li, K.; Lian, L.; Yang, N.; Qu, L. Temporal expression and DNA hypomethylation profile of CD30 in Marek's disease virus-infected chicken spleens. *Poult. Sci.* **2015**, *94*, 1165–1169. [[CrossRef](#)]
11. Bacon, L.D.; Witter, R.L.; Crittenden, L.B.; Fadly, A.; Motta, J. B-haplotype influence on Marek's disease, Rous sarcoma, and lymphoid leukosis virus-induced tumors in chickens. *Poult. Sci.* **1981**, *60*, 1132–1139. [[CrossRef](#)]
12. Kaufman, J. The simple chicken major histocompatibility complex: Life and death in the face of pathogens and vaccines. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2000**, *355*, 1077–1084. [[CrossRef](#)] [[PubMed](#)]
13. Koch, M.; Camp, S.; Collen, T.; Avila, D.; Salomonsen, J.; Wallny, H.-J.; van Hateren, A.; Hunt, L.; Jacob, J.P.; Johnston, F.; et al. Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity* **2007**, *27*, 885–899. [[CrossRef](#)] [[PubMed](#)]
14. Hunt, H.D.; Jadhao, S.; Swayne, D.E. Major histocompatibility complex and background genes in chickens influence susceptibility to high pathogenicity avian influenza virus. *Avian Dis.* **2010**, *54* (Suppl. S1), 572–575. [[CrossRef](#)] [[PubMed](#)]
15. Liu, H.C.; Kung, H.J.; Fulton, J.E.; Morgan, R.W.; Cheng, H.H. Growth hormone interacts with the Marek's disease virus SORF2 protein and is associated with disease resistance in chicken. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9203–9208. [[CrossRef](#)] [[PubMed](#)]
16. Liu, H.C.; Cheng, H.H.; Tirunagaru, V.; Sofer, L.; Burnside, J. A strategy to identify positional candidate genes conferring Marek's disease resistance by integrating DNA microarrays and genetic mapping. *Anim. Genet.* **2001**, *32*, 351–359. [[CrossRef](#)] [[PubMed](#)]

17. Liu, H.C.; Niikura, M.; Fulton, J.E.; Cheng, H.H. Identification of chicken lymphocyte antigen 6 complex, locus E (LY6E, alias SCA2) as a putative Marek's disease resistance gene via a virus-host protein interaction screen. *Cytogenet. Genome Res.* **2003**, *102*, 304–308. [[CrossRef](#)]
18. Smith, J.; Sadeyen, J.R.; Paton, I.R.; Hocking, P.M.; Salmon, N.; Fife, M.; Nair, V.; Burt, D.W.; Kaiser, P. Systems analysis of immune responses in Marek's disease virus-infected chickens identifies a gene involved in susceptibility and highlights a possible novel pathogenicity mechanism. *J. Virol.* **2011**, *85*, 11146–11158. [[CrossRef](#)]
19. Meydan, H.; Yildiz, M.A.; Dodgson, J.B.; Cheng, H.H. Allele-specific expression analysis reveals CD79B has a cis-acting regulatory element that responds to Marek's disease virus infection in chickens. *Poult. Sci.* **2011**, *90*, 1206–1211. [[CrossRef](#)]
20. Li, D.F.; Lian, L.; Qu, L.J.; Chen, Y.M.; Liu, W.B.; Chen, S.R.; Zheng, J.X.; Xu, G.Y.; Yang, N. A genome-wide SNP scan reveals two loci associated with the chicken resistance to Marek's disease. *Anim. Genet.* **2013**, *44*, 217–222. [[CrossRef](#)]
21. Wang, D.; Sun, S.; Heidari, M. Marek's disease vaccine activates chicken macrophages. *J. Vet. Sci.* **2018**, *19*, 375–383. [[CrossRef](#)]
22. Chakraborty, P.; Kuo, R.; Vervelde, L.; Dutia, B.M.; Kaiser, P.; Smith, J. Macrophages from Susceptible and Resistant Chicken Lines have Different Transcriptomes following Marek's Disease Virus Infection. *Genes* **2019**, *10*, 74. [[CrossRef](#)] [[PubMed](#)]
23. Han, B.; He, Y.; Zhang, L.; Ding, Y.; Lian, L.; Zhao, C.; Song, J.; Yang, N. Long intergenic non-coding RNA GALMD3 in chicken Marek's disease. *Sci. Rep.* **2017**, *7*, 10294. [[CrossRef](#)] [[PubMed](#)]
24. You, Z.; Zhang, Q.; Liu, C.; Song, J.; Yang, N.; Lian, L. Integrated analysis of lncRNA and mRNA repertoires in Marek's disease infected spleens identifies genes relevant to resistance. *BMC Genom.* **2019**, *20*, 245. [[CrossRef](#)]
25. Tian, F.; Luo, J.; Zhang, H.; Chang, S.; Song, J. MiRNA expression signatures induced by Marek's disease virus infection in chickens. *Genomics* **2012**, *99*, 152–159. [[CrossRef](#)] [[PubMed](#)]
26. Parnas, O.; Corcoran, D.L.; Cullen, B.R. Analysis of the mRNA targetome of microRNAs expressed by Marek's disease virus. *MBio* **2014**, *5*, e01060-13. [[CrossRef](#)] [[PubMed](#)]
27. Hicks, J.A.; Liu, H.C. Impact of HVT Vaccination on Splenic miRNA Expression in Marek's Disease Virus Infections. *Genes* **2019**, *10*, 115. [[CrossRef](#)]
28. Neerukonda, S.N.; Tavlarides-Hontz, P.; McCarthy, F.; Pendarvis, K.; Parcels, M.S. Comparison of the Transcriptomes and Proteomes of Serum Exosomes from Marek's Disease Virus-Vaccinated and Protected and Lymphoma-Bearing Chickens. *Genes* **2019**, *10*, 116. [[CrossRef](#)]
29. Luo, J.; Mitra, A.; Tian, F.; Chan, S.; Zhang, H.; Cui, K.; Yu, Y.; Zhao, K.; Song, J. Histone methylation analysis and pathway predictions in chickens after MDV infection. *PLoS ONE* **2012**, *7*, e41849. [[CrossRef](#)]
30. Mitra, A.; Luo, J.; He, Y.; Gu, Y.; Zhang, H.; Zhao, K.; Cui, K.; Song, J. Histone modifications induced by MDV infection at early cytolytic and latency phases. *BMC Genom.* **2015**, *16*, 311. [[CrossRef](#)]
31. Heifetz, E.M.; Fulton, J.E.; O'Sullivan, N.P.; Arthur, J.; Cheng, H.; Wang, J.; Soller, M.; Dekkers, J.C.M. Mapping QTL affecting resistance to Marek's disease in an F6 advanced intercross population of commercial layer chickens. *BMC Genom.* **2009**, *10*, 20. [[CrossRef](#)]
32. Kranis, A.; Gheyas, A.A.; Boschiero, C.; Turner, F.; Yu, L.; Smith, S.; Talbot, R.; Pirani, A.; Brew, F.; Kaiser, P.; et al. Development of a high density 600K SNP genotyping array for chicken. *BMC Genom.* **2013**, *14*, 59. [[CrossRef](#)]
33. McElroy, J.P.; Dekkers, J.C.; Fulton, J.E.; O'Sullivan, N.P.; Soller, M.; Lipkin, E.; Zhang, W.; Koehler, K.J.; Lamont, S.J.; Cheng, H.H. Microsatellite markers associated with resistance to Marek's disease in commercial layer chickens. *Poult. Sci.* **2005**, *84*, 1678–1688. [[CrossRef](#)] [[PubMed](#)]
34. Heifetz, E.M.; Fulton, J.E.; O'Sullivan, N.P.; Arthur, J.A.; Wang, J.; Dekkers, J.C.; Soller, M. Mapping quantitative trait loci affecting susceptibility to Marek's disease virus in a backcross population of layer chickens. *Genetics* **2007**, *177*, 2417–2431. [[CrossRef](#)] [[PubMed](#)]
35. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (accessed on 1 October 2013).
36. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]

37. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. 1000 Genome Project Data Processing Subgroup. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
38. Cingolani, P.; Platts, A.; le Wang, 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: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)]
39. Mosig, M.O.; Lipkin, E.; Khutoreskaya, G.; Tchourzyna, E.; Soller, M.; Friedmann, A. A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling in a daughter design, using an adjusted false discovery rate criterion. *Genetics* **2001**, *157*, 1683–1698.
40. Fernando, R.L.; Nettleton, D.; Southey, B.R.; Dekkers, J.C.; Rothschild, M.F.; Soller, M. Controlling the proportion of false positives in multiple dependent tests. *Genetics* **2004**, *166*, 611–619. [[CrossRef](#)] [[PubMed](#)]
41. Lipkin, E.; Strillacci, M.G.; Eitam, H.; Yishay, M.; Schiavini, F.; Soller, M.; Bagnato, A.; Shabtay, A. The Use of Kosher Phenotyping for Mapping QTL Affecting Susceptibility to Bovine Respiratory Disease. *PLoS ONE* **2016**, *11*, e0153423. [[CrossRef](#)]
42. Weller, J.I. *Quantitative Trait Loci Analysis in Animals*, 2nd ed.; CABI Publishing: London, UK, 2009.
43. Kuo, R.I.; Tseng, E.; Eory, L.; Paton, I.R.; Archibald, A.L.; Burt, D.W. Normalized long read RNA sequencing in chicken reveals transcriptome complexity similar to human. *BMC Genom.* **2017**, *18*, 323. [[CrossRef](#)]
44. Biggs, P.M.; Payne, L.N. Studies on Marek's Disease. I. Experimental Transmission. *JNCI J. Natl. Cancer Inst.* **1967**, *39*, 267–280. [[CrossRef](#)]
45. Khare, M.L.; Grun, J.; Adams, E.V. Marek's disease in Japanese quail—A pathological, virological and serological study. *Poult. Sci.* **1975**, *54*, 2066–2068. [[CrossRef](#)] [[PubMed](#)]
46. Bettridge, J.M.; Lynch, S.E.; Brena, M.C.; Melese, K.; Dessie, T.; Terfa, Z.G.; Desta, T.T.; Rushton, S.; Hanotte, O.; Kaiser, P.; et al. Infection-interactions in Ethiopian village chickens. *Prev. Vet. Med.* **2014**, *117*, 358–366. [[CrossRef](#)] [[PubMed](#)]
47. Kim, D.; Pertea, G.; Trapnell, C.; Pimentel, H.; Kelley, R.; Salzberg, S.L. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **2013**, *14*, R36. [[CrossRef](#)] [[PubMed](#)]
48. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
49. Trapnell, C.; Hendrickson, D.G.; Sauvageau, M.; Goff, L.; Rinn, J.L.; Pachter, L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* **2013**, *31*, 46–53. [[CrossRef](#)]
50. Perumbakkam, S.; Muir, W.M.; Black-Pyrkosz, A.; Okimoto, R.; Cheng, H.H. Comparison and contrast of genes and biological pathways responding to Marek's disease virus infection using allele-specific expression and differential expression in broiler and layer chickens. *BMC Genom.* **2013**, *14*, 64. [[CrossRef](#)]
51. Hindorff, L.A.; Sethupathy, P.; Junkins, H.A.; Ramos, E.M.; Mehta, J.P.; Collins, F.S.; Manolio, T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9362–9367. [[CrossRef](#)]
52. Semagn, K.; Beyene, Y.; Warburton, M.L.; Tarekegne, A.; Mugo, S.; Meisel, B.; Sehabiague, P.; Prasanna, B.M. Meta-analyses of QTL for grain yield and anthesis silking interval in 18 maize populations evaluated under water-stressed and well-watered environments. *BMC Genom.* **2013**, *14*, 313. [[CrossRef](#)]
53. Fulton, J.E.; McCarron, A.M.; Lund, A.R.; Pinegar, K.N.; Wolc, A.; Chazara, O.; Bed'Hom, B.; Berres, M.; Miller, M.M. A high-density SNP panel reveals extensive diversity, frequent recombination and multiple recombination hotspots within the chicken major histocompatibility complex B region between BG2 and CD1A1. *Genet. Sel. Evol.* **2016**, *48*, 1. [[CrossRef](#)]
54. Vallejo, R.L.; Bacon, L.D.; Liu, H.C.; Witter, R.L.; Groenen, M.A.; Hillel, J.; Cheng, H.H. Genetic mapping of quantitative trait loci affecting susceptibility to Marek's disease virus induced tumors in F2 intercross chickens. *Genetics* **1998**, *148*, 349–360. [[PubMed](#)]
55. Yonash, N.; Bacon, L.D.; Witter, R.L.; Cheng, H.H. High resolution mapping and identification of new quantitative trait loci (QTL) affecting susceptibility to Marek's disease. *Anim. Genet.* **1999**, *30*, 126–135. [[CrossRef](#)] [[PubMed](#)]

56. Cheng, H.; Niikura, M.; Kim, T.; Mao, W.; MacLea, K.S.; Hunt, H.; Dodgson, J.; Burnside, J.; Morgan, R.; Ouyang, M.; et al. Using integrative genomics to elucidate genetic resistance to Marek's disease in chickens. *Dev. Biol.* **2008**, *132*, 365–372. [[CrossRef](#)]
57. Goto, R.M.; Wang, Y.; Taylor, R.L., Jr.; Wakenell, P.S.; Hosomichi, K.; Shiina, T.; Blackmore, C.S.; Briles, W.E.; Miller, M.M. BG1 has a major role in MHC-linked resistance to malignant lymphoma in the chicken. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16740–16745. [[CrossRef](#)] [[PubMed](#)]
58. Du, Q.; Tian, J.; Yang, X.; Pan, W.; Xu, B.; Li, B.; Ingvarsson, P.K.; Zhang, D. Identification of additive, dominant, and epistatic variation conferred by key genes in cellulose biosynthesis pathway in *Populus tomentosa*. *DNA Res.* **2015**, *22*, 53–67. [[CrossRef](#)] [[PubMed](#)]
59. Pei, G.; Sun, H.; Dai, Y.; Liu, X.; Zhao, Z. Investigation of multi-trait associations using pathway-based analysis of GWAS summary statistics. *BMC Genom.* **2019**, *20*, 79. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).