Discovery of Somatic Driver Variations of MDV-induced Transformations and Genetic Elements Demonstrating Resistance to Marek's Disease via Cytogenetic Analysis, Genome-Wide Selection, and Next-Generation Sequencing

Alec Steep
Submitted: July 28th, 2014

Committee:
Eran Andrechek
Titus Brown
Hans Cheng
Jerry Dodgson
Michele Fluck

Abstract

Marek's Disease Virus (MDV) is a highly oncogenic alphaherpes virus that induces Marek's Disease (MD) in chickens. MDV is characterized most notably by T-cell lymphomas. Consequently, primary infectious disease control strategies for MD focus on vaccination against MDV-induced tumorigenesis and optimization of genetic resistance to MD. However, selective pressures—such as high-density poultry rearing practices, vaccination control, and incorporated genetic resistance to MD—have induced MDV evolution and increased MDV virulence. Vaccination strategies are particularly susceptible to MDV evolution as shown by multiple vaccine breaks throughout the latter half of the twentieth century. The last vaccine break occurred as recently as the late 1990s, and in the wake of this break, the poultry industry has come to rely on only a handful of vaccines.

Additional measures of control are needed to ensure MD containment and prevent future vaccine breaks. Future research must characterize MDV-induced driver mutations. Current vaccines suppress T-cell lymphomas, but more information is needed on the specific somatic alterations that contribute to transformation. This proposal will examine driver mutations and gene resistance by using cytogenetic analysis, SNP arrays, and Next-generation sequencing of DNA and RNA. This methodological combination has yet to be applied to MDV and is expected to reveal the requisite somatic alterations that drive transformation and tumorigenesis. These results will likely be used to guide future infectious disease control strategies.

Table of Contents

Research Plan

Specific Aims

This project intends to identify the somatic variations responsible for MDV-transformation events, specifically driver mutations, and the alleles, genes, and pathways involved in genetic resistance to Marek's Disease.

Objective 1: Characterize large-scale MDV-induced genomic alterations using single nucleotide polymorphism (SNP) arrays and cytogenetics.

Genomic material from 72 MDV-induced tumor samples will undergo cytognetic analysis and SNP microarray analysis. Cytogenetics will reveal MDV-integration profiles, tumor clonality, and large-scale alterations. Tumor DNA will undergo microarray analysis on a custom 15K Affymetrics SNP array to reveal select SNPs, large-scale alterations (e.g. CNVs and LOH), tumor clonality, and alleles involved in MD resistance.

Objective 2: Characterize MDV-induced somatic alterations responsible for transformation and tumorigenesis.

Illumina sequenced DNA from 22 MDV-induced tumor samples will reveal somatic alterations driving transformation including single nucleotide variants, insertions and deletions, copy number variations, structural variants, and gene fusions. Variants considered driver mutations may reveal novel targets for MD control strategies (e.g. vaccine development and incorporation of genetic resistance).

Objective 3: Characterize MDV-induced gene expression alterations responsible for transformation and tumorigenesis.

Illumina sequenced RNA from 22 MDV-induced tumor samples will reveal alterations in gene expression and transcription products to reveal the genes, alleles, and pathways altered in transformation and confirm suspected genomic candidate alterations.

Background and Significance

Global consumer demand for poultry products, especially meat, is at its highest. Poultry is the leading meat consumed in the United States and around the world. It is considered one of the most dynamic animal commodities where production has increased by 436% since 1970(1). The United States has the largest poultry industry in the world and is the largest poultry meat producer, exporter, and consumer. The United States Department of Agriculture (USDA) expects these trends to continue into 2019 and beyond (2).

This demand has transformed poultry cultivation. Poultry raising practices have changed dramatically since the early 1950s, when the average chicken was raised in relatively small backyard flocks, into the high-density chicken rearing industry of today, which relies on certain breeds for genetic improvement (1). Up to 90% of the rise in chicken rearing productivity is attributable to advanced poultry breeding programs and the selection of certain production traits (e.g. growth-enhancement and disease resistance)(2,3).

Despite these advances, the modern landscape of the poultry industry faces new challenges. The industry has consolidated and the common practice of high-density chicken rearing has reduced genetic diversity (1). Therefore, the fast-growing industry remains focused on controlling infectious disease to sustain economic viability, enhance animal welfare, and maintain consumer confidence in poultry products (2). One especially serious threat to the industry is Marek's Disease. MD is a lymphoproliferative, infectious disease that is caused by the highly oncogenic alpha-herpes Marek's Disease Virus. Pathogenic strains of MDV can be found in almost all countries with a developed poultry industry. MDV control is difficult because of its capricious nature to spontaneously overcome vaccines and become increasingly virulent. Annual world-wide losses associated with MD equate to roughly 1-2 billion US dollars, which is likely an underestimate given chronic under-reporting (4). MD remains one of the top infectious poultry diseases of concern in the United States and worldwide.

Research into Marek's Disease began in in 1907 when József Marek, a Hungarian veterinarian at the Budapest Veterinary School, characterized a generalized polyneuritis in four roosters (5). This disease was later discovered to show neoplastic attributes (6,7). Furthermore, its pathology closely resembled the retroviral-induced lymphatic leucosis. It was not until the discovery of MD's herpes virus etiology in the late 1960s that the polyneuritis discovered by Marek was distinguishable from lymphatic leucosis (8). This discovery ultimately led to the development of a vaccine against MDV. The vaccine did not stop the proliferation of MDV, however, but stopped the formation of neoplasms in infected chickens. This development made it the first immune prophylaxis to be used against a cancer, otherwise known as the first cancer vaccine (9).

As discussed, MD is extremely virulent and thus the severity of MD has changed drastically since its discovery. Originally described as a sporadic chronic disease, MD has evolved into several forms of aggressive, acute and peracute diseases. In addition to the chronic polyneuritis and visceral lymphomas observed in 1907, MDV has evolved to cause a combination of immunosuppressive, inflammatory, and lymphoproliferative lesions in chickens (Figure 1) (5,10). Clinical signs include paralysis, skin leucosis, dermatitis, cachexia, neurological deficits (e.g. ataxia and torticollis), inflammatory brain and nerve lesions, bursal and thymic atrophy, gross visceral lymphomas, stunting, and death (10). The degree of clinical

severity is dependent upon the virulence of the virus pathotype; ranging from mildly virulent (m) to virulent (v), very virulent (vv) and very virulent plus (vv+). The classification of virulence was established based on the response of the MDV pathotype to commercially available vaccines and combinations thereof, and it seems there is a close relationship between virulence and vaccine breaks (Figure 2) (5,11–13). The ever-increasing virulence of MDV is attributable to the use of vaccines in a high-density rearing environment. Modern poultry housing units are densely packed with chickens. The ebb and flow of poultry consistently introduces naïve hosts, providing MDV with a continuous reservoir, which has a reduced genetic diversity. These practices have created an environment ideal for aggressive selection and adaptation of MDV, creating the need for additional control strategies (4). In response to this need, we aim to reveal the genetic etiology of tumorigenesis, one of the contributing factors to MDV-induced fatality. A genetic approach to tumor prevention may offer an alternative to the vicious cycle of vaccine resistance in treating Marek's Disease.

The pathogenesis of MDV-induced tumors begins when infectious dust contaminated with MDV, which is shed from other infected birds, is inhaled (Figure 3). The virus is transported from the lungs to the lymphoid tissue of the spleen, thymus, and bursa of Fabricius by macrophages. It is in these visceral organs where MDV targets B lymphocytes for cytolytic infection. From there, the virus most commonly spreads to T-lymphocytes (14), but may also target macrophages (15).

During the initial event when the extracellular enveloped virus seeks to enter its first target cell, it likely binds to cellular receptors via glycoprotein B (gB) in combination with other glycoproteins (16). Certain cellular receptor molecules also likely play a role in infection (e.g. heparin sulfate) (17). When the virus seeks to leave an infected cell and enter an uninfected cell, it requires direct contact between both cells (16). Evidence suggests that an intercellular bridge forms through which the virus passes to the uninfected cell (18). Glycoproteins E, I, and M (gE, gI, and gM) also play a necessary role in infected- to uninfected-cell transfer (19,20).

Near the end of the cytolytic replication phase, infected lymphocytes carry MDV to the feather follicle epithelium (FFE), where productive infection occurs. Infected birds shed the extracellular virus from the FFE for an average of 10 days post infection (14).

In total, there are three stages of MDV-cell interactions: productive, latent, and transforming. During productive infection, also known as cytolytic infection, the vast majority of viral DNA is replicated, proteins are synthesized, and virus particles are produced. An intranuclear inclusion body forms, resulting in subsequent lysis of the cell (14). The viral host shut off protein, UL41, has been shown to likely be responsible for the initiation of the lytic process (21). UL41 degrades host mRNA, shuts off host protein synthesis, and regulates viral gene expression (22,23). Near the end of the cytolytic replication phase, infected lymphocytes carry MDV to the feather follicle epithelium (FFE), where shedding of the extracellular virus begins 10 days post infection on average (14,24).

After the cytolytic stage, MDV enters latency in infected lymphocytes and integrates into the telomeric regions of chromosomes in CD4+ T cells, and sometimes CD8+ T cells and B cells (25–27). The reactivation of early and late genes in latently infected cells is characterized by the upregulation of transcripts of ICP4, a protein that acts as a transactivator of genes associated

with lytic infection, and downregulation of latency-associated transcripts (LATs), resulting in a return to the lytic stage of infection (28,29).

In susceptible and/or unvaccinated birds, infected T cells, primarily CD4+ cells but rarely CD8+ cells, can undergo a transformation event resulting in subsequent lymphomagenesis in various visceral organs. Cells having undergone MDV-induced transformation commonly exhibit MD tumor-associated surface antigens (MATSAs), which are not present on cells undergoing MDV cytolytic replication (30,31). The only successfully cloned and identified MATSA is CD30, which is specifically associated with transformed CD4+ T cells (32,33). High expression of CD30 is characteristic of MD lymphomas; therefore, CD30 may be a component of a critical intracellular signaling pathway in transformed cells (34).

MDV induced transformation events are dependent on the establishment of the latent state, which is induced through a number of mechanisms but none more crucial than the integration of MDV into host chromosomes. This has been corroborated by lymphoma formation efficiency, which has been shown to directly correlate with the number of cells with latently infected genomes. Furthermore, one of these latently infected cells inevitably out-competes other cells typically resulting in tumors of monoclonal origin (35,36).

Integration is an important tactic for herpes viruses in animals; it ensures viral genome maintenance and replication during cell division (37,38), evasion from the immune system (27), and horizontal infection between hosts in a population (39). MDV is no exception, it is found exclusively in the integrated state during latency primarily in CD4+ T cells, although CD4-, CD8-, CD8+T cells, and B cells can also harbor latent MDV (27,35,37,40–43). MDV integration is preferential and likely exclusive to telomeric regions of host chromosomes (36,41,44); this is partially explained by the telomeric repeats (TMRs) found at either distal region of the linear MDV genome. Viral TMRs are homologous to host telomere sequences (TTAGGG)n and allow for targeted MDV integration (45,46). During latency, few viral gene products are produced: estimated between about 10-30 genes (47–49). Arguably the most important MDV latently expressed protein is Meq. Meq represses expression of lytic viral gene products, an essential scenario for latency induction.

Through mechanisms not entirely understood, MDV integration into host chromosomes induces somatic alterations in the chicken genome. It is widely believed that viral oncogenes in alliance with somatic driver mutations contribute to MDV-induce lymphomagenesis. The exact scenario of lymphomagenesis is in development, and we aim to confirm and expand upon the current model (Figure 4) (10).

The onset of transformation and tumorigenesis seems to be fuelled by genes carried on the MDV genome that, through aggressive expression, interact with host cell genomic material and machinery. The most powerful known MDV-encoded oncogene is *Meq*, *M*arek's *E*coRI-*Q*-encoded protein (50). In very virulent and very virulent plus pathotypes of MDV, Meq is a basic leucine zipper (bZIP) protein encoded as a 339-amino-acid unspliced open reading frame (51); our discussion will focus on this form of Meq. If *Meq* is deleted or mutated to substantially perturb function, tumorigenesis does not ensue; therefore, *Meq* is necessary for MDV transformation (52,53). However, Meq is not sufficient for MDV transformation, as Meq is encoded and expressed on attenuated, nononcogeneic MDV strains (54). Our discussion of

MDV transformation will focus primarily on Meq, but will also include other viral and host cell gene products that contribute to lymphamogenesis.

Because Meq is a bZIP protein, it can dimerize with other bZIP proteins or itself to form hetero- or homodimers, both of which are essential for MDV transformation (Figure 5) (52,55,56). Meq can heterodimerize with c-Jun, CREB, ATF-1, ATF-2, ATF-3, Fra-2, Jun-B, Jun-D, and NFIL3 (57,58). The Meq-c-Jun heterodimer has been the most extensively characterized as it seems to demonstrate the most oncogenic behavior of dimers. Meq stabilizes c-Jun, a known cellular proto-oncogene, and increases the efficacy of c-Jun signaling (57) resulting in enhanced cellular proliferation, suppression of apoptosis, and increased mobility of latently infected CD4+ T cells (59,60). Meq can also bind to other cellular proteins outside of its leucine zipper domain including: CDK-2, p53, Rb (50), HSP-70 (61), CtBP-1 (52), and SKP-2 (62). The interaction most thoroughly characterized is that of Meq and CtBP-1, which is essential for MDV transformation (52). CtBP-1 is a repression complex protein that represses transcription via chromatin remodeling (63). It is likely that Meq recruits CtBP-1 to specific loci during latency and transformation to influence repression of apoptosis (57) and enhancement of differentiation (60,64) in latent CD4+ T cells.

Meq and Meq dimers have been shown to transcriptionally regulate genes in several pathways critical for oncogenesis and/or involved in apoptotic repression including: the extracellular signal-regulated kinase/mitogen-activated protein kinase (ERK/MAPK), Jak-STAT, and ErbB pathways (65). Unsurprisingly, Meq has also been shown to upregulate the promoter of CD30, one of the Marek's tumor-associated surface antigens (MATSAs) (34). In addition to its full-length protein, Meq can be spliced resulting in a gene fusion between the first 100 amino acids of Meq and either the exon 2 or 3 of viral interleukin 8 (vIL8): Meq/vIL8 and Meq/vIL8Δexon3 (10,66,67). These splice variants can induce proliferation of fibroblasts and macrophages, promote transformation of CD4+ T cells through interactions with CtBP-1, and repress gene promoters in the latent MDV genome involved in cell lysis (10).

The above interactions revolve around the activity of Meq; however there are other gene products, transcripts, and microRNAs suspected to contribute to oncogenesis. These include but are not limited to: viral telomerase RNA (vTR), cellular TR for telomerase reverse transcriptase (cTERT) (68,69), viral interleukin 8 (vIL8) (70–73), *UL36* encoding ubiquitin-specific protease domain on the viral major tegument protein (MTP) (74), and MDV microRNAs (75). In our endeavor to reveal the genomic players involved in oncogenesis, we will watch for the aforementioned contributors of transformation and hopefully add more comprehension to the current model.

Preliminary Studies/Progress Report

We plan to characterize the MDV-induced genome integration events and somatic genetic alterations responsible for transformation and subsequent tumorigenesis. Much has been done to lay the groundwork for such ambitious goals. The Avian Disease and Oncology Laboratory (ADOL) and other collaborators have sought to genomically characterize MD tumor profiles by examining the alleles, genes, and pathways affected by MDV-induced somatic alterations; and the quantitative trait loci associated with MD resistance. An ongoing

collaboration between the Cheng and Delany labs has shed light on MDV integration and how it affects tumor lineages.

In the early 1990's Delecluse and Hammerschmidt discovered that it was possible for MDV to integrate into the chicken genome (35). However, more work was needed to better understand the context of MDV integration; the Cheng, Delany, and Hunt labs proposed to fill this void. They investigated integration events in late stage tumors within birds by using multi-color fluorescence in situ hybridization (FISH) with probes specific for either MDV or host chromosomes (27). The study concluded several major findings. First, MDV integrates into host chromosome telomeres (Figure 6); this observation was independently confirmed (76). It was originally hypothesized that MDV would specifically integrate into mega-telomeres, and while this was observed, integration events were also found in a number of average-sized telomeres in different chromosomes. Furthermore, it appears that MDV may preferentially integrate into certain chromosomes (i.e., GGA 4, 6, 9, 12, and 20). Figure 6 illustrates MDV integration preferences. One interesting candidate chromosome for future study is GGA 9, in which both homologs demonstrated MDV integration ability. This is especially appealing because GGA 9 encodes a single copy of the chicken telomerase RNA (cTR) component of telomerase, and integrated MDV encodes an additional two copies of TR (vTR). Viral TR has a high sequence identity (85-88%) to cTR (68,77,78). Viral integration sites have proven useful markers in determining tumor clonality (79); that is, do multiple tumors in a bird derive from a single transformation event (monoclonal origin) or multiple independent transformation events (polyclonal origin)? Robinson and colleagues discovered that tumors within birds frequently develop from a monoclonal origin and are related (Figure 7) (27).

To determine when MDV integrates during infection and in which cell types, the Cheng and Delany labs switched their focus from late stage tumor samples to cells targeted for MDV infection at time points representative of the lytic cycle (1-7 dpi), the establishment of latency (7-14 dpi), and the transformation period (14-21 dpi) (5,16,80). Early in infection, even in the first week, Robinson saw that both B and T cells could harbor integration events, showing that MDV integrates as part of its natural life cycle (80). In congruence with earlier work, MDV was shown to integrate only in telomeres and never in non-telomeric chromosomal DNA. It was also found that early in infection MDV integration was witnessed in a wider array of chromosomes than in more senior post transformation tumor cells. In addition, newly infected tissues did not demonstrate the clonal profiles seen in tumors, but rather highly variable profiles. These observations confirmed and expanded our understanding of MDV integration. We learned that integration occurs at high frequency and is preferential to telomeres on favored chromosomes. In addition, MDV can begin integration much earlier than previously thought. And finally, although integration is a hallmark of MDV-induced cancer, integration alone is not sufficient for MDV transformation (80).

More work is needed to understand the essential role integration plays in tumorigenesis, but we suspect that early integration serves as a mechanism for latency and that subsequent additional somatic mutations are essential for transformation and tumorigenesis. We have reason to suspect that integration is vital for tumorigenesis as deletion or mutation of MDV telomeric repeats (vTTAGGG), a necessary genetic element for integration, substantially decreases the incidence of tumors, but does not significantly weaken viral replication during the

lytic cycle (27,76). We will employ advanced genomic technologies in cooperation with previous collaborative efforts to investigate how structural alterations and somatic variations may arise from specific integration profiles, and how these influence the progression of tumor cell lineages.

To understand how the chicken genetically interacts with MDV in the narrative of MD we must characterize quantitative trait loci (QTLs) and regions of interest associated with resistance to MD and/or involved with tumorigenesis. Multiple approaches have been utilized in an effort to increase confidence in candidate genes, alleles, and pathways. Thus far, this approach has identified many strong candidates, specifically three MD resistance genes: *GH1*—encodes growth hormone 1 protein (81), *LY6E* aka *SCA2*—encodes lymphocyte antigen 6E protein (82,83), and *CD74*—encodes HLA class II histocompatibility antigen gamma chain protein (84). These results should grant increased accuracy and guide future, more precise analyses with advanced technology to understand the molecular mechanisms associated with tumor formation and their relationship to MD resistance.

These combined efforts include genome-wide select SNP arrays and QTL scans, selective sweeps for MD genetic resistance, transcript profiling, allele-specific expression (ASE) screens, and chromatin immunoprecipitation (ChIP) followed by NGS. All efforts use the same genetic background, our experimental models: White Leghorn chickens lines 6 and 7 and their F1 progenies; lines 6x7 and 7x6. Lines 6 and 7 are MD resistant and susceptible, respectively, and are over 99% inbred (85,86). Resulting F1 progeny are heterozygous for resistant and susceptible alleles at specifically considered loci, many of which include candidate genes. A designed exception to complete genetic heterozygousity is at the genetic origin of the major histocompatibility complex (MHC), which is a locus shown to greatly influence MD incidence (87). Therefore, the remaining loci may be surveyed for their collective contribution to MD incidence.

In a pilot study, the Avian Disease and Oncology Laboratory attempted to gain a global view of tumor sample genomes. Two Affymetrix SNP arrays, the chicken genome wide 600K array and the custom 15K, were used to genotype 10 tumor samples from six line 6x7 F1 progeny (4 birds had paired tumors). In addition, 12 uninfected controls were genotyped: three from line 6 parents, three from line 7 parents, and six from 6x7 F1 progeny. The 15K array and the 600K array have the capacity to survey 8,896 SNPs and 145,985 SNPs, respectively. The vast majority of these SNPs are completely fixed in and differ between parental lines and are therefore heterozygous in F1 progeny, by design. SNPs that deviate from the expected specific heterozygousity in F1 progeny tumor samples should expose a somatic alteration, likely a large loss-of-heterozygousity (LOH) variation.

These SNPs are spaced throughout the genome and therefore allow us to draw conclusions from a global genome perspective:

- There was consistent and limited somatic alterations between samples. To determine if these loci involve LOH and/or copy number variations (CNVs) with the same array technology, we will utilize the methods described in Laurie et al. (2012) and Jacobs et al. (2012) (88,89).
- The average size of large somatic alterations is 100-350 Kb based on SNP distributions.

- Six of eight paired tumors showed genotypes suggestive of monoclonal origin, consistent with the results from cytogenetic analysis from the Delany lab (27).
- Comparisons between paired tumors should help filter out passenger SNPs and technical errors. Shared SNPs between tumor-pairs will help hone in on the region of interest. We may then compare independent transformation events for driver mutations in tumors of different clonality.
- The 15K array demonstrated power comparable to the 600K array in detecting regions with possible LOH and tumor origin; this can be attributed to large blocks of linkage disequilibrium.
- Arrays did not survey the male sex chromosomes (Z chromosomes) and future studies should include the Z chromosome in examination.

One region of interest stood out especially, a SNP on GGA 2 at position 6.2 Mb. Nine of ten tumor samples demonstrated a homozygous line 7 (susceptible) genotype. The region of interest is about 46 Kb in size based on the nearest heterozygous SNPs. This region contains a number of candidate genes and will be examined in future investigations.

This pilot study confirmed that the custom 15K SNP array demonstrates sufficient power and is adequate for future array analysis, better utilizing monetary resources as the 15K array is substantially less expensive than the 600K array. SNP microarrays grant us a global perspective of the genome at key positions of interest guiding more precise efforts with advanced technology.

The widely used MD control strategy is vaccination; however, selection of genetic resistance is an additional powerful means of control, especially because it augments vaccination strategies. The Cheng lab set out to identify genes and certain gene products associated with MD resistance. They examined QTL alleles associated with resistance to MD. Multiple approaches (QTL scans, selective sweeps, transcript profiling, ASE screens, and chromatin immunoprecipitation) were used in order to find candidate genes and/or pathways of high confidence; using multiple approaches harnesses the unique strengths of each approach granting more power of analysis. In addition, given the mass amount of data, multiple approaches acted as filtering techniques and made data analysis more manageable.

Genome-wide QTL scans were performed on about 1,100 6x7 F6 MD resource population. Using the F6 population took advantage of prior SNP arrays as DNA from these birds were also genotyped. The custom affymetrix 15K array includes 2,256 SNPs associated with ASE in response to MDV infection, 4,497 SNPs associated with selective sweeps for MD resistance, and the remaining 8,097 SNPs were previously validated and used to increase precision of analysis. ASE associated SNPs demonstrate a transcriptional response to MDV infection; when allelic variation is observed, there must be a polymorphic *cis*-acting element present for the gene of interest (2).

Selective sweep analysis for MD genetic resistance was performed by sequencing line 6 and 7 genomes with paired end reads of average size 2.5 Kb. Reads were aligned to the chicken reference genome and SNPs were analyzed. Both lines were divergently selected for from the same origin population; therefore, divergently fixed SNP haplotype blocks and their respective lengths, chromosome haplotype length (CHL), were examined. The degree of selection for a particular allele is indicative of the integrity of linkage disequilibrium (LD), essentially how

'fixed' the allele is (90). Results indicate that about 25% of ASE SNPs are in or near genes in regions of selective sweep. Genomic regions that share association with ASE and selective sweeps show highly significant association suggestive of MD resistance.

QTL screens offer an approximation of gene location in the genome relative to the detectable marker, often with poor resolution. To increase resolution, differential gene expression was measured with the expectation that genes and pathways differentially expressed and suggestive of MD resistance and/or response to infection would reveal candidate gene locations when combined with QTL mapping (91,92). Therefore, differential gene expression was measured in response to infection and between lines of different and defined MD resistance. This collaborative approach revealed many candidate genes and pathways associated with MD resistance and/or MDV infection (91,92), which collectively suggest that the degree of host immune response was directly correlated with vulnerability to tumorigenesis. This makes sense because select immune cells, mainly CD4+ T lymphocytes—the primary targets for transformation—congregate to areas of infection.

Two ASE screens were performed on F1 progeny (line 6x7) in response to MDV infection. The entire chicken genome was screened with the exception of MHC genes, since line 6x7 progeny are homozygous for MHC alleles (93,94). Half of the birds were infected with MDV at 2 weeks of age and RNA samples from spleen were collected at 1, 4, 7, 11, 13, and 15 dpi. In the first study, replicate RNA was pooled and analyzed from a single time point. In the second study, 7 individual RNA samples were analyzed from infected and uninfected birds. SNPs and allelic ratios were identified in both studies and if the allelic ratio changed in response to MDV infection, then there was a *cis*-acting regulatory element affecting the expression of the SNP associated gene. The second study determined 4,528 SNPs in 3,718 genes demonstrating ASE variation in response to MDV infection and 70% of first study results confirmed ASE variation, suggesting that RNA sequencing is an effective means of measuring valid ASE SNPs. Differential gene expression was also measured from RNA sequenced data. The genes and pathways associated with MD genetic resistance were as expected. Several pathways demonstrated both pathway enrichment and significant ASE, particularly of early pathway genes (94). In key regulatory pathways, upstream factors similar to the above-mentioned enriched genes demonstrating ASE may influence expression of subsequent downstream factors (95). However, this experiment did not address the somatic alterations likely associated with transformation that are believed to be downstream in key pathways.

In addition to the collaborative efforts above, chromatin immunoprecipitation (ChIP) was performed on Meq homo- and heterodimers followed by next generation sequencing (NGS) analysis to identify regions of the chicken genome to which Meq binds. Meq is considered one of the most oncogenic genes in the scenario of MDV transformation, and identification of the genes and pathways influenced by Meq helped the Cheng lab to reveal candidates involved in transformation. In addition, RNA expression of gene sets was measured with microarrays between cells expressing Meq and those that did not. The correspondence between differential gene expression and regions from ChIP revealed 351 candidate genes within 2 Kb of Meq-binding sites. Pathways analysis suggested enrichment of genes in the mitogen-activated protein kinase (MAPK) and WNT signaling pathways. This information can be combined with tumor genome profiling (objective 2) to better understand MDV induced transformation,

because MD tumor formation is likely driven by somatic mutaions and Meq-directed regulatory changes.

The above tests and collaborations between them have revealed, in part, how MDV likely influences the chicken genome, gene products, and pathways. This information will guide our analysis and hopefully corroborate our results; but most importantly we hope to expand upon our understanding of the forces driving MDV transformation and genetic elements associated with MD resistance.

Research Design and Methods

We aim to both improve upon existing and generate alternative means of control of MD. This requires investigation of the etiology of MD, specifically the genetic origin of MDV-induced tumorigenesis. Cytogenetics, molecular genetics, and genomic analysis enable us to investigate what somatic alterations and integration events are at the origin of MDV-induced T-cell lymphomas and strengthen our understanding of MD tumorigenesis.

Evidence suggests that certain integration sites are preferred in tumor lineages and tumors are mostly monoclonal in individual birds. MDV integration is necessary for transformation although singularly insufficient because integration is variable and does not solely guarantee transformation (27), suggesting that additional somatic alterations are necessary for transformation events and tumorigenesis. In addition, as mentioned above, MDV oncogenes (e.g. Meq) regulate many host genes and pathways. We aim to reveal the somatic alterations necessary for tumorigenesis and the interplay between somatic alterations, MDV-induced regulation of gene expression and pathways, and MDV integration.

Our first objective is to characterize large-scale somatic alterations and MDV integration on 72 MD tumors from male birds using a custom Affymetrix 15K SNP chip array and cytogenetic analyses, respectively. Together these approaches will reveal a global perspective of SNPs, CNVs, LOH, translocations, aneuploidy, and other large-scale somatic alterations, as well as their potential relationship to each other. Resulting data from objective 1 will guide further analysis to identify somatic mutations via characterization of sequenced DNA (objective 2) and to elucidate gene expression alterations from sequenced RNA (objective 3). Objectives 2 and 3 entail a more precise investigation of DNA and RNA from 22 tumors selected based on somatic alteration signatures revealed in objective 1. DNA sequencing grants detection precision at the single nucleotide level, allowing us to decipher: somatic mutations, genes, and alleles. RNA sequencing will illustrate pathways associated with MDV-induced transformation events and presumably confirm suspected candidate genomic alterations. In addition, this information will help us to confirm and further investigate results from objective 1: the LOH and CNVs from the SNP microarrays and the MDV integration sites. In sum, we gain a broad genomic view of large-scale mutations and MDV integration sites by using SNP arrays and cytogenetics, respectively. These results will guide us to investigate alterations in DNA and RNA expression more precisely via NGS. By combining and corroborating results from arrays, cytogenetics, and NGS we hope to identify genetic alterations responsible for MDV-induced transformation.

Experimental Birds, Materials, and Tissue Sampling

All birds in this experiment, lines 7x6 F_1 progeny, will provide true biological replicates due to high levels of inbreeding in parental lines 6 and 7, of which both genomes have been

extensively characterized. Use of the heterozygous 7x6 F_1 progeny will allow us to characterize tumor specific somatic mutations in reference to both genetic backgrounds.

We will subcutaneously infect ~200 F1 line 7x6 chicks at hatch with 1,000 pfu JM/102W strain MDV, which is classified as the virulent (vMDV) pathotype (11), and should yield sufficient monoclonal tumors that are fairly homogenous in regard to tumor cell concentration. Birds will be necropsied for gross late-stage tumors after 8 weeks post inoculation or until moribund. Larger tumors are preferred to easily discriminate between tumor and surrounding tissue. Collected tumor samples will be divided: a large portion in cell suspension (cytogenetic analysis), and remaining portions in RNAlater (RNA sequence analysis) and frozen at -80 degrees C (DNA microarray and sequence analysis). For the genomic control (DNA microarray and sequencing analysis), blood will be sampled from six line 6x7 F1 progeny not infected with MDV (unchallenged). Nucleated red blood cells (RBCs) in these samples will provide the control DNA necessary. For the transcriptional control (RNA sequencing analysis), we will sample RNA from 15 unchallenged birds. The RNA will be isolated from purified CD4+ T cells, the cells most likely to be transformed by MDV, thereby providing an accurate representation of RNA in this lymphocyte microenvironment without the influence of MDV transformation. At 2, 4 and 6 weeks of age 5 unchallenged birds will be euthanized and necropsied for CD4+ T cells collection and verification of no MD, respectively.

Objective 1: SNP Array and Cytogenetic Screens

Together SNP arrays and cytogenetic screens will provide insight toward potential large-scale somatic alterations (i.e. LOH and CNVs) in MD tumors as well and whether tumors from the same bird are monoclonal. The initial 72 tumors and 24 controls will be screened via a custom Affymetrix 15K SNP chip with 9,042 fully informative markers, which are spaced once per 120Kb of the chicken genome, on average. The 15K array is sufficient due to cost and large regions of linkage disequilibrium. Tumors from different tissue types (e.g. heart, spleen, thymus) will be sampled, but we suspect the majority of tumors to be gonadal, because JM/102W preferentially induce large, fairly homogenous gonadal tumors (96). Each 96-well plate will contain 24 controls consisting of three sets of 8 technical replicates from: one line 6 bird, one line 7 bird, and one 7x6 F1 bird. Replicates will be randomly assigned to each well to test for row bias. Parental lines 6 and 7 are fixed for heterozygous alleles at each of the 9,042 SNPs. Line 6, 7, and 7x6 F1 represent homozygous AA, homozygous BB, and heterozygous AB genotypes, respectively. Microarray signal strengths of each SNP from tumor samples will be compared to signal strengths of known controls, providing especially useful insight into CNVs and LOH (88,89). Given that MD tumors consistently arise within 3-6 weeks, we expect alterations not to hide at low frequency and therefore we should have sufficient power to detect them; in other words, the cells within fairly homogenous tumors will not have sufficient time to accumulate passenger mutations that would distract from the discovery of driver mutations. We will also screen for candidate alterations from aforementioned prior efforts. One example is the SNP on GGA 2 at position 6.2 Mb (homozygous for line 7 susceptible alleles), which is highly prevalent in preliminary data (9 out of 10 tumors). We will determine GGA 2 allelic frequency in all tumor and control samples via Pyrosequencing.

Tumor samples will be filtered based on their cytogenetic profiles. We will select for samples that suggest monoclonal integration events specifically on chromosomes that have full or near complete sequence assembly. Many chromosomes in the chicken are too small and/or compact to properly annotate on the reference chicken (Gallus gallus) genome (i.e. 16, 25, 28 and higher, and W); alterations on these chromosomes would be difficult to detect in downstream NGS analysis because they could not be properly mapped and assembled. We will also select for diverse integration profiles that may differ in which annotated chromosome and/or region of the chromosome the virus integrates.

Results from SNP arrays will be filtered based on our understanding of MDV-induced tumorigenesis as well as the extensive literature on human tumorigenesis. Regions with somatic alterations of sufficient frequency (allelic frequency of greater than 10%) in tumors will be compared to:

- Genes associated with resistance to MD and which allele (line 6 or line 7) confers resistance (85).
- Genes that cooperate or are regulated directly or indirectly by Meq.
- Known oncogenes and tumor suppressor genes, especially those from The Cancer Genome Atlas (TCGA). We will examine the number of alleles affected (haplosufficient or haploinsuffient), the mode of action (dominant or recessive), the altered function of the gene (e.g. gain-of-function or loss-of-function), and the gene group (oncogene or tumor suppressor gene).
- Genes within and that interact with specific pathways, such as the MAPK pathway that includes three leucine zipper proteins (bZIP) (e.g c-Jun), and the DNA binding sites of bZIP proteins (e.g. AP-1 sites).
- Genes that influence telomeres and/or telomerase.

Samples to undergo cytogenetic and SNP array screening will originate from late-stage tumors, and will likely possess karyotype abnormalities. Previous efforts revealed that tetraploidy and aneuploidy were a common characteristic of MD tumor cell karyotypes (27). We will investigate the appearance of abnormal karyotypes based on viral integration sites to determine if certain chromosomes are targeted and if certain integration events are preferred. And if so, how do such relationships influence tumor cell mitosis, particularly segregation.

Objectives 2 and 3: NGS of DNA and RNA from MDV-induced tumor samples.

DNA and RNA from 22 MD tumors, specifically selected from the 72 MD tumors mentioned above, will be sequenced. DNA sequencing will identify integration sites and somatic alterations, which include nucleotide substitutions, small insertions and deletions, copy number alterations, and chromosomal rearrangements (Figure 9) (97). RNA sequencing will allow profiling of gene expression. Tumor DNA and RNA sequence data will be compared to 6 DNA and 15 RNA controls, which will also be sequenced. In most human cancer genomics studies, tumor DNA and RNA are compared to DNA and RNA from surrounding tissue samples. However, the control samples we will use are from different chickens. We argue that it is not necessary to have controls for each individual given the high level of inbreeding and genetic consistency of F1 progeny and parental lines. Furthermore, parental lines 6 and 7 have

been genomically sequenced with over 60x coverage providing valuable information. DNA and RNA from samples will be sequenced via Illumina HiSeq 2000 with 150 base paired end reads to generate 150+ Gb of sequence equivalent to 15+x genome (DNA) coverage per sample. DNA and RNA alterations will be screened for deviation from the heterozygous 1:1 allelic ratio (genome) and from the RNA control allelic ratio (transcriptome). Collectively, sequence information from tumor DNA and RNA, when compared to their respective controls and pre-existing line 6 and 7 data, will allow us to detect somatic genomic mutations and factors regulating gene expression.

Read trimming, quality control, mapping and alignment. Sequenced tumor and control samples will be aligned to a reference genome and compared against each other to determine somatic alterations and differential gene expression. Sequence will be analyzed bioinformatically utilizing pipelines similar to those used in human cancer genomics studies (Figure 8). Reads will be checked for quality, trimmed of low quality bases and adapters, and rechecked for quality via Trimmomatic (98). Trimmed reads of ample quality will then be mapped and assembled to the Gallus gallus 4 (Galgal4) reference genome via the Tophat package (99). Read quality will be further analyzed in relation to mapping and alignment via SAMStat (100). Aligned reads will be analyzed through a series of algorithms and software packages for detection of variants and gene expression (Figure 8). It is empirically known that using several independent algorithms in combination, rather than just a single algorithm, to detect a given variant type (e.g. indels) is less likely to generate false positives (101). We will use the combined algorithm approach based on the cutting edge but still tried and tested methods of Chen et al. (2014), as there is no consensus for optimal algorithm combinations (101). False positives will also be limited by: ensuring alterations have at least three reads, comparing genomic (DNA) and transcriptomic (RNA) data, and comparing paired samples to target driver mutations.

SNV detection. Single nucleotide variants (SNVs) are the most frequent alterations in cancer genomes (101). Of the numerous SNV detection algorithms, we will use SomaticSniper (102) and Bambino (103) because they were designed to detect somatic mutations using tumor and matched normal genomic sequences. Resulting SNVs will be compared to the databases previously explained above in SNP array analysis and to the database of previously identified chicken SNPs in the UCSC genome browser.

Indel detection. It is challenging to detect indels because they are difficult to map (104) and they occur at low frequency compared to SNVs (105,106). In addition, many algorithms suffer from lack of precision and cannot detect medium-sized indels (101). We will use four algorithms collectively to eliminate false positives: SAMtools (107), GATK (108), Pindel (109), and Bambino (103). Pindel stands out from other algorithms because it achieves high precision and sensitivity by allowing mismatches (101).

Confirmation of ploidy and detection of CNVs. Aligned reads will be analyzed for total read coverage (RC) at each locus to confirm ploidy from objective 1 analysis and to detect CNVs (Figure 9). We will divide per coordinate RC by average genomic RC (X), which should be about 15. Per coordinate copy number will be calculated by CNV=2(RC/X). CNV is expected to be 2—a value greater than 2 will be a gain in copy number. We will then measure CNVs in sliding windows of 50 SNPs to increase power and average noise. A t-test will be determined

for detected duplications per line, sample, and locus and t-test significance will be determined via permutation test (110) of 1000 randomizations of sample IDs across lines. Detected CNVs will be compared to previously determined CNVs in lines 6 and 7 and a custom Agilent comparative genomic hybridization (CGH) array (111).

Detection of LOH, structural variants, and gene fusions. Structural changes in the genome (e.g. chromosome deletions, insertions, and translocations) affect cancer genes by resulting in a fusion transcript or transcriptional dysregulation. Whole genome sequencing (WGS) will allow us to expand upon objective 1 data to characterize somatic rearrangements and their breakpoints with base-pair resolution. Paired-end sequencing tends generate many false positives (101), therefore we will use a collective analysis approach using two or more algorithms from the following: BreakDancer (112), CREST (113), GASV-Pro (114), and Genome STRiP (115). In addition, we will supplement data with local assembly of candidate regions.

To detect LOH we will compute minor allele frequency (MAF) in windows containing 400 SNPs. The combination of 15x coverage and 400 SNPs per window will generate 4,800 genotypes to be analyzed on average per coordinate and sample allowing for high precision estimation of MAF. A t-test will be performed on average MAFs per line, sample, and coordinate in the same manner as described above with CNVs and significance (p value) determined. Gene fusions arising from structural rearrangements are a major mechanism driving tumor formation (Ding et al., 2014). We will employ two or more of the recently developed algorithms to detect gene fusions from RNA-seq data: TopHat-fusion (116), deFuse (117), MapSplice (118), ChimeraScan (119), and BreakFusion (120).

Determination of driver mutations and pathways. The next challenge in analysis is to determine which of the detected somatic alterations are likely driver mutations (Figure 8). To distinguish driver mutations from passenger mutations, we will employ three strategies: variant prediction; recurrence and frequency assessment; and pathway or network analysis (101). We will use SIFT (121) and polyPhen (122) to detect deleterious mutations in coding regions. We will identify SNPs in non-coding regions 2 Kb upstream of transcriptional start sites and label associated genes, especially if they are suspected to associate with Meq homo- or heterodimer-induced regulation of transcription. Somatic alterations will be compared to 15K SNP data (objective 1), TCGA database, and the chicken alterations on the UCSC genome browser. DNA and RNA variations will be tested for pathway association (e.g. MAPK and WNT signaling) or enrichment via DAVID (123).

Validation of mutations. After we have exhausted our bioinformatics analysis we should end up with a manageable amount of candidate driver variations. To validate these alterations were actually present in tumor samples, we will incorporate specially designed Pyrosequencing arrays on all 72 MD tumor samples and addition female tumor samples.

Figures

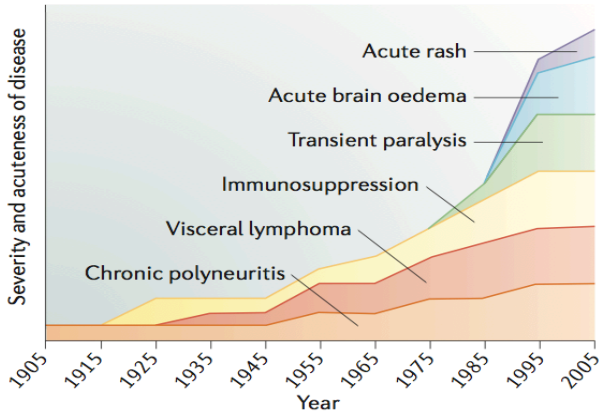


Figure 1: The changing clinical picture of MD as a result of the evolution of MDV virulence (5).

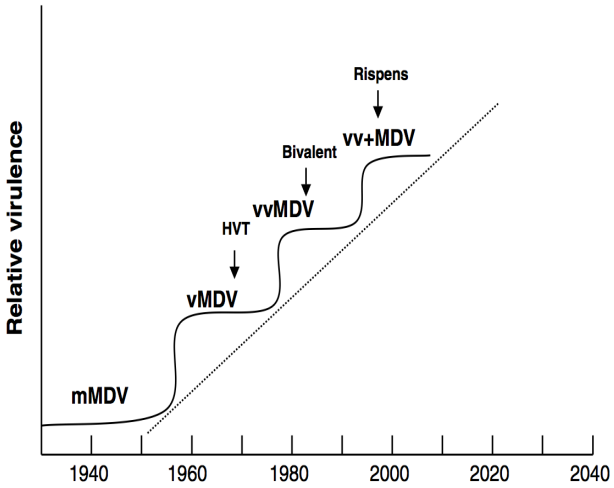


Figure 2: The evolution of MDV virulence in relation to vaccine breaks. HVT, herpes virus of turkeys; Bivalent, HVT and serotype 2 (SB-1) vaccines; Rispens, CVI988 strain (11,14).

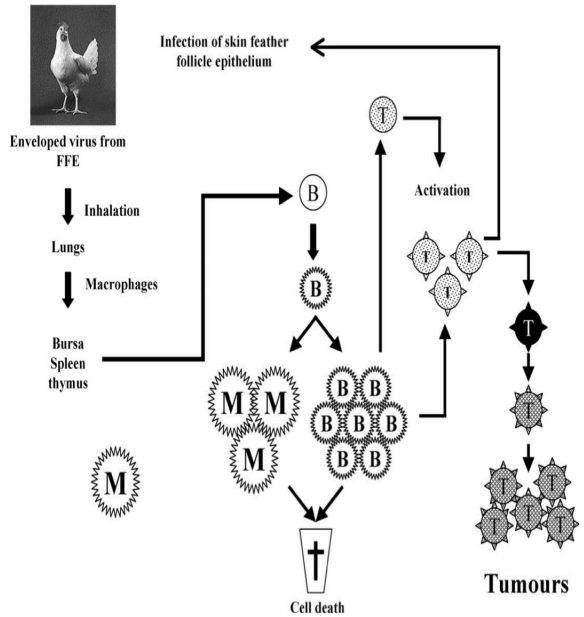


Figure 3: The pathogenesis of MD (14).

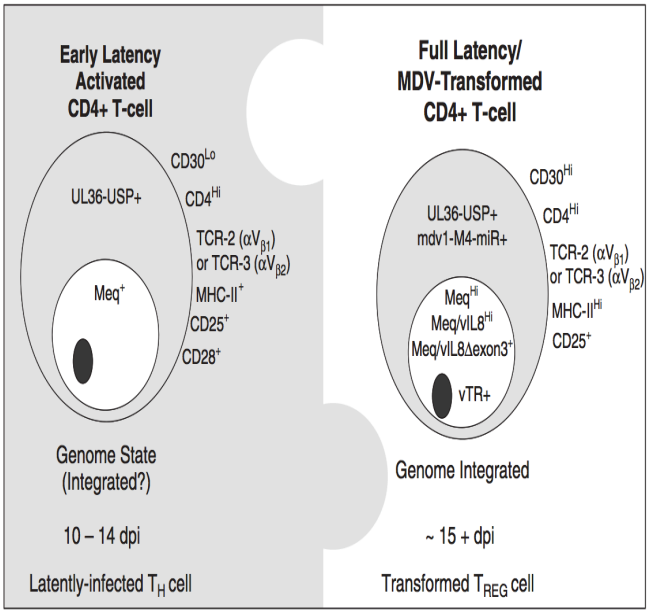


Figure 4: The developing scenario of MDV transformation (10).

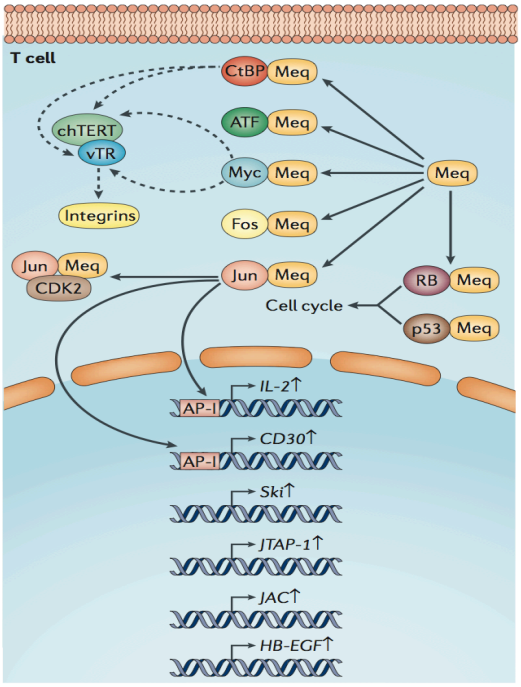


Figure 5: The role of Meq in MDV transformation (5).

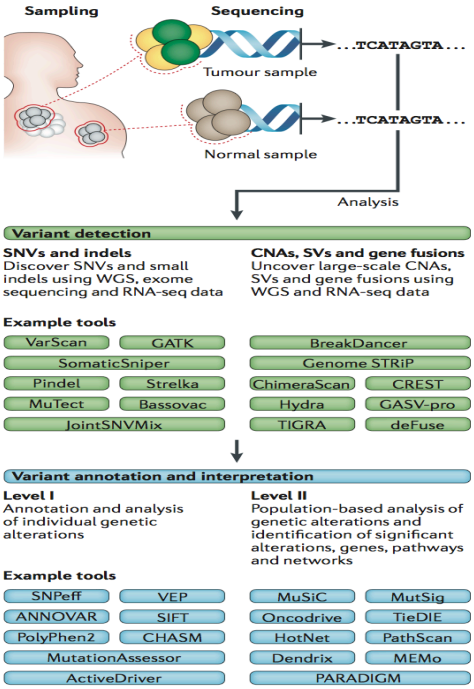
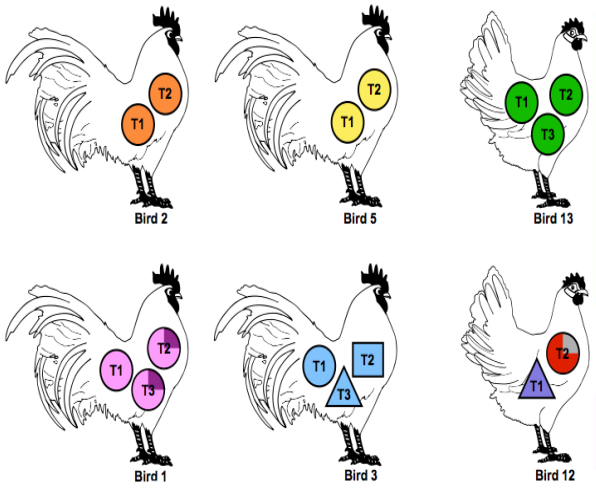


Figure 8: Bioinformatic analysis of DNA from tumor and control samples in human cancer genomics (101).



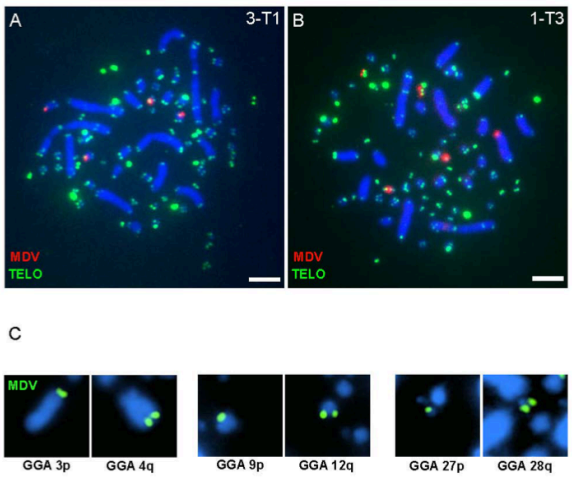


Figure 7: Tumor lineage analysis of MDV-induce lymphomas. The mapped MDV integration sites in 15 tumors in 6 birds provided a means to characterize tumor lineage, mostly monoclonal (27).

Figure 6: Localization of MDV integration at the telomeres of chicken chromosomes. Panels **A** and **B** show MDV integration sites (*red*) in the telomeres (*green*) of chicken chromosomes (*blue*) in MDV-induced lymphoma tumor cells. Panel **C** demonstrates that MDV integrations (*green*) occurred in both macro- and micro-chromosomes (*blue*) and did not preferentially discriminate between p and q arms (27).

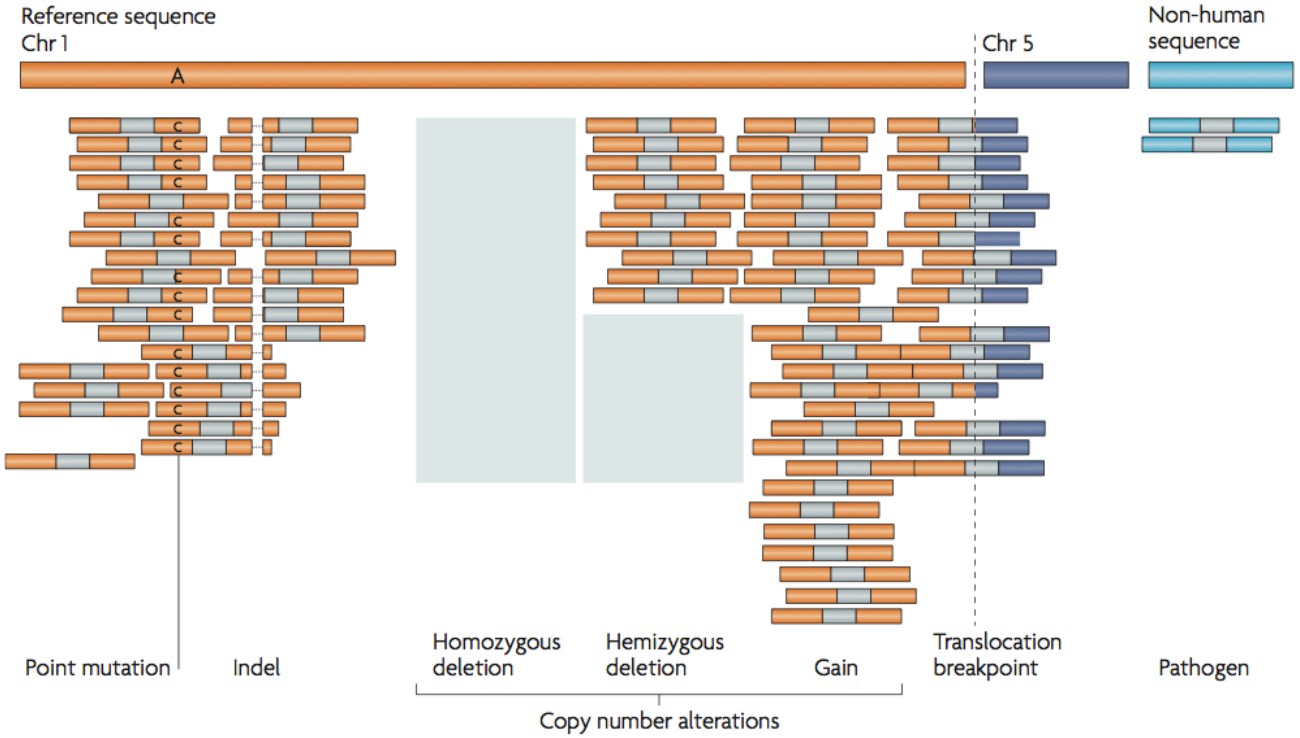


Figure 9: Genome alterations detectable via second-generation sequencing in Humans (97).

Timetable

Birds and Samples	2014
Cytogenetics and SNP Screens	2014-2015
Next Generation Sequencing	2014-2015
SNP Validation	2015
Analyses	2014-2016

Potential Funding Sources and Rationale

This project is funded through USDA Agriculture and Food Research Initiative (AFRI) grant 113.29.13.30. AFRI grants fund research that address important aspects of food and agriculture. This project aims to enhance control of Marek's Disease.

Literature Cited

1. Muir WM, Wong GK-S, Zhang Y, Wang J, Groenen MA, Crooijmans RP, et al. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. Proc Natl Acad Sci. 2008;105(45):17312–7.
2. Cheng HH. Genome biology of Marek'sdisease: Viral integration and genome alterations in genetically resistant and susceptible stocks. AFRI grant submission; 2013.
3. Havenstein GB, Ferket PR, Qureshi MA. Growth, livability, and feed conversion of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets. Poult Sci. 2003 Oct;82(10):1500–8.
4. Davison TF, Nair V, Institute for Animal Health (Great Britain). Marek's disease an evolving problem [Internet]. Amsterdam; Boston: Elsevier; 2004 [cited 2014 Jul 26]. Available from: http://site.ebrary.com/id/10177015
5. Osterrieder N, Kamil JP, Schumacher D, Tischer BK, Trapp S. Marek's disease virus: from miasma to model. Nat Rev Microbiol. 2006 Apr;4(4):283–94.
6. Pappenheimer AM, Dunn LC, Cone V. STUDIES ON FOWL PARALYSIS (NEUROLYMPHOMATOSIS GALLINARUM) : I. CLINICAL FEATURES AND PATHOLOGY. J Exp Med. 1929 Jan 1;49(1):63–86.
7. Pappenheimer AM, Dunn LC, Seidlin SM. STUDIES ON FOWL PARALYSIS (NEUROLYMPHOMATOSIS GALLINARUM) : II. TRANSMISSION EXPERIMENTS. J Exp Med. 1929 Jan 1;49(1):87–102.
8. Churchill AE, Biggs PM. Agent of Marek's disease in tissue culture. Nature. 1967 Jul 29;215(5100):528–30.
9. Churchill AE, Payne LN, Chubb RC. Immunization against Marek's disease using a live attenuated virus. Nature. 1969 Feb 22;221(5182):744–7.
10. Robertson ES, editor. Cancer Associated Viruses [Internet]. Boston, MA: Springer US; 2012 [cited 2014 Jul 26]. Available from: http://www.springerlink.com/index/10.1007/978-1-4614-0016-5
11. Witter RL. Increased Virulence of Marek's Disease Virus Field Isolates. Avian Dis. 1997 Jan;41(1):149.
12. Witter RL. Avian tumor viruses: persistent and evolving pathogens. Acta Vet Hung. 1997;45(3):251–66.
13. Witter RL. The changing landscape of Marek's disease. Avian Pathol. 1998 Apr;27(sup1):S46–S53.
14. Nair V. Evolution of Marek's disease – A paradigm for incessant race between the pathogen and the host. Vet J. 2005 Sep;170(2):175–83.
15. Barrow AD, Burgess SC, Baigent SJ, Howes K, Nair VK. Infection of macrophages by a lymphotropic herpesvirus: a new tropism for Marek's disease virus. J Gen Virol. 2003 Oct;84(Pt 10):2635–45.

16. Schat KA, Nair V. Diseases of poultry. 12th ed. Saif YM, Barnes HJ, editors. Ames, Iowa: Blackwell Pub. Professional; 2008.
17. Lee SI, Ohashi K, Sugimoto C, Onuma M. Heparin inhibits plaque formation by cell-free Marek's disease viruses in vitro. J Vet Med Sci Jpn Soc Vet Sci. 2001 Apr;63(4):427–32.
18. Kaleta EF, Neumann U. [Investigations on the mode of transmission of the herpesvirus of turkeys in vitro]. Avian Pathol J WVPA. 1977;6(1):33–9.
19. Schumacher D, Tischer BK, Reddy SM, Osterrieder N. Glycoproteins E and I of Marek's disease virus serotype 1 are essential for virus growth in cultured cells. J Virol. 2001 Dec;75(23):11307–18.
20. Tischer BK, Schumacher D, Messerle M, Wagner M, Osterrieder N. The products of the UL10 (gM) and the UL49.5 genes of Marek's disease virus serotype 1 are essential for virus growth in cultured cells. J Gen Virol. 2002 May;83(Pt 5):997–1003.
21. Lupiani B, Lee LF, Reddy SM. Protein-coding content of the sequence of Marek's disease virus serotype 1. Curr Top Microbiol Immunol. 2001;255:159–90.
22. Taddeo B, Roizman B. The virion host shutoff protein (UL41) of herpes simplex virus 1 is an endoribonuclease with a substrate specificity similar to that of RNase A. J Virol. 2006 Sep;80(18):9341–5.
23. Skepper JN, Whiteley A, Browne H, Minson A. Herpes simplex virus nucleocapsids mature to progeny virions by an envelopment --> deenvelopment --> reenvelopment pathway. J Virol. 2001 Jun;75(12):5697–702.
24. Calnek BW, Adldinger HK, Kahn DE. Feather follicle epithelium: a source of enveloped and infectious cell-free herpesvirus from Marek's disease. Avian Dis. 1970 May;14(2):219–33.
25. Calnek BW, Schat KA, Ross LJ, Shek WR, Chen CL. Further characterization of Marek's disease virus-infected lymphocytes. I. In vivo infection. Int J Cancer J Int Cancer. 1984 Mar 15;33(3):389–98.
26. Lesnik F, Chudý D, Bogdan J, Vrtiak OJ, Rudic M. [Testing the immunogenicity of the dermal antigen in Marek's disease virus]. Veterinární Medicína. 1978 Jul;23(7):421–30.
27. Robinson CM, Hunt HD, Cheng HH, Delany ME. Chromosomal integration of an avian oncogenic herpesvirus reveals telomeric preferences and evidence for lymphoma clonality. Herpesviridae. 2010;1(1):5.
28. Abujoub AA, Coussens PM. Evidence that Marek's disease virus exists in a latent state in a sustainable fibroblast cell line. Virology. 1997 Mar 17;229(2):309–21.
29. Yamaguchi T, Kaplan SL, Wakenell P, Schat KA. Transactivation of latent Marek's disease herpesvirus genes in QT35, a quail fibroblast cell line, by herpesvirus of turkeys. J Virol. 2000 Nov;74(21):10176–86.
30. Powell PC, Payne LN, Frazier JA, Rennie M. T lymphoblastoid cell lines from Marek's disease lymphomas. Nature. 1974 Sep 6;251(5470):79–80.
31. Witter RL, Stephens EA, Sharma JM, Nazerian K. Demonstration of a Tumor-Associated Surface Antigen in Marek's Disease. 1975 Jul;115:177–83.
32. Burgess SC, Kaiser P, Davison TF. A novel lymphoblastoid surface antigen and its role in Marek's disease (MD). In: Silva RF, Cheng HH, Coussens PM, Lee LF, Velicer LF, editors.

Current Research on Marek's Disease. Michigan State University, East Lansing, Michigan: Kennett Square, Pa. : American Association of Avian Pathologists, c1996; 1996.
33. Ross N, O'Sullivan G, Rothwell C, Smith G, Burgess SC, Rennie M, et al. Marek's disease virus EcoRI-Q gene (meq) and a small RNA antisense to ICP4 are abundantly expressed in CD4+ cells and cells carrying a novel lymphoid marker, AV37, in Marek's disease lymphomas. J Gen Virol. 1997 Sep;78 (Pt 9):2191–8.
34. Burgess SC, Young JR, Baaten BJG, Hunt L, Ross LNJ, Parcells MS, et al. Marek's disease is a natural model for lymphomas overexpressing Hodgkin's disease antigen (CD30). Proc Natl Acad Sci U S A. 2004 Sep 21;101(38):13879–84.
35. Delecluse H-J, Hammerschmidt W. Status of Marek's disease virus in established lymphoma cell lines: herpesvirus integration is common. J Virol. 1993;67(1):82–92.
36. Delecluse HJ, Schüller S, Hammerschmidt W. Latent Marek's disease virus can be activated from its chromosomally integrated state in herpesvirus-transformed lymphoma cells. EMBO J. 1993 Aug;12(8):3277–86.
37. Hall CB, Caserta MT, Schnabel K, Shelley LM, Marino AS, Carnahan JA, et al. Chromosomal integration of human herpesvirus 6 is the major mode of congenital human herpesvirus 6 infection. Pediatrics. 2008 Sep;122(3):513–20.
38. Li M, Mizuuchi M, Burke TR, Craigie R. Retroviral DNA integration: reaction pathway and critical intermediates. EMBO J. 2006 Mar 22;25(6):1295–304.
39. Jones C. Alphaherpesvirus latency: its role in disease and survival of the virus in nature. Adv Virus Res. 1998;51:81–133.
40. Gulley ML, Raphael M, Lutz CT, Ross DW, Raab-Traub N. Epstein-Barr virus integration in human lymphomas and lymphoid cell lines. Cancer. 1992 Jul 1;70(1):185–91.
41. Luppi M, Marasca R, Barozzi P, Ferrari S, Ceccherini-Nelli L, Batoni G, et al. Three cases of human herpesvirus-6 latent infection: integration of viral genome in peripheral blood mononuclear cell DNA. J Med Virol. 1993 May;40(1):44–52.
42. Schat KA, Chen CL, Calnek BW, Char D. Transformation of T-lymphocyte subsets by Marek's disease herpesvirus. J Virol. 1991 Mar;65(3):1408–13.
43. Lee SI, Ohashi K, Morimura T, Sugimoto C, Onuma M. Re-isolation of Marek's disease virus from T cell subsets of vaccinated and non-vaccinated chickens. Arch Virol. 1999;144(1):45–54.
44. Arbuckle JH, Medveczky MM, Luka J, Hadley SH, Luegmayr A, Ablashi D, et al. The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. Proc Natl Acad Sci U S A. 2010 Mar 23;107(12):5563–8.
45. Kishi M, Harada H, Takahashi M, Tanaka A, Hayashi M, Nonoyama M, et al. A repeat sequence, GGGTTA, is shared by DNA of human herpesvirus 6 and Marek's disease virus. J Virol. 1988 Dec;62(12):4824–7.
46. Secchiero P, Nicholas J, Deng H, Xiaopeng T, van Loon N, Ruvolo VR, et al. Identification of human telomeric repeat motifs at the genome termini of human herpesvirus 7: structural analysis and heterogeneity. J Virol. 1995 Dec;69(12):8041–5.
47. Silver S, Tanaka A, Nonoyama M. Transcription of the Marek's disease virus genome in a nonproductive chicken lymphoblastoid cell line. Virology. 1979 Feb;93(1):127–33.

48. Tanaka A, Silver S, Nonoyama M. Biochemical evidence of the nonintegrated status of Marek's disease virus DNA in virus-transformed lymphoblastoid cells of chicken. Virology. 1978 Jul 1;88(1):19–24.
49. Maray T, Malkinson M, Becker Y. RNA transcripts of Marek's disease virus (MDV) serotype-1 in infected and transformed cells. Virus Genes. 1988 Oct;2(1):49–68.
50. Kung HJ, Xia L, Brunovskis P, Li D, Liu JL, Lee LF. Meq: an MDV-specific bZIP transactivator with transforming properties. Curr Top Microbiol Immunol. 2001;255:245–60.
51. Shamblin CE, Greene N, Arumugaswami V, Dienglewicz RL, Parcells MS. Comparative analysis of Marek's disease virus (MDV) glycoprotein-, lytic antigen pp38- and transformation antigen Meq-encoding genes: association of meq mutations with MDVs of high virulence. Vet Microbiol. 2004 Sep 8;102(3-4):147–67.
52. Brown AC, Baigent SJ, Smith LP, Chattoo JP, Petherbridge LJ, Hawes P, et al. Interaction of MEQ protein and C-terminal-binding protein is critical for induction of lymphomas by Marek's disease virus. Proc Natl Acad Sci U S A. 2006 Feb 7;103(6):1687–92.
53. Lupiani B, Lee LF, Cui X, Gimeno I, Anderson A, Morgan RW, et al. Marek's disease virus-encoded Meq gene is involved in transformation of lymphocytes but is dispensable for replication. Proc Natl Acad Sci U S A. 2004 Aug 10;101(32):11815–20.
54. Ajithdoss DK, Reddy SM, Suchodolski PF, Lee LF, Kung H-J, Lupiani B. In vitro characterization of the Meq proteins of Marek's disease virus vaccine strain CVI988. Virus Res. 2009 Jun;142(1-2):57–67.
55. Suchodolski PF, Izumiya Y, Lupiani B, Ajithdoss DK, Gilad O, Lee LF, et al. Homodimerization of Marek's disease virus-encoded Meq protein is not sufficient for transformation of lymphocytes in chickens. J Virol. 2009 Jan;83(2):859–69.
56. Suchodolski PF, Izumiya Y, Lupiani B, Ajithdoss DK, Lee LF, Kung H-J, et al. Both homo and heterodimers of Marek's disease virus encoded Meq protein contribute to transformation of lymphocytes in chickens. Virology. 2010 Apr 10;399(2):312–21.
57. Levy AM, Gilad O, Xia L, Izumiya Y, Choi J, Tsalenko A, et al. Marek's disease virus Meq transforms chicken cells via the v-Jun transcriptional cascade: a converging transforming pathway for avian oncoviruses. Proc Natl Acad Sci U S A. 2005 Oct 11;102(41):14831–6.
58. Reinke AW, Grigoryan G, Keating AE. Identification of bZIP interaction partners of viral proteins HBZ, MEQ, BZLF1, and K-bZIP using coiled-coil arrays. Biochemistry (Mosc). 2010 Mar 9;49(9):1985–97.
59. Katiyar S, Jiao X, Wagner E, Lisanti MP, Pestell RG. Somatic excision demonstrates that c-Jun induces cellular migration and invasion through induction of stem cell factor. Mol Cell Biol. 2007 Feb;27(4):1356–69.
60. Buza JJ, Burgess SC. Modeling the proteome of a Marek's disease transformed cell line: a natural animal model for CD30 overexpressing lymphomas. Proteomics. 2007 Apr;7(8):1316–26.
61. Zhao Y, Kurian D, Xu H, Petherbridge L, Smith LP, Hunt L, et al. Interaction of Marek's disease virus oncoprotein Meq with heat-shock protein 70 in lymphoid tumour cells. J Gen Virol. 2009 Sep;90(Pt 9):2201–8.

62. Kim S. The Meq Protein of Marek's Disease Virus (MDV) binds to the chicken SKP-1 protein. 2010; University of Delaware, Newark, DE.
63. Chinnadurai G. Transcriptional regulation by C-terminal binding proteins. Int J Biochem Cell Biol. 2007;39(9):1593–607.
64. Shack LA, Buza JJ, Burgess SC. The neoplastically transformed (CD30hi) Marek's disease lymphoma cell phenotype most closely resembles T-regulatory cells. Cancer Immunol Immunother CII. 2008 Aug;57(8):1253–62.
65. Subramaniam S, Johnston J, Preeyanon L, Brown CT, Kung H-J, Cheng HH. Integrated Analyses of Genome-Wide DNA Occupancy and Expression Profiling Identify Key Genes and Pathways Involved in Cellular Transformation by a Marek's Disease Virus Oncoprotein, Meq. J Virol. 2013 Aug 15;87(16):9016–29.
66. Peng Q, Zeng M, Bhuiyan ZA, Ubukata E, Tanaka A, Nonoyama M, et al. Isolation and characterization of Marek's disease virus (MDV) cDNAs mapping to the BamHI-I2, BamHI-Q2, and BamHI-L fragments of the MDV genome from lymphoblastoid cells transformed and persistently infected with MDV. Virology. 1995 Nov 10;213(2):590–9.
67. Peng Q, Shirazi Y. Characterization of the protein product encoded by a splicing variant of the Marek's disease virus Eco-Q gene (Meq). Virology. 1996 Dec 1;226(1):77–82.
68. Fragnet L, Blasco MA, Klapper W, Rasschaert D. The RNA Subunit of Telomerase Is Encoded by Marek's Disease Virus. J Virol. 2003 May 15;77(10):5985–96.
69. Fragnet L, Kut E, Rasschaert D. Comparative functional study of the viral telomerase RNA based on natural mutations. J Biol Chem. 2005 Jun 24;280(25):23502–15.
70. Cortes PL, Cardona CJ. Pathogenesis of a Marek's disease virus mutant lacking vIL-8 in resistant and susceptible chickens. Avian Dis. 2004 Mar;48(1):50–60.
71. Cui X, Lee LF, Reed WM, Kung H-J, Reddy SM. Marek's disease virus-encoded vIL-8 gene is involved in early cytolytic infection but dispensable for establishment of latency. J Virol. 2004 May;78(9):4753–60.
72. Cui X, Lee LF, Hunt HD, Reed WM, Lupiani B, Reddy SM. A Marek's disease virus vIL-8 deletion mutant has attenuated virulence and confers protection against challenge with a very virulent plus strain. Avian Dis. 2005 Jun;49(2):199–206.
73. Parcells MS, Lin SF, Dienglewicz RL, Majerciak V, Robinson DR, Chen HC, et al. Marek's disease virus (MDV) encodes an interleukin-8 homolog (vIL-8): characterization of the vIL-8 protein and a vIL-8 deletion mutant MDV. J Virol. 2001 Jun;75(11):5159–73.
74. Jarosinski K, Kattenhorn L, Kaufer B, Ploegh H, Osterrieder N. A herpesvirus ubiquitin-specific protease is critical for efficient T cell lymphoma formation. Proc Natl Acad Sci U S A. 2007 Dec 11;104(50):20025–30.
75. Burnside J, Bernberg E, Anderson A, Lu C, Meyers BC, Green PJ, et al. Marek's disease virus encodes MicroRNAs that map to meq and the latency-associated transcript. J Virol. 2006 Sep;80(17):8778–86.
76. Kaufer BB, Jarosinski KW, Osterrieder N. Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. J Exp Med. 2011 Mar 14;208(3):605–15.

77. Lee LF, Wu P, Sui D, Ren D, Kamil J, Kung HJ, et al. The complete unique long sequence and the overall genomic organization of the GA strain of Marek's disease virus. Proc Natl Acad Sci U S A. 2000 May 23;97(11):6091–6.
78. Tulman ER, Afonso CL, Lu Z, Zsak L, Rock DL, Kutish GF. The Genome of a Very Virulent Marek's Disease Virus. J Virol. 2000 Sep 1;74(17):7980–8.
79. Raab-Traub N, Flynn K. The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation. Cell. 1986;47(6):883–9.
80. Robinson CM, Cheng HH, Delany ME. Marek's disease virus and chicken host genome interactions: Viral genome integration occurs early post-infection and over a timeframe associated with latency, yet integration alone is not sufficient for cellular transformation. Submitted. 2014;
81. Liu H-C, Kung H-J, Fulton JE, Morgan RW, Cheng HH. Growth hormone interacts with the Marek's disease virus SORF2 protein and is associated with disease resistance in chicken. Proc Natl Acad Sci. 2001;98(16):9203–8.
82. Liu H-C, Niikura M, Fulton JE, Cheng HH. Identification of chicken lymphocyte antigen 6 complex, locus E (<i>LY6E</i>, alias <i>SCA2</i>) as a putative Marek’s disease resistance gene via a virus-host protein interaction screen. Cytogenet Genome Res. 2003;102(1-4):304–8.
83. Mao W, Hunt HD, Cheng HH. Cloning and functional characterization of chicken stem cell antigen 2. Dev Comp Immunol. 2010 Mar;34(3):360–8.
84. Niikura M, Liu HC, Dodgson JB, Cheng HH. A comprehensive screen for chicken proteins that interact with proteins unique to virulent strains of Marek's disease virus. Poult Sci. 2004;83(7):1117–23.
85. Stone HA. Use of Highly Inbred Chickens in Research. Washington, D. C.: United States Departmnet of Agriculture; 1975 Jul p. 1–22. Report No.: 1514.
86. Bacon LD, Hunt HD, Cheng HH. A review of the development of chicken lines to resolve genes determining resistance to diseases. Poult Sci. 2000;79(8):1082–93.
87. Briles WE, Stone HA, Cole RK. Marek's disease: effects of B histocompatibility alloalleles in resistant and susceptible chicken lines. Science. 1977 Jan 14;195(4274):193–5.
88. Laurie CC, Laurie CA, Rice K, Doheny KF, Zelnick LR, McHugh CP, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. Nat Genet. 2012 May 6;44(6):642–50.
89. Jacobs KB, Yeager M, Zhou W, Wacholder S, Wang Z, Rodriguez-Santiago B, et al. Detectable clonal mosaicism and its relationship to aging and cancer. Nat Genet. 2012 May 6;44(6):651–8.
90. Maynard J, Haigh J. The hitch-hiking effect of a favourable gene. Genet Res. 2007 Dec;89(5-6):391.
91. Liu H-C, Cheng HH, 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(6):351–9.
92. Morgan RW, Sofer L, Anderson AS, Bernberg EL, Cui J, Burnside J. Induction of host gene expression following infection of chicken embryo fibroblasts with oncogenic Marek's disease virus. J Virol. 2001 Jan;75(1):533–9.

93. MacEachern S, Muir WM, Crosby SD, Cheng HH. Genome-Wide Identification and Quantification of cis- and trans-Regulated Genes Responding to Marek's Disease Virus Infection via Analysis of Allele-Specific Expression. Front Genet [Internet]. 2012 [cited 2014 Jul 26];2. Available from: http://www.frontiersin.org/Journal/10.3389/fgene.2011.00113/full
94. Perumbakkam S, Muir WM, Black-Pyrkosz A, Okimoto R, Cheng HH. 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 Genomics. 2013;14:64.
95. Sackton TB, Lazzaro BP, Clark AG. Genotype and Gene Expression Associations with Immune Function in Drosophila. Begun DJ, editor. PLoS Genet. 2010 Jan 8;6(1):e1000797.
96. Cheng HH, Dunn JR. Personal Communcation. 2014.
97. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010 Oct;11(10):685–96.
98. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug 1;30(15):2114–20.
99. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012 Mar;7(3):562–78.
100. Lassmann T, Hayashizaki Y, Daub CO. SAMStat: monitoring biases in next generation sequencing data. Bioinforma Oxf Engl. 2011 Jan 1;27(1):130–1.
101. Ding L, Wendl MC, McMichael JF, Raphael BJ. Expanding the computational toolbox for mining cancer genomes. Nat Rev Genet. 2014 Aug;15(8):556–70.
102. Larson DE, Harris CC, Chen K, Koboldt DC, Abbott TE, Dooling DJ, et al. SomaticSniper: identification of somatic point mutations in whole genome sequencing data. Bioinformatics. 2012 Feb 1;28(3):311–7.
103. Edmonson MN, Zhang J, Yan C, Finney RP, Meerzaman DM, Buetow KH. Bambino: a variant detector and alignment viewer for next-generation sequencing data in the SAM/BAM format. Bioinforma Oxf Engl. 2011 Mar 15;27(6):865–6.
104. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. 2008 Nov;18(11):1851–8.
105. Lunter G. Probabilistic whole-genome alignments reveal high indel rates in the human and mouse genomes. Bioinforma Oxf Engl. 2007 Jul 1;23(13):i289–296.
106. Cartwright RA. Problems and solutions for estimating indel rates and length distributions. Mol Biol Evol. 2009 Feb;26(2):473–80.
107. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.
108. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;20(9):1297–303.
109. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinforma Oxf Engl. 2009 Nov 1;25(21):2865–71.

110. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. Genetics. 1994 Nov;138(3):963–71.
111. Crooijmans RPMA, Fife MS, Fitzgerald TW, Strickland S, Cheng HH, Kaiser P, et al. Large scale variation in DNA copy number in chicken breeds. BMC Genomics. 2013;14:398.
112. Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat Methods. 2009 Sep;6(9):677–81.
113. Wang J, Mullighan CG, Easton J, Roberts S, Heatley SL, Ma J, et al. CREST maps somatic structural variation in cancer genomes with base-pair resolution. Nat Methods. 2011 Aug;8(8):652–4.
114. Sindi SS, Onal S, Peng LC, Wu H-T, Raphael BJ. An integrative probabilistic model for identification of structural variation in sequencing data. Genome Biol. 2012;13(3):R22.
115. Handsaker RE, Korn JM, Nemesh J, McCarroll SA. Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. Nat Genet. 2011 Mar;43(3):269–76.
116. Kim YK, Bae G-U, Kang JK, Park JW, Lee EK, Lee HY, et al. Cooperation of H2O2-mediated ERK activation with Smad pathway in TGF-beta1 induction of p21WAF1/Cip1. Cell Signal. 2006 Feb;18(2):236–43.
117. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MGF, et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. PLoS Comput Biol. 2011 May;7(5):e1001138.
118. Wang K, Singh D, Zeng Z, Coleman SJ, Huang Y, Savich GL, et al. MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. Nucleic Acids Res. 2010 Oct;38(18):e178.
119. Iyer MK, Chinnaiyan AM, Maher CA. ChimeraScan: a tool for identifying chimeric transcription in sequencing data. Bioinforma Oxf Engl. 2011 Oct 15;27(20):2903–4.
120. Chen K, Wallis JW, Kandoth C, Kalicki-Veizer JM, Mungall KL, Mungall AJ, et al. BreakFusion: targeted assembly-based identification of gene fusions in whole transcriptome paired-end sequencing data. Bioinforma Oxf Engl. 2012 Jul 15;28(14):1923–4.
121. Ng PC, Henikoff S. Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet. 2006;7:61–80.
122. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010 Apr;7(4):248–9.
123. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.