Chapter 1. Application of functional genomics for discovery of new health products

1.1 Introduction

The availability of extensive genomic information has revolutionised the study of biology, with profound effects on understanding the basis of many human diseases and in developing screening, diagnostic, and treatment strategies. Complete genome sequencing of experimental model organisms such as the mouse has also allowed rapid progress in both basic and applied biology. Our goal in establishing this project was to bring some of the benefits of these great advances in genome science to the study of chicken biology. Specifically, we wanted to use genomics based technologies to advance our understanding of how the chicken responds to attack by pathogens. By combining capabilities in genomics and immunology we aimed to advance our knowledge of the host-pathogen interaction with a view to identifying new opportunities to intervene in the process. Such intervention strategies, for example using cytokines or soluble receptor mimics, would have the potential for development as new health products for the industry.

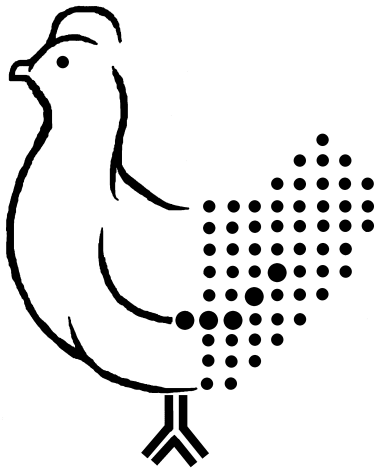
The publication of the first draft of the chicken genome in 2005, the second year of this project, was of great benefit to us and allowed us to quickly develop whole genome microarrays suitable for measuring the expression of essentially all genes.

1.2 Establishment of a National Centre for Poultry Genomics

During the course of this project we have established and developed a facility which we have called the Australian Centre for Poultry Immunogenomics. This facility provides cutting-edge capabilities in genomic technologies applied to chicken research. Its core capabilities revolve around the analysis of gene expression by microarray and real-time PCR technologies. The development of this core capability has required a range of resources and research inputs:

- a suite of advanced equipment
- ongoing development of the probe set used to construct microarrays
- assessment and optimisation of microarray printing and microarray formats
- method development for RNA extraction, labelling and microarray probing

- bioinformatics expertise for data extraction and analysis
- refinement of gene annotations to improve data interpretation
- functional analysis of gene ontologies and pathway mapping
- experimental design



The rationale for developing this capability is that gene expression is a fundamental aspect of organismal development and response that can be quantitatively assessed.

The structure of an organism is defined by (a) the cells that make up each organ and (b) how those cells signal and interact with each other. The structure of a cell is in turn driven by the complement of proteins that make up the cell and direct its formation, functional properties and response to environmental changes and assaults from external agents such as pathogens. All the proteins expressed within a cell are encoded by the organisms DNA. The DNA is transcribed to give an RNA copy which is then translated to produce the cellular proteins. By measuring the amount of each RNA produced by a cell we have a surrogate measurement of the proteins that are being actively synthesised by the cell or tissue. It is, in principle, fairly straightforward to measure all RNA levels whereas the measurement of all protein levels is technically very challenging and difficult to implement on a large scale, particularly for relatively poorly characterised organisms. Therefore, gene expression analysis is the method we chose to implement in our studies of host-pathogen interactions.

By comparing RNA derived from cells, tissues, or whole organisms subjected to different conditions, for example comparing normal cells to cells infected with a virus, it is possible to build up a picture of

how the cells respond to changing conditions. By understanding how cells respond to particular stimuli and challenges we are in a better position to discover, develop and test intervention strategies. In this way we anticipate that this technological capability will inform our efforts to improve the health and productivity of poultry.

1.3 Equipment

CSIRO has made a major investment (~$2 million) in equipping the centre, with the initial purchase of a microarray scanner, microarray printer, colony picking robot and PCR machines. Subsequent investments have added a sophisticated liquid handling robot, further standard PCR machines and multiple real-time PCR instruments plus other ancillary equipment. A next generation, high-throughput genome sequencer, already used in a number of CRC projects, has been added to the equipment repertoire. Most recently a new high-resolution microarray scanner, needed for data capture from the latest microarray format, has been added to the resources of the centre.

1.4 Probe set development

Over the life of the project we have endeavoured to source and maintain the best microarray technologies available (Table 1-1). The development of these chicken microarray resources has been largely funded by direct support from the CRC. We initially started with a library of cDNA clones focused on genes expressed within the lymphoid tissues. This library of 2880 clones was sourced from Dr Joan Burnside (University of Delaware, USA). Production of microarrays from this resource was very complex as clones had to be grown up, plasmid DNA extracted and then PCR products produced for spotting on to microarrays. This set provided a basic resource and gave us the opportunity to learn and establish the basic technologies required for microarray production and use.

Such tissue or functionally focused microarrays do not fully exploit the potential of microarray analysis. One of the great strengths of microarray technology is that it is as easy to probe the expression of 10,000 genes as it is to probe 100 genes. Therefore, rather than predicting which genes might be important and only assaying those it is just as easy to populate the microarray with as much genetic content as possible. We extended the power of the microarrays by next buying in a set of 13,000 cDNA probes prepared by the Roslin Institute. This set was essentially a complete set of all genes that had, at that stage, been identified in chickens.

The publication of the genome sequence of the chicken revealed the complete gene set encoded within chicken DNA(http://www.ncbi.nlm.nih.gov/genome/guide/chicken/). This allowed the development of a whole genome microarray that was designed and produced as a collaborative effort amongst various leading chicken research laboratories around the world. The design was completed by the Roslin Institute with input on gene content from other groups. We provided sequences for genes of particular interest to us plus a series of microRNA sequences that we wanted to include on the array. This set of array probes took on a new format, rather than using PCR products derived from cDNA clones, the new 20,000 element microarray was constructed using chemically synthesised long oligonucleotides. Such a format allows greater quality control and made it easier to add new content as our knowledge of the genome increased.

There were considerable technical challenges in printing microarrays using the cDNA and long oligonucleotide resources and the printing equipment that we had so the final version of the arrays, that we have been using most recently, have utilised the great advances made in commercial array printing. We first used commercially synthesised arrays designed by us and produced by CombiMatrix. These arrays were of high quality but had limited capacity to increase the number of gene probes on the array. We have subsequently changed to NimbleGen arrays which are of high quality and have much higher content. We have been able to custom design these commercially produced arrays and populate each array with 140,000 probes meaning that each array can have a good degree of replication of particular probes and multiple probes per gene.

Table 1-1. Development of microarray resources.

Array Name	Probe type	No. of probes	Source	Advantages
Lymphoid array	PCR of cDNA	2,880	Burnside	First available
Genome array	PCR of cDNA	13,000	Roslin Institute	All cloned genes
Whole Genome Array	Long oligonucleotide	20,000	Operon	Content of whole chicken genome
Combi-Chicken	Long oligonucleotide (in situ synthesis)	21,000	CombiMatrix	Higher quality arrays
Nimble-Chicken	Long oligonucleotide (in situ synthesis)	140,000	NimbleGen (Roche)	High quality arrays, large number of probes

1.5 Optimisation of microarray printing

The printing of microarrays using the various probe sets required the testing and assessment of a number of parameters. We determined the optimal DNA concentrations and buffer conditions to use and assessed the performance of a large number of microarray substrates sourced from a range of producers around the world. The commercial microarray substrates varied in optical quality and surface flatness and had a variety of different coatings to promote the attachment of the DNA probes. We also tested the production of our own microarray substrates using a number of published methods. We concluded that the UltraGAPS II slide manufactured by Corning was the best substrate for our purposes. During the course of the project we also used a number of different pin formats in the microarray printer and found that Point Technologies PT2500 pins had advantages in terms of consistency of spot morphology and printing speed and capacity.

1.6 Method development for RNA extraction, labelling and microarray probing

We evaluated a number of published methods and commercial kits for RNA isolation. Most methods and kits worked reasonably well and it was clear that the biggest cause of variation was the state of the tissue that was sampled, the type of tissue and the way in which tissue was stored before processing. The highest quality RNA preparations are generally obtained from fresh tissue samples that are immediately processed. However, immediate processing is often not possible and so tissue samples must be stored. Storage at 4^{o}C or -20^{o}C is really not effective but storage at -80^{o}C can be effective. In most cases we have used RNAlater to stabilise tissue; this works by salting out proteins, thus removing RNase activity. In general such samples can be stored at 4^{o}C for extended periods of time as long as the tissue is cut into small pieces (e.g. 2 mm cubes) to allow rapid infiltration by the high salt solution.

We evaluated a range of methods for producing cDNA from the RNA samples and then labelling the cDNA with fluorescent dyes for detection on the microarray. We used both enzymatic and chemical methods of labelling and tested the performance of Cy-dyes and Alexa-dyes before eventually settling on Cy3 labelling using the ULS cDNA Synthesis and Labelling Kit (Kreatech Technologies). Early arrays used the dual colour format hybridising samples labelled either with Cy3 or Cy5 to the same array. We discarded this approach in favour of single colour/single sample arrays because the dual colour arrays introduced extra problems for array analysis, largely because of differences in the sample labelling efficiency of the two different dyes.

We investigated a variety of techniques and equipment for hybridising and washing arrays in order to optimise the conditions used for each of the different array formats. With our recent move to the commercially printed array formats these aspects are largely fixed by the requirements and recommendations of the commercial platform.

1.7 Bioinformatics for data extraction and analysis

To extract useful data from a hybridised microarray the array must be scanned at an appropriate setting so that maximum signal intensity is achieved without over-exposure and loss of a linear signal response. Individual spots must then be recognised and signal intensity data extracted. There then follows a process of data normalisation, to allow data across multiple arrays to be compared, and analysis to identify gene probes that have detectable signals (above background) which can then be compared across arrays (samples) to define expression patterns. Data analysis is usually directed towards finding genes that have a statistically significant difference in expression level between different treatment groups. At its simplest the output is a list of differentially expressed genes.

For all steps in data extraction and analysis there are a wide variety of different approaches, methods and software packages that can be applied. During the course of this project we have evaluated many methods and software packages, both community developed freeware packages and commercial software. For the majority of the project we have used our in-house developed methods for data normalisation and the commercial software package, GeneSpring, for data analysis. More recently we have changed to Genowiz for data analysis. Genowiz is a less expensive package and the program developers are much more responsive and willing to address our specific needs.

1.8 Refinement of gene annotations

In order to interpret the lists of differentially expressed genes that are the primary output of the initial microarray analysis it is useful to have an understanding of the potential role of each gene in cellular functions. As the size of the probe set that we were using on the microarrays increased there was an increase in the number of genes for which there was little or no meaningful annotation. A genes annotation gives an indication of the function/activity of the encoded protein. Within the gene databases that have been developed over the last 30 years there are many genes that have been identified based on the cloning and characterisation of mRNA sequences but for which the protein that they are predicted to encode is of completely unknown function. Hence, such genes are of limited

value in interpreting experimental results. For the first genome array that we used only about 30% of the genes represented on the array had a meaningful annotation. A similar lack of annotation was found with the whole genome array when first released. We, and the wider poultry research community, have put a good deal of effort into refining the annotation of the chicken gene set. There is now a level of useful annotation for approximately 75% of the gene content. The vast majority of this annotation has been done by homology to the mouse and human genomes so there are still relatively few genes that are annotated in chicken due to direct experimental evidence. This major increase in the degree of annotation makes a significant difference to the interpretation of the microarray results.

1.9 Gene ontologies and pathway mapping

Gene annotations are most useful when they allow the placement of the encoded gene product into gene ontologies. At its highest level gene ontology describes the functional properties of the gene product in terms of location within the cell, molecular function and the biological process in which it is involved. By understanding these features of a gene product it is more likely that that functional significance of differential gene expression patterns can be interpreted and put into a context that drives forward our understanding of the biological processes that are important in the experiments we are carrying out. Associated with the assignment of gene ontologies is the ability to place the gene products under investigation into their place within biochemical pathways which outline the metabolic and signalling pathways occurring in a cell.

The assignment of gene ontologies is a large ongoing effort amongst a wide community of interest in the genomic sciences (http://www.geneontology.org/). A range of software tools have been developed to allow the assignment, analysis, and presentation of gene ontologies and biochemical pathways. We have investigated most of these tools and found them to be of variable relevance and utility for use with chicken genomic information. Many of the tools are specifically designed for use with particular species and cannot be easily adapted for use on other species. There is only a limited range of tools which can be used with chicken data and we have endeavoured to utilise these and develop our own databases for tracking and using gene ontology information to assist with the functional analysis of chicken gene expression data.

1.10 Experimental design

Over the course of the project our knowledge and understanding of the important elements and considerations that must be taken into account when designing a successful microarray experiment have has evolved and been refined. The three key considerations in designing a microarray experiment are the sample size, the type of sample used and the way these samples are used on the microarray chip.

1.10.1 Sample size

During the course of this project our views on experimental design for microarray analysis have evolved and matured. As with any experimental technique that relies on statistical analysis the sample size is an important consideration. The more samples that can be analysed the more reliable the results are likely to be and the finer the resolution that can be derived from the experiment. However, sample size is limited by the cost of analysis of each sample and sometimes by limitations in terms of the number of samples that can be physically handled through the experimental procedures both before and after sampling for microarray analysis. Our general conclusion is that any one treatment group should provide at least five independent samples, thus giving a reasonable level of biological replication. Technical replicates, in which the same sample is analysed a number of times, can give a guide to the robustness of the methods being used but does not replace the need for biological replicates.

1.10.2 Sample type

Gene expression is a highly dynamic process that can change rapidly over time. For this reason it is essential that appropriate samples are used for microarray analysis. Generation of samples must be very well controlled and assessed. This can present real challenges, for example the chicken's response to a particular pathogen can be quite variable both in terms of the degree and timing of a response, and hence having closely matched samples suitable for analysis can be difficult to achieve. Gene expression changes are an early biological response to a challenge or change in conditions and the key changes in gene expression often occur before any overt changes in the animal/tissue/cell is obvious at a gross level. Therefore, it is generally most useful to take samples early on in the process

under investigation and take more samples then will be needed for the expression analysis so that subsequent biological assays can be carried out to confirm that a response is underway and allow the selection of well matched samples. In general we would advise that at least 10 biological replicates should be taken to generate six samples suitably matched for microarray analysis. It is particularly useful to sample the system under study over an appropriate time course in which key events are anticipated to be happening. The provision of appropriate samples was a particular concern for microarray analysis experiments that we carried out in the early phase of this project. For a number of studies the sampling and analysis that was carried out prior to expression analysis was totally inadequate with the result that little or no useful information could be generated. Experience has taught us that it is not worth expending time and resources on inadequate samples – no matter how precious the samples might be, useful results will not be obtained unless the samples meet our minimum criteria.

1.10.3 Application of samples to microarrays

When we commenced the project the overwhelming number of published studies used two-colour arrays to study gene expression. In this method two samples are separately labelled with different fluorescent dyes and then hybridised to one microarray chip. This methodology was first used to reduce costs by allowing expensive microarrays to be used for two samples. However, as the method became established it was justified more on the basis of giving better experimental results. The theory was that by using two samples on an array a control sample and a test sample could be directly compared. However, it became obvious to us that the two different fluorescent dyes used to label samples behaved differently and so the two samples could not be simply compared; a lot of experimental design considerations and bioinformatic power had to go into allowing for this ”dye bias”. Complex loop-design experiments incorporating dye-swap sample analysis were carried out in order to address the issues associated with using samples labelled with different dyes. We came to the conclusion that the best solution was to move away from two-colour experiments and simplify the process by using a single sample with a single dye on each microarray chip. In this way no allowance had to be made for the different properties of different dyes and, as all samples and arrays were treated in a similar manner, all could be directly compared after data normalisation. This meant that in principle microarrays could be compared not only within a single experiment but across all experiments thus offering significant advantages.

1.10.4 Analysis of host response to infection

Our primary interest in establishing the technological capabilities encapsulated within the Australian Centre for Poultry Immunogenomics was to marry expertise in gene expression analysis with an in depth understanding and capability in immunology to apply to the study of host/pathogen interactions in the chicken.

Our initial plans were to investigate the chicken response in a number of infection models, using both bacterial and viral pathogens. In the first phase of the project we collected samples from disease models using *Mycoplasma* and *Pastuerella* as representative bacterial infections. On the viral side we initially took samples from Marek's Disease Virus (MDV) infection studies and from Chicken Anaemia Virus (CAV) infected birds and we also investigating influenza infected birds.

1.11 Mycoplasma gallisepticum infection

Samples for microarray analysis of *M. gallisepticum* infected tissue were supplied by Phil Markam (Melbourne University). Initially this experiment consisted of 20 samples (10 control and 10 infected) from one time point. Half the birds were infected with the temperature sensitive mutant TS11 and half with the wild-type strain AP3. All the infected and 5 of the control samples were hybridised and subseqent analysis showed that there were only marginal differences between the two groups. After consultation with Phil Markam (Melbourne University) four new samples from two earlier time points were analysed (2 infected samples from both 3dpi and 8dpi). Although the group sizes were small we went ahead and attempted to analyse the data. The number of genes which were identified as differentially expressed under various comparisons are shown in Table 1-2.

Table 1-2. Differentially expressed genes in Mycoplasma infection experiment.

Comparison	**Number of genes regulated > 4-fold**
TS11(day3) vs. TS11 Control(day3)	44
TS11(day7) vs. TS11 Control(day7)	35
TS11(all) vs. TS11 Control(all)	39
AP3(day3) vs. AP3 Control(day3)	225
AP3(day7) vs. AP3 Control(day7)	113
AP3(all) vs. AP3 Control(all)	215
TS11(all) vs. AP3(all)	3755
TS11(all)+AP3(all) vs. TS11&AP3Cont.(all)	15
AP3 Control(day3) vs. TS11 Control(day3)	2114
AP3 Control(day7) vs. TS11 Control(day7)	5019
AP3 Control(all) vs. TS11 Control(all)	4683

The comparison of both the infection groups to the appropriate uninfected control gave reasonable numbers with the indication that the temperature sensitive mutant had little effect on the host tissue whereas the wild-type strain had more pronounced effects. However, the comparison of the two control groups (AP3 Control and TS11 Control) at both time points gave very large numbers of differentially expressed genes even though the groups should be essentially identical. As these numbers make no biologogical sense we did not pursue this analysis further.

This aborted experiment indicated the central importance of ensuring that all samples are taken at appropriate times and under well controlled conditions. It is also apparent that it is of great benefit to have a series of independent biological measures to ensure that the samples are well matched and, in the case of infection studies, that all the test subjects have been adequately infected. One of the major advantages of global microarrays are that they can bring strong statistical power to the analysis of the system under study. That in turn means that samples within a group need to be well correlated and/or there needs to be sufficient group sizes to control for in-group variation.

1.12 Pasteurella multocida infection

Tissue samples from *P. multocida* infection studies carried out by Monash University (John Boyce) were studied by microarray analysis. Total RNA was extracted from liver and spleen samples, transcribed into cDNA, directly labelled with Cy3 and hybridised to chicken whole genome long oligonucleotide arrays. Analysis of the data showed that there was little correlation between samples

of the same treatment groups. Low correlation within treatment groups does not allow a confident analysis of the expression differences between different treatment groups. The low correlation was most likely due to degradation of the RNA sampled from the tissues. To test if the degradation occurred during the isolation and labelling step this process was repeated and gave similar results. This indicated that, although the samples were stored in RNAlater, the RNA in the source material was already degraded. The concern is that during the necropsies it took too long to get samples into the RNAlater buffer to stabilise the RNA content of the tissues. RNA from a second set of samples was isolated but has not been used to probe the latest version of our array platform. This continuing work is awaiting the arrival of our new microarray scanner.

1.13 Marek's Disease Virus infection

We obtained MDV infected tissue samples from a trial conducted by Andrew Bean. The MDV trial was scaled down from the originally planned trial due to the limited availability of SPF eggs at the time. The global gene expression analysis experiment was based around the four time points collected during the time course experiment which was run at the end of 2005. We had extracted RNA samples from tissue samples from all time points of this experiment and six samples from one time point had been hybridised along with the general reference. Unfortunately it became apparent that all the birds sampled were infected with CAV and hence, because of the likely confounding effects of a double viral infection, there was no point in progressing further with the gene expression analysis

1.14 Re-alignment of priorities

The difficulties with certain disease models and the inability to fully control the nature of experiments in terms of the number of samples available, the timing of sampling and the integrity of tissue samples, meant that we re-evaluated our goals and re-aligned our tasks to focus on a successful disease model over which we had greater control. Subsequent work has focused on the analysis of the chicken response to CAV infection.

1.15 CAV infection

CAV, a member of the *Circovidae* family, *Gyrovis* genus (Pringle, 1999), is a non-enveloped, icosahedral virus of about 25 nm in diameter with a negative sense single-stranded circular DNA genome. The CAV genome consists of 2.3 kb with three open reading frames (ORFs) encoding for

VP1, the major structural protein, VP2, a scaffolding protein and VP3, a non-structural protein called apoptin, due to its ability to induce apoptosis (Noteborn, 2004).

Chicken anaemia virus has been recognised since the late 70's as a pathogen that mainly causes clinical disease in chickens prior to 3 weeks of age. The virus is associated with symptoms such as anaemia, lymphoid depletion, haemorrhages and immunosuppression in young birds (Miller and Schat, 2004). Subclinical infection in older chickens does not generally cause anaemia, however it still results in significant economic losses. One of the more important aspects of subclinical CAV infection is its ability to suppress the immune system, resulting in more serious problems when associated with other viruses (Fehler, 2001; Zanella et al., 2001;Davidson et al., 2004). CAV infection has also readily been associated with secondary bacterial infections and vaccine failures (Simionatto et al., 2006). Second to economic losses in the commercial poultry industry, is the prevention of CAV infection in specific-pathogen-free (SPF) flocks. It is imperative that vaccine manufacturers maintain CAV free SPF flocks. Unfortunately, seroconversion of CAV in SPF flocks often happens on or after sexual maturity, resulting in costly losses to the vaccine manufacturers and the SPF industry. Reasons for this late seroconversion are unknown.

A series of experiments have been carried out to investigate the response to CAV infection both in the whole animal and in cultured cells. In vitro studies in cultured cells simplifies the analysis because only a single cell type is present and essentially all cells can be infected, this means that a consistent basal expression level and a consistent response to infection can be expected across the sample. In vivo studies in whole animals are more challenging; only key tissues can be investigated and because there are a variety of different cell types within a tissue and variable levels of infection the overall response measured by gene expression is likely to be more varied and to some extent masked by the natural cellular variation.

1.16 CAV infection of MSB1 cells

CAV infection of cultured MSB1 cells was investigated at two time points post-infection, 24 and 48 hours. Six independent cell cultures (three CAV infected and three controls) per time point were used in single colour microarrays to detect genes that consistently responded to infection with CAV. At 24 h pi 70 % of cells displayed signs of infection and at 48 h pi all cells showed signs of infection and a number of cells had become apoptotic. Control cell cultures at both time points displayed no visible signs of infection.

All microarrays successfully hybridised and passed our quality control measures. A total of 463 and 449 genes were found to be differentially regulated at 24 h pi and 48 h pi respectively by applying a 2-fold cut off and a false discovery rate (FDR) of 0.05 (Morgan et al., 2001). Interestingly only two genes were found to be common across both time points, suggesting that the *in vitro* gene expression response to CAV infection changes markedly over the two time points studied. These rapid changes are further supported by the large number of clusters determined by cluster analysis, which revealed 30 distinct clusters, based on their similar response to infection. In addition, to the 30 clusters over 40 % of the genes were unassigned, signifying a high level of variation in the *in vitro* response between the two time points studied. Several of the genes found to be regulated in this experiment were further analysed by q-PCR and while the absolute fold change in mRNA abundance was not identical to the microarray, directional changes were consistent in 5 out of 6 genes tested.

The major differences between virally infected and uninfected cells were centred on genes involved in inflammation, apoptosis and antiviral activity. A number of cytokine pathways were detected at both points of infection. Interestingly, at 24 h pi a number of genes involved in cytokine pathways are up-regulated in response to infection, whereas at 48 h pi they are all down regulated. At 24 h pi IL-1R, IL-22Rα and IFNα are up-regulated. These two cytokine receptors are implicated in the hosts inflammatory immune response (Schneider et al., 2000; Weining et al., 1998; Kaiser, 2007), and IFNα is known to possess antiviral properties (Staeheli et al., 2001; Kaiser, 2007). Interleukin 1 has been previously linked to CAV infection, showing a reduction in IL-1 production in infected spleen macrophages, attributing this decrease to interference with immune competence at all levels (McConnell et al., 1993). However, here we found an increase in the production of the IL-1 receptor, suggesting an inflammatory response, a known characteristic of the IL-1 family (Kaiser, 2007). At 48 h pi IL-18Rα is up-regulated, whereas IL-2 and IL-7R are down-regulated. IL-18 is a member of the IL-1 family, thus this up-regulation is likely to be involved in an inflammatory response. Both IL-2 and IL-7 have roles in proliferation and development of T cells, consequently, down-regulation of these genes would potentially effect the overall growth of MSB1 cells. In addition, IL-2 is necessary for the development of T cell immunologic memory, one of the unique characteristics of the immune system (Lillehoj et al., 2001). By down-regulating IL-2 it is possible that CAV is circumventing this vital immune function, which may result in immunosupression of the host cells. Not surprisingly, similar antiviral cytokine responses have been shown in a chicken's response to Marek's disease virus, including the induction of IL-1β, IFNα and IL-18 (Kaiser et al., 2003;Heidari et al., 2008).

Chemokines are part of the cytokine family and are defined by their chemotactic activity towards different cells of the immune system (Sick et al., 2000). Chemokines are also known for their role in angiogenesis and at 24 h pi the angiogenic factor K60 is up regulated. This up regulation of K60 may be related to the up regulation of the IL-1 receptor as K60 has been shown to be strongly induced by

the pro-inflammatory cytokine IL-1β (Weining et al., 1998;Introna et al., 1993;Wu et al., 2007). K60 is a member of CXCL class of chemokine which are known for their role in inflammation. At 48 h pi there are three chemokine receptors that are down-regulated including CXCL receptor 5, CCXC receptor 1 and a CC chemokine receptor cluster. Potentially this reduction in chemokine activity may be related to a reduction in the inflammatory response caused by CAV at the later time point of virus infection.

As mentioned earlier only two genes were found to be regulated at both time points studied. The first gene was MHC class IV antigen. MHC class IV is one of three parts of the MHC in chickens and has been implicated in genetic resistance to Marek's disease virus, together with MHC Rfp-Y (Guillemot et al., 1988;Weigend et al., 2001). Both the MHC class IV antigen and MHC Rfp-Y class I α chain are up-regulated at 24 h pi and may play a role in the cells effort to respond the infection. The second gene is CASK interacting protein, this gene is initially down-regulated but is 2-fold up-regulated at the later time point of infection. Up-regulation of CASK has been previously associated with a reduced rate of cell growth (Qi et al., 2005) and may be contributing to a decrease in cell growth at 48 h pi.

Apoptosis has been previously implicated in CAV infection and it is believed that apoptin is one of the elements responsible for inducing the apoptotic pathway (McConnell et al., 1993;Noteborn et al., 1994;Adair, 2000;Schat, 2003;Noteborn, 2004). A previous study investigating the underlying mechanisms of CAV induced apoptosis has suggested that apoptin induced apoptosis is independent of p53 and upstream caspases, reporting that caspase-3 and Bcl-2 are involved but not essential (Noteborn, 2004). In this study p53 was not differentially expressed at either time point, however Bcl-2 like protein and caspase 9 were up-regulated at 24 h pi. Caspase-9 has been shown previously to be involved in the production of caspase-3 (Cho and Choi, 2002), thus it is possible that we have captured an early snapshot of the apoptotic pathway. Interestingly, two other genes that have been previously associated with apoptosis were up-regulated at 24 h pi, being TIA-1 (Forch and Valcarcel, 2001) and CIAPIN1 (Shibayama et al., 2004; Hao et al., 2006). TIA-1 is an RNA-binding domain-containing protein that is suggested to promote apoptosis (Forch and Valcarcel, 2001), whereas CIAPIN1 is predicted to be a cytokine induced anti-apoptosis molecule (Shibayama et al., 2004). It would be useful to investigate these two genes in relation to apoptin and CAV induced apoptosis. Largely, this investigation has corroborated the occurrence of apoptosis induced by CAV infection.

Some of the other genes detected in this experiment that have been influenced by CAV infection include ChT1 thymocyte antigen precursor, T-cell receptor β, genes involved in transcription, Cathepsin L and heat shock proteins. CAV is known to cause cell death and thymic depletion. Down-regulation of ChT1 and T-cell receptor β may be linked to the depletion of thymocytes in the thymus post CAV infection (Adair, 2000; Schat, 2003). Genes involved in transcription were up-regulated at

24 h pi (BHLH-PAS, NR5A2, RUNX2/CBFA1, HSTF1 and TCF-9) and down-regulated at 48 h pi (PBXIA, B2 and POU2F3). Up-regulation of Cathepsin L in response to viral infection has been reported previously in chickens infected with Marek's disease virus (Ruby et al., 2006). At both time points there were heat shock genes (HSP2, HSP25 and HSP47) that displayed reduced expression, these genes are often stress induced (Wang et al., 2002), so it is intriguing as to why they are down regulated. It is possible that these genes are involved in modulation of the inflammatory response (Ruby et al., 2006) therefore, their reduction would allow the virus to induce an inflammatory response.

Six microRNAs were shown to be down regulated 48 h pi, including gga-mir-137 which until now was only a predicted micro RNA in chicken. While microRNAs are thought to play a large role in regulating gene expression (Xu et al., 2006), it is difficult to say how these six micro RNAs are involved in CAV infection, due to the lack of functional information available for these small molecules. However, it is interesting that the viral infection has down regulated all of the microRNAs detected, potentially reducing their control over gene expression in MSB1 cells. Perhaps the down regulation of these microRNAs is linked to CAV's ability to induce immunosuppression.

The clear picture arising from the in vitro infection study is that at 24 h pi many of the immune genes are up-regulated, suggesting the cell is responding well to the virus. However, by 48 h pi many of the immune pathways have been diminished, possibly a reflection of CAV's ability to induce immunosupression.

1.17 CAV in vivo infection

A series of in vivo CAV infection experiments have been carried out. A time course experiment was performed to investigate the kinetics of the chicken response to CAV infection. Thymus samples were taken 7, 14, 21, and 28 days after infection. The extracted and labelled RNA was used to interrogate the cDNA based Genome Array. A large number of gene expression differences, comparing CAV infected birds with time matched uninfected controls, were seen at each time point (Figure 1-1). The most difference in gene expression was seen at day 21 with 1557 genes differentially expressed by a factor of 2-fold or greater. Approximately 5-10% of the differentially expressed genes overlapped at different time points with only 7 being differentially expressed at the first three time points.

Another way to look at the data is to compare the virally infected samples at different time points (Figure 1-2). This is potentially a more complex interaction as the time course of response to viral infection is superimposed on the natural time-course changes that are occurring in the maturing tissue.

One clear conclusion from this analysis is that the day 14 and 21 pi samples are much more closely related to each other than to the other time points. This is in accord with the other biological measures (Chapter 3) which indicate that the peak of thymus damage occurs about 18 days pi.

Samples for microarraying were taken from a second CAV infection experiment (CAV5) and used to interrogate the long oligonucleotide based Whole Genome Array, the same array used to analyse the samples from the in vitro CAV infection experiment. Given that the previous microarray and biological data, indicating the time-course of the infection process, the 21 day pi time point was chosen for detailed analysis. There were 1669 genes differentially expressed by a factor of at least 2-fold ($P = 0.05$) between the CAV infected and uninfected samples (5 birds/group). Of these genes 491 were also found to be differentially expressed in the time course CAV experiment. This represents a very high level of similarity, especially given that different array platforms were used for the two experiments. A number of key genes were found in both experiments including IL-6 and IL-12, Caspase 8/10(apoptosis), along with cell surface markers CD3 and CD80. IL-6 and Caspase are upregulated in the infected thymus whereas IL-12, CD3 and CD80 are down regulated. Direct measurement of the cell surface markers by fluorescence activated cell sorting (FACS) also showed that in the infected birds there were lower percentages of cells carrying these markers. The IL-6 and IL-12 findings were confirmed by quantitative RT-PCR.

The genes that are differentially expressed in the thymus of CAV infected birds are quite diverse with genes in many different biochemical pathways being represented. Some selected pathways which were represented include:

MAPK signalling pathway (20)
TGF-beta signalling pathway (7)
Jak-STAT signalling pathway (8)
Wnt signalling pathway (8)
Cytokine-cytokine receptor interaction (11)
ECM-receptor interaction (10)
Cell adhesion molecules (CAMs) (9)
Neuroactive ligand-receptor interaction (15)
Endocytosis (9)
Regulation of actin cytoskeleton (14)
Cell cycle (13)
Apoptosis (7)
p53 signalling pathway (6)
Focal adhesion (18)
Insulin signalling pathway (6)
Adipocytokine signalling pathway (9)
Toll-like receptor signalling pathway (6)
NOD-like receptor signalling pathway (5)
(The numbers in brackets indicate the number of differentially expressed genes in that pathway)

It can be seen that many different pathways are affected in CAV infected birds, including a preponderance of cellular signalling pathways, apoptosis and cellular architecture. A comparison of the differentially expressed gene lists generated in the in vitro and in vivo infection experiments showed a moderate degree of cross over with 70 genes in common, including genes in the MAPK signalling pathway and genes involved in apoptosis.

In all these microarray studies we find many differentially expressed genes which are not well annotated, that is we have no indication as to the function of the encoded protein. This makes it very difficult to draw much value from this gene set but with our gradual accumulation of data there are some genes in this class which we repeatedly see coming up. Such genes may represent common host responses to infection and some may have utility as potential therapeutic agents. The challenge in the future will be to do extensive meta-analysis of our various data sets to pick the best candidates and devise ways of discovering the function of the genes. Outside of the CRC project we are currently planning to chose some candidate genes for structural analysis, by over expression, purification, crystallization and x-ray diffraction studies, to see if this can give us some clues to function. This is a route by which we may be able to discover new therapeutic agents.

1.18 Use of chicken microarrays for characterisation of gene expression in other bird species

The chicken genome and the microarray resources which we have available are valuable assets which place chicken researchers at a considerable advantage to those working on other species of bird. We investigated whether our whole genome chicken microarray could be used with other bird samples. Our immediate interest was to determine if duck samples could be used as there are important questions to be addressed in terms of the difference in virulence of avian influenza strains in ducks compared to chickens. By understanding the strengths and limitations of cross species microarrays we will be able to elucidate the power of the arrays to address important biological issues in diverse bird species. Cross-species gene-expression comparison is a powerful tool that may alleviate the lack of genomic information for a range of different avian species. By utilising existing microarrays it is possible to study gene expression in closely related species; this technique is termed cross species hybridisation (CSH). CSH has been previously used to investigate gene expression across a number of different species including humans and cattle (Adjaye et al. 2004), humans and primates (Wallace et al. 2007), sheep and cattle (Kijas et al. 2006) and a number of plant species (Bar-Or et al. 2006). CSH highlights the wider unconventional use of microarrays, proving that this technique can be flexible; however one must keep in mind the limits of this application, with regard to species specific genes.

In addition to studying gene expression profiles in closely related species, CSH can be used to explore comparative genomics. Comparative genomics provides an opportunity to ascertain relationships between gene function and location in a range of organisms (Griffin et al. 2008). Moreover, CSH allows insight into conservation of functional elements and the tracing of evolutionary phylogenies by way of comparing both closely and distantly related species. Comparative genomic studies in birds may help to develop detailed genomic information in a wide range of bird species. The significance of such genomic information is highlighted by the potential it offers in the study of important issues such as the recent outbreaks of avian influenza. Research on such diseases is hindered by the lack of genomic information available for many of the avian species known to be capable of infection by the virus. Alleviating this lack of detailed genomic information will inevitably assist in understanding the differences of immune responses in a range of avian species and allow better treatment and control strategies to be implemented.

We collected spleen samples from a range of different bird species, isolated RNA and used our standard techniques to label cDNA and then interogaate the long oligonucleotide Whole Genome Array. All of these avian samples were successfully hybridised to the whole genome chicken microarray and this is demonstrated in the summary statistics in Table 1-3. All arrays passed our quality control measures with the magpie geese samples displaying the least amount of spots higher than one standard deviation above background at 67 %. As expected chicken spleen samples performed the best with 78 % spots passing quality control (it is not expected that hybridising chicken to the whole chicken genome microarray will ever result in 100 % spot hybridisation as not all genes in the genome will be expressed in any one sample). Background median intensities were similar for all arrays except for the duck arrays, where the values were higher. The duck spleen samples also displayed a larger net intensity and background standard deviation than all the other microarrays. The distributions of the $\log_{10}$ transformed normalised spot intensities for all control samples are plotted in Figure 1-3. These plots show similar trends between all samples except the magpie goose and duck arrays. The magpie goose plot shows a shift highlighting the higher number of low intensity spots on these arrays. The duck plot suggests that these samples have a greater number of high intensity spots when compared with chicken. Overall the avian microarrays displayed a significant range of spot intensities across the chicken microarray.

Table 1-3. Summary statistics for avian samples hybridised to the whole genome chicken microarray

	Chicken	Duck	Kookaburra	Tawny frogmouth	Magpie goose	Starling
Spots > 1 standard deviation above background	33,759	32,483	32,659	31,154	28,748	31,390
% of spots above background	78	75	76	72	67	73

To further support the use of the whole chicken genome array for other avian species we specifically investigated the expression levels of a test set of chicken genes with known immunological functions. Each of the avian species provided significant net signal intensity for a variety of the genes investigated. A number of the genes were more highly expressed in the other avian species when compared to chicken. The defensin gallinacin 1 gave significant net intensities for all birds studied, with particularly high values for both duck and tawny frogmouth. A number of the genes including IL-8, IFNγ, IFNλ and caspase 3 gave low or no net intensity signals across all samples. In summary, all avian samples tested here provided significant net signal intensities for a number of the chicken genes with known immunological functions.

A major motivation for studying the cross-species use of the chicken microarrays was to determine if we could use the arrays to investigate the course of viral influenza infections in ducks. The study of viral influenza infections in ducks is of great interest because, although they are easily infected, they are much less severally affected by the virus. Could an understanding of how ducks deal with influenza infections give us clues as to how chickens could be assisted to overcome the deleterious affects of infection?

We carried out a small pilot scale experiment to compare H5N1 infected and uninfected ducks. Five-week-old Pekin ducks were challenged with a Vietnamese H5N1 strain (A/Muscovy duck/Vietnam/453/2004); each dose contained approximately $10^{7.2}$ median egg infectious doses (EID_{50}). Spleen samples were collected 2 days post infection. Infected samples were confirmed using viral titres (data not shown). RNA was recovered and processed through our standard microarray processing pipeline. A total of 2103 genes were differentially expressed ($p = 0.05$), 685 genes were up regulated in the H5N1 infected ducks and a further 244 genes were down regulated when compared to the uninfected controls (above 1.8 fold). The very clear differentiation of the infected and uninfected samples is obvious from the condition tree (Figure 1-4), which placed each group on separate arms of the tree. This condition tree also highlights the large number of regulated genes in the duck samples,

and in particular identified key immune genes IL-1β, GAL4 and MHC class II that are regulated in the H5N1 infected samples.

This pilot-scale trial indicated the value of chicken microarrays for cross-species studies in other avians. It gives us a valuable resource to not only study chicken gene expression but also a useful resource to study gene expression in other bird species which present issues, such as avian influenza, of relevance to the industry. In this regard we have also successfully used the chicken microarray to study the development of the crop in brooding pigeons – this has been done with a view to discovering new biologicals which may have applications in chickens.

1.19 Contributions to other research areas

The genomics capability was developed so that the technology was available to all poultry researchers in Australia, not only for our studies on host-pathogen interactions. The project has contributed directly to other CRC projects and other associated work with a focus on chicken genomics. (See Appendix for publications list).

1.20 Development and biochemistry of the gut

One of the other major areas of application is in the emerging field of nutrigenomics; the use of genomics technologies to study nutrition. We have pursued this area of research and established a capability in this field with the assistance of a PhD project (Barb Konsak) that has been associated with our fundamental genomics work.

The PhD project work has focused on characterising baseline gene expression profiles in the gut to inform us about the biological processes that are occurring and how we should sample the gut to get a good picture of gene expression dynamics. Initial studies focused on the amount of sampling across the gut that was needed in order to characterise spatial differences in gene expression. Because the small intestine is divided into physically distinct regions (duodenum, jejunum and ileum) we were interested to see if there were major differences in expression between sections. We were also concerned that there may be gradients of expression within particular gut sections and so anticipated that it would be important to standardise the exact part of each gut section that was sampled in subsequent experiments. On looking at gene expression across the small intestine we found that each gut section was quite homogeneous and there were only relatively subtle gene expression changes between gut sections (e.g. comparing duodenum and ileum). This means that sampling is relatively simple; no great care needs to be taken to ensure sampling of exactly equivalent segments. In terms of

the implications for understanding the biology of the gut these results suggest that there are few gross functional differences between different regions of the small intestine. The apparent regional specificity that has been assigned to the digestion and absorption of certain nutrients may be more a reflection of changes in composition of the digesta as it is processed along the length of the gut rather than a reflection of the availability of essential enzymes and transporters.

A comparison of gene expression in the duodenum and ileum of 14 day old chickens (functionally mature) revealed 772 genes that were differentially expressed at a level of 2-fold or greater, with only 54 of those genes regulated at greater than 4-fold. Pathway analysis of the differentially expressed genes showed genes in many different categories. Genes involved in calcium and zinc absorption and fatty acid transport were found to be more highly expressed in the duodenum whereas vitamin B12, lipid and bile absorption genes were higher in the ileum. The greater than 4-fold differentially expressed genes of known function are shown in Table 1-4. A large number of genes were expressed similarly between the duodenum and ileum. The vast majority of genes expressed in the gut were expressed at equal levels in the duodenum and ileum; key genes such as those encoding glucose transporters, sucrase isomaltase, apolipoproteins, low density lipoproteins, fatty acid transporters, peptidases and peptide transporters were expressed at low yet similar levels between the duodenum and ileum.

Table 1-4. Genes regulated greater than 4 fold between the duodenum and ileum. Red represents higher gene expression and green represents lower gene expression.

		Accession Number			
Functional Category	**Identification**	**Microarray**	**GenBank**	**D**	**I**
Lipid Metabolism					
	Similar to Glycerol-3-phosphate dehydrogenase	RIGG19803	426881		
Transport					
	Fatty acid binding protein 6, ileal (gastrotropin)	RIGG10528	416154		
	Solute carrier family 25 member 6	RIGG20374	374072		
	Solute carrier family 13 (sodium/sulfate symporters), member 1	RIGG08655	770198		
	Transient receptor potential cation channel subfamily M, member 2	RIGG08244	424835		
Structural					
	Kinesin family member 21B	RIGG10057	421178		
	MAM domain containing glycosylphosphatidylinositol anchor 1	RIGG02174	421431		
	ADAM metallopeptidase with thrombospondin type 1 motif, 6	RIGG09354	427160		
	Similar to kazrin	RIGG02535	426211		
	Similar to cadherin EGF LAG seven-pass G-type receptor 1	RIGG17500	426206		
	Actin, alpha, cardiac muscle 1	RIGG15340	423298		
	Actin, alpha 2, smooth muscle, aorta	RIGG13312	423787		
Protein Metabolism					
	Alanyl (membrane) aminopeptidase	RIGG14412	395667		
Neuroendocrine					
	Neuronal growth regulator 1	RIGG08976	395662		
	Similar to olfactory receptor, family 5, subfamily U member 1	RIGG16585	430778		
Immune/Defense					
	Similar to CD59 protein	RIGG20426	423148		
Other					
	Nucleolar protein 7, 27kDa 98%	RIGG06264	420835		
	GTP binding protein 4 (GTPBP4), mRNA	RIGG04138	420458		
	Synovial sarcoma translocation gene on chromosome 18-like 1	RIGG08178	419225		
	Similar to KIAA0556 protein	RIGG10097	426427		
	Similar to CGI-90 protein	RIGG18792	420213		
	SRY (sex determining region Y)-box 7	RIGG19247	771337		
	Erythrocyte membrane protein band 4.1	RIGG10458	396493		
	Similar to Ac2-059	RIGG07933	420020		
	HECT domain containing 3	RIGG15570	424589		
	Mesoderm induction early response 1, family member 2	RIGG02722	770192		
	Differentially expressed in FDCP 6 homolog (mouse)	RIGG11208	419893		
	Ring finger protein 151	RIGG18846	420226		
	PDZ domain containing 11	RIGG12106	422153		
	Similar to zinc finger RNA binding protein	RIGG10593	427010		
	Quiescin Q6 sulfhydryl oxidase 1	RIGG11955	373914		
	Similar to Zic3 protein	RIGG08378	422251		
	Similar to vacuolar protein sorting 13A isoform B	RIGG18377	427050		
	Poly (ADP-ribose) polymerase family, member 11	RIGG17907	418264		

Analysis of the top 100 highly expressed genes identified genes involved in cell communication, metabolic processes, transport, regulation of biological processes, response to stimulus, catalytic activity, signal transducer activity and binding.

We have also investigated the development of the gut by studying in detail the gene expression profile of the ileum from the immature gut at the day of hatch and comparing it to the functionally mature gut at 14 days of age. There are significant differences in the gene expression profiles at the two time points. Using the cDNA Genome Array we found a total of 7770 genes expressed in the ileum, of which 1364 are only expressed in the immature gut, 2541 were more highly expressed in the immature gut and 583 were more highly expressed in the mature gut. We had commenced this study with the hypothesis that the immature gut would have limited function which would then develop in to more complex functionality during maturation. However, our results tell a very different story with many more genes expressed in the immature gut than in the mature gut. This suggests that the immature gut has broad biochemical potential which, during development, is sculptured and reduced to produce the focused tissue function of the mature gut.

These basic studies on gene expression in the gut were designed to give use experience in collecting and analysing such data, knowledge on how best to sample, and provide a body of information about the general patterns of gene expression in the gut. We have gone on to use this information to study gene expression in the gut in the context of nutritional studies.

1.21 Gene expression in differentially performing birds

We have commenced applying our genomics capabilities to the study of birds with differential performance characteristics, first as part of CRC project 05-02 and more recently in the short strategic project 09-20. Details of this work is detailed in the respective final project reports but in brief we found in 09-20 that 2272 genes were differentially expressed in the jejenum when comparing birds with high and low AME values. These genes were involved in many different metabolic pathways. Key genes involved in metabolism, nutrient and ion transport, growth, gut health and gut hormone action have been identified. No genes were highly differentially expressed (> 10-fold), suggesting that the differences seen are modulating the activity of pathways rather than turning whole pathways on or off. A large number of the genes expressed in the gut encode proteins involved in basic cellular growth and metabolism and have key roles in the ongoing rapid tissue renewal which is a prominent feature of the gut. In future analysis it may be useful to focus in more detail on those genes which encode proteins involved in direct nutrient scavenging and use. This would include sampling and analysing other organs, such as the pancreas, that are producing key digestive enzymes. It will now be important

to investigate how reproducible these findings are and then to determine if any of these gene expression differences play a causal role in determining the efficiency of energy use or whether they are rather an effect of different energy flows. This area of work is showing great promise and should be actively pursued in future work.

1.22 Cataloguing microRNAs of the chicken

Development of the group's expertise in genomics facilitated work to investigate key new regulatory molecules present in the chicken. Within the life of this project it was found that most eukaryotic organisms encode a number of different classes of small non-translated RNAs that play various functions in the genome. The most widely studied of these molecules are the microRNAs (miRNA); 21-23 nucleotide RNAs that play key roles in controlling the translation of messenger RNAs. Because of their central role in controlling the expression of many genes it is important that they are studied as it may reveal new ways to manipulate, control and understand the expression of important agricultural traits. In model systems such as drosophila and the mouse these miRNAs have become very well studied and characterised. When we commenced our studies of miRNAs in the chicken the only chicken miRNAs that had been identified had been found bioinformatically by analysing the genomic sequence and picking predicted miRNAs by homology to those known in mice and humans. We used the capabilities offered by the newly developed Next Generation Sequencing technologies to directly characterise small RNA samples from chickens and hence identify 488 new miRNAs in the chicken. 430 of these new miRNAs appear to be specific for the bird lineage. Together with the 121 previously predicted miRNAs our study brought the total number of known chicken miRNAs to 609 (Glazov et al. 2008).

We are using this newly acquired information to look at the roles of individual miRNAs in important biological processes such as haematopoiesis and sex determination, with the aim of being able to manipulate aspects of these characteristics. We are also doing extensive genomic analysis to try to predict target sequences for miRNA action and to correlate changes in miRNA expression with changes in messenger RNA expression.

1.23 FIGURES:

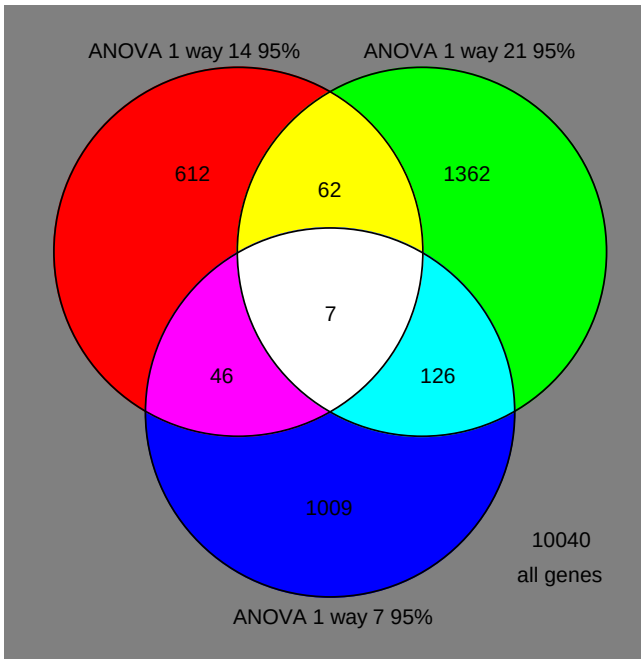


Figure 1-1. Venn diagram detailing, at each time point, the number of genes differentially expressed in the thymus when CAV infected samples were compared to uninfected controls (P = 0.05). For example at day 7 pi (bottom circle) there were a total of 1188 genes that were differentially regulated comparing the viral infected birds with uninfected birds. Of those genes 7 were found to be regulated at each time point (the centre section), 126 were also regulated at day 21 pi and 46 were shared with day 14 pi. 10,040 genes gave good signals but were not differentially expressed at any of the three time points.

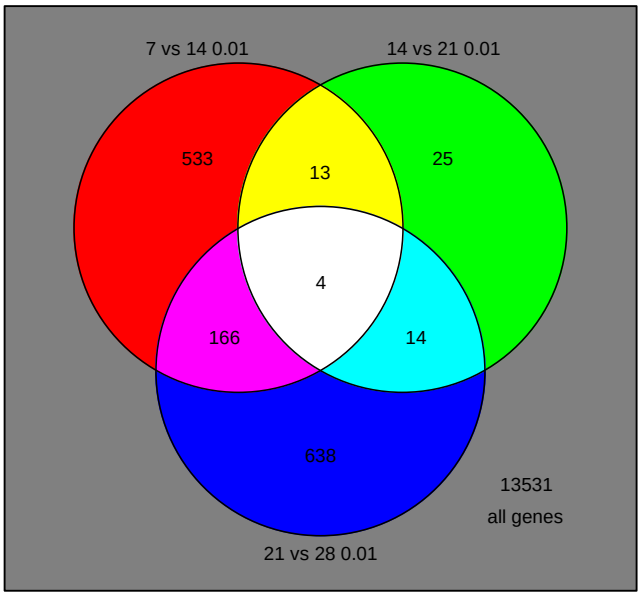


Figure 1-2. Venn diagram showing the number of genes found to be differentially expressed when comparing samples from infected birds at different time points pi.

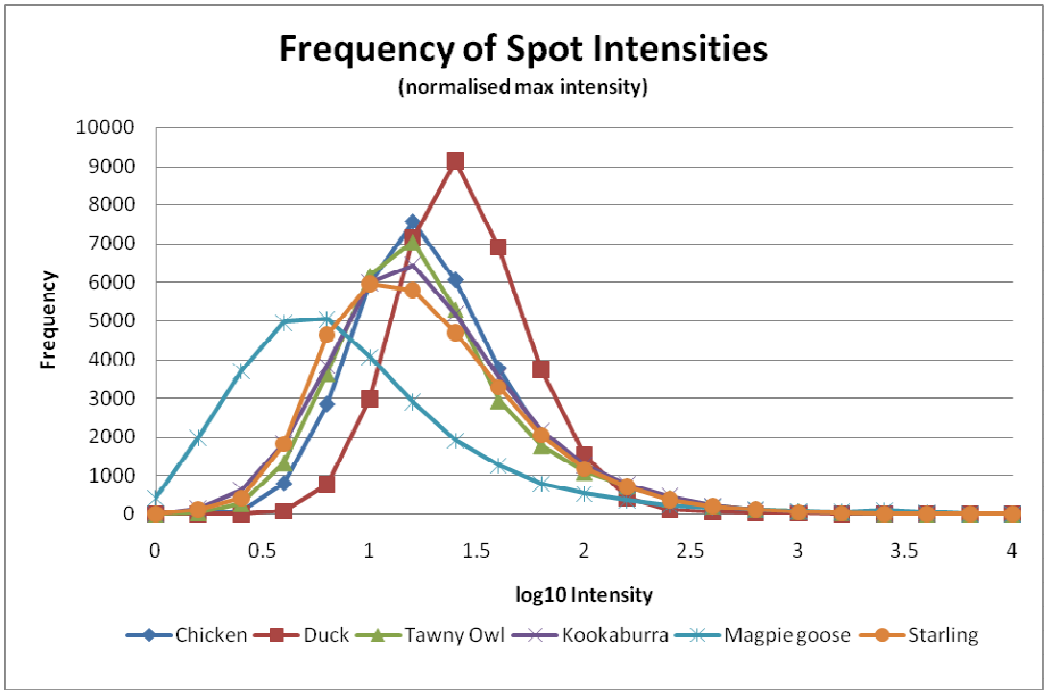


Figure 1-3. Frequency of spot intensity. Frequency and range of average values of net normalised signals for each avian species tested on the whole genome chicken array are shown. Each plot is the average of at least two independent hybridisations.

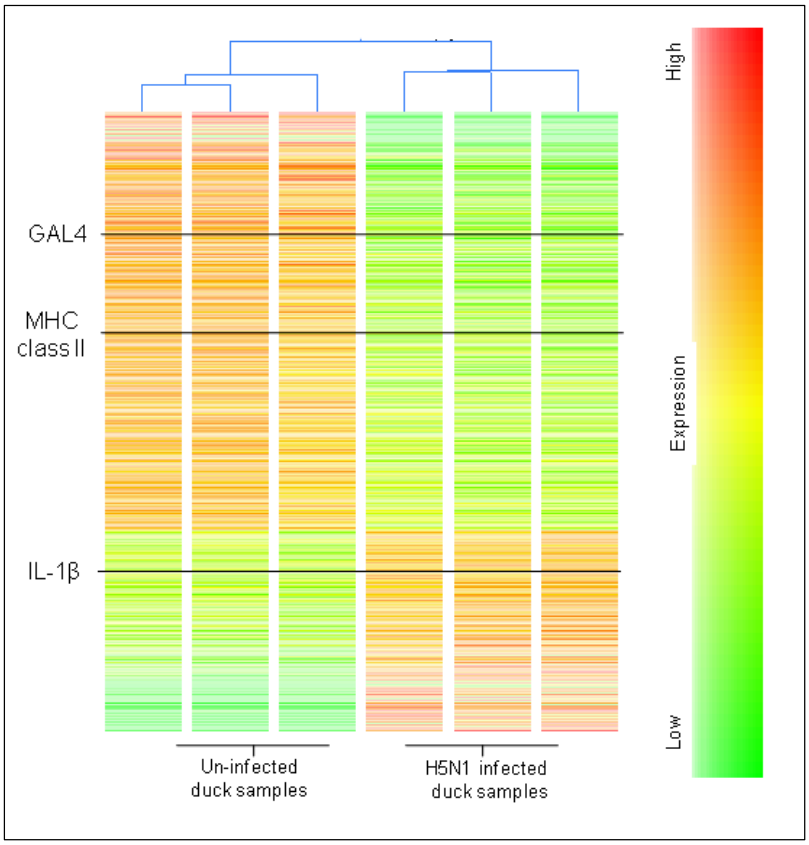


Figure 1-4. Condition tree of H5N1 infected and uninfected ducks.Average signal log intensities are presented, highlighting the strong relationship between the H5N1 infected samples compared to the uninfected duck controls. Immune genes of particular interest are shown in black.

1.24 Discussion

Our major goal was to establish a facility and the technical expertise to apply genomic approaches to the study of chicken biology. This has been achieved; we have a well resourced laboratory and through a long process of experimentation, trial and error, reading, seeking advice from experts and integrating all sources of information we have developed an effective microarray analysis pipeline to investigate gene expression in chickens. We now have a good handle on essential experimental details such as the overall design of effective studies, sample numbers, sample quality, the need for independent biological measurements and the best data analysis methods to use. During the course of the project the core microarray format has been improved and updated to take advantage of the latest technological advances in the field. The publication of the chicken genome sequence in 2005 was a landmark event in apply genomic approaches to chicken studies because, for the first time, it gave us access to essentially all the genetic information encoded in the chicken. This allowed the design and construction of a Whole Genome Microarray to allow the expression monitoring of all genes. The genomic and immunological capabilities are drawn together under the banner of the Australian Centre for Poultry Immunogenomics.

Our first application of our genomic capabilities was directed towards the study of host-pathogen interactions. We collected samples from a number of pathogen infection studies that were being conducted by various teams working within the CRC but we soon discovered that, for various reasons the samples collected were not particularly well suited for expression analysis. In some cases the integrity of mRNA had not been maintained and in other cases the level of infection was variable or the timing of sampling was inappropriate. Many of the most interesting and directly relevant gene expression changes are generally early events in the host's response to pathogen infection and so samples need to be taken early. As a consequence of these issues most of our early studies, although useful learning exercises, did not result in useful analyses. After review we refocused our work and concentrated on the analysis of CAV infections, a model system over which we had greater control.

We conducted studies of both in vitro infected cell cultures and in vivo chicken infections. In vitro infections are useful to study because a single homogeneous cell population can be used leading to very robust results. Our studies were very informative and showed that early in infection the host mounts a strong immune response to the virus but the cell then quickly becomes immunosuppressed, a noted feature of CAV infection. One of the issues of interest to us was how well the in vitro cell culture model reflects the infection that occurs in whole animals. We found a considerable overlap in the genes that were differentially expressed in the two model systems. This informed the direct immunological studies, implying that the in vitro model was a reasonable system to use. The gene

expression analysis gave a broad indication of the kinetics of the host response and indicated when and how intervention strategies such as IFN-λ could be used (Chapter 3).

We have gone on to apply our genomics capabilities to studies on the chicken gut, first doing basic characterisation of gene expression across the gut and during development and then looking at differential gene expression in the small intestine of birds at the extremes of the performance continuum. We found only a moderate number of gene expression differences when comparing duodenum and ileum, indicating that the tissues are quite similar in their functional composition. During the maturation process, from newly hatched chick to a mature gut, the number of genes that are expressed is reduced indicating that initially the tissue has a broader biochemical potential that is sculptured and refined to produce the mature tissue. It will be interesting to see if this time course of development has any correlates with the differences that are seen in gene expression patterns in differentially performing birds, i.e. can some aspects of differential performance be caused by differences in the rate of gut maturation.

The genomic capabilities have also been of more general use, extending out to relevant work in other birds species and in-depth analysis of important aspects of gene control in chickens (miRNA work).

Publications arising from the chicken genomic work

1. Crowley, T.M., Haring, V.R., Burggraaf, S. and Moore, R.J. (2009) Application of whole genome chicken microarrays for avian immunology. BMC Genomics , **10**(Suppl 2), S3.
2. Glazov, E.A., Cottee, P.A., Barris, W.C., Moore, R.J., Dalrymple, B.P., and Tizard, M.L. (2008). A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. Genome Research **18**, 957-964.
3. Tizard, M., Moore, R., Lambeth, L., Lowenthal, J., Doran, T. (2007). Manipulation of small RNAs to modify the chicken transcriptome and direct productivity traits. Cytogenetic and Genome Research, **117**, 158-164.
4. Crowley, T.M. and Moore, R.J. (2006). Microarrays: chipping away at the mysteries of chicken genomics. Poultry Digest vol. Feb/Mar, pp. 18-22,54.
5. Moore, R.J., Doran, T.J., Wise, T.G., Riddell, S., Granger, K., Crowley, T.M., Jenkins, K.A., Karpala, A.J., Bean, A.G.D., and Lowenthal, J.W. (2005). Chicken functional genomics: an overview. Aust. J. Exper. Agric. **45**, 749-756.

A number of other publications are in preparation or have been submitted

1.25 Implications

- We have developed a genomics capability which has contributed to a number of other projects and has ongoing applications in a wide range of projects.

- The genomics work has informed our ongoing studies of CAV infection in chickens and has assisted in identifying and appropriately using and studying IFN-λ as a therapeutic agent.

1.26 Recommendations

- The genomics capacity is a core capability that needs to maintained to service the needs of multiple projects that are anticipated to be supported in the new CRC. This capability is not, in its own right, directed at producing specific commercial outcomes for the industry but is rather an enabling technology used by other more specifically focused projects.

- We see major opportunities in continuing host-pathogen studies, more focused work on monitoring the changes induced by therapeutic treatments such as IFN-λ, nutrigenomic and productivity studies, and in monitoring and understanding the effects of more advanced genomic manipulations such as those envisioned in the chicken sex selection work and our ongoing work to produce avian influenza resistant chickens.

References

Adair, B. M. (2000). Immunopathogenesis of chicken anemia virus infection. *Dev Comp Immunol* **24**, 247-55.

Akiyama, Y. and Kato, S. (1974). Two cell lines from lymphomas of Marek's disease. *Biken J* **17**, 105-16.

Adjaye, J., Herwig, R., Herrmann, D., Wruck, W., Benkahla, A., Brink, T.C., Nowak, M., Carnwath, J.W., Hultschig, C., Niemann, H., Lehrach, H. (2004) Cross-species hybridisation of human and bovine orthologous genes on high density cDNA microarrays. *BMC Genomics* **5**, 83.

Bar-Or, C., Bar-Eyal, M., Gal, T.Z., Kapulnik, Y., Czosnek, H., Koltai, H. (2006) Derivation of species-specific hybridization-like knowledge out of cross-species hybridization results. *BMC Genomics* **7**, 110.

Cho, S. G. and Choi, E. J. (2002). Apoptotic signaling pathways: caspases and stress-activated protein kinases. *J Biochem Mol Biol* **35**, 24-7.

Dar, A., Munir, S., Vishwanathan, S., Manuja, A., Griebel, P., Tikoo, S., Townsend, H., Potter, A., Kapur, V., and Babiuk, L. A. (2005). Transcriptional analysis of avian embryonic tissues following infection with avian infectious bronchitis virus. *Virus Res* **110**, 41-55.

Davidson, I., Shkoda, I., Elkin, N., Ayali, G., Hamzani, E., Kaas, N., Smith, B., Borochovitch, H., Gilat, G., Krispin, H., Kedem, M., and Perk, S. (2004). Chicken Infectious anemia in young broiler flocks in Israel. **59**, 78-82.

Fehler, F. (2001). CAV infection in older chickens: an apathogenic infection? *In* "Proceedings of the Second International Symposium on Infectious Bursal Disease and Chicken Infectious Anaemia" pp. 391-394. Rauischholzhausen, Germany.

Forch, P. and Valcarcel, J. (2001). Molecular mechanisms of gene expression regulation by the apoptosis-promoting protein TIA-1. *Apoptosis* **6**, 463-8.

Griffin, D.K., Robertson, L.B., Tempest, H.G., Vignal, A., Fillon, V., Crooijmans, R.P., Groenen, M.A., Deryusheva, S., Gaginskaya, E., Carre, W., Waddington, D., Talbot, R., Volker, M., Masabanda, J.S., Burt, D.W. (2008) Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *BMC Genomics* **9**,168.

Glazov, E.A., Cottee, P.A., Barris, W.C., Moore, R.J., Dalrymple, B.P., and Tizard, M.L. (2008). A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. Genome Research **18**, 957-964.

Guillemot, F., Billault, A., Pourquie, O., Behar, G., Chausse, A. M., Zoorob, R., Kreibich, G., and Auffray, C. (1988). A molecular map of the chicken major histocompatibility complex: the class II beta genes are closely linked to the class I genes and the nucleolar organizer. *EMBO J* **7**, 2775-85.

Hao, Z., Li, X., Qiao, T., Du, R., Zhang, G., and Fan, D. (2006). Subcellular localization of CIAPIN1. *J Histochem Cytochem* **54**, 1437-44.

Heidari, M., Zhang, H. M., and Sharif, S. (2008). Marek's Disease Virus Induces Th-2 Activity During Cytolytic Infection. *Viral Immunol* **21**, 203-14.

Introna, M., Breviario, F., d'Aniello, E. M., Golay, J., Dejana, E., and Mantovani, A. (1993). IL-1 inducible genes in human umbilical vein endothelial cells. *Eur Heart J* **14 Suppl K**, 78-81.

Kaiser, P. (2007). The avian immune genome a glass half-full or half-empty? *Cytogenet Genome Res* **117** , 221-30.

Kaiser, P., Underwood, G., and Davison, F. (2003). Differential cytokine responses following Marek's disease virus infection of chickens differing in resistance to Marek's disease. *J Virol* **77**, 762-8.

Karaca, G., Anobile, J., Downs, D., Burnside, J., and Schmidt, C. J. (2004). Herpesvirus of turkeys: microarray analysis of host gene responses to infection. *Virology* **318**, 102-11.

Kijas, J.W., Menzies, M., Ingham, A. (2006) Sequence diversity and rates of molecular evolution between sheep and cattle genes. *Animal Genetics*, **37**, 171-4.

Lillehoj, H. S., Min, W., Choi, K. D., Babu, U. S., Burnside, J., Miyamoto, T., Rosenthal, B. M., and Lillehoj, E. P. (2001). Molecular, cellular, and functional characterization of chicken cytokines homologous to mammalian IL-15 and IL-2. *Vet Immunol Immunopathol* **82**, 229-44.

McConnell, C. D., Adair, B. M., and McNulty, M. S. (1993). Effects of chicken anemia virus on cell-mediated immune function in chickens exposed to the virus by a natural route. *Avian Dis* **37**, 366-74.

Miller, M. M. and Schat, K. A. (2004). Chicken infectious anemia virus: an example of the ultimate host-parasite relationship. **48**, 734-745.

Morgan, R. W., Sofer, L., Anderson, A. S., Bernberg, E. L., Cui, J., and Burnside, J. (2001). Induction

of host gene expression following infection of chicken embryo fibroblasts with oncogenic Marek's disease virus. *J Virol* **75**, 533-9.

Munir, S. and Kapur, V. (2003). Regulation of host cell transcriptional physiology by the avian pneumovirus provides key insights into host-pathogen interactions. *J Virol* **77**, 4899-910.

Noteborn, M. H. (2004). Chicken anemia virus induced apoptosis: underlying molecular mechanisms. *Vet Microbiol* **98**, 89-94.

Noteborn, M. H., Todd, D., Verschueren, C. A., de Gauw, H. W., Curran, W. L., Veldkamp, S., Douglas, A. J., McNulty, M. S., van der eb, A., and Koch, G. (1994). A single chicken anemia virus protein induces apoptosis. *J Virol* **68**, 346-51.

Pringle, C. R. (1999). Virus taxonomy at the XIth International congress of Virology, Sydney, Australia, 1999. **144**, 2065-2069.

Qi, J., Su, Y., Sun, R., Zhang, F., Luo, X., Yang, Z., and Luo, X. (2005). CASK inhibits ECV304 cell growth and interacts with Id1. *Biochem Biophys Res Commun* **328**, 517-21.

Rajeevan, M. S., Vernon, S. D., Taysavang, N., and Unger, E. R. (2001). Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *J Mol Diagn* **3**, 26-31.

Ruby, T., Whittaker, C., Withers, D. R., Chelbi-Alix, M. K., Morin, V., Oudin, A., Young, J. R., and Zoorob, R. (2006). Transcriptional profiling reveals a possible role for the timing of the inflammatory response in determining susceptibility to a viral infection. *J Virol* **80**, 9207-16.

Sarson, A. J., Abdul-Careem, M. F., Zhou, H., and Sharif, S. (2006). Transcriptional analysis of host responses to Marek's disease viral infection. *Viral Immunol* **19**, 747-58.

Schat, K. A. (2003). Chicken infectious anemia. *In* "Diseases of Poultry" (Y. M. Saif, H. J. Barnes, A. M. Fadly, J. R. Glisson, L. R. McDougald, and D. E. Swayne, Eds.), pp. 182-202. Iowa State Press, Ames, Iowa.

Schneider, K., Puehler, F., Baeuerle, D., Elvers, S., Staeheli, P., Kaspers, B., and Weining, K. C. (2000). cDNA cloning of biologically active chicken interleukin-18. *J Interferon Cytokine Res* **20**, 879-83.

Shibayama, H., Takai, E., Matsumura, I., Kouno, M., Morii, E., Kitamura, Y., Takeda, J., and Kanakura, Y. (2004). Identification of a cytokine-induced antiapoptotic molecule anamorsin essential for definitive hematopoiesis. *J Exp Med* **199**, 581-92.

Sick, C., Schneider, K., Staeheli, P., and Weining, K. C. (2000). Novel chicken CXC and CC chemokines. *Cytokine* **12**, 181-6.

Simionatto, S., Lima-Rosa, C. A., Binneck, E., Ravazzolo, A. P., and Canal, C. W. (2006). Characterization and phylogenetic analysis of Brazilian chicken anaemia virus. *Virus Genes* **33**, 5-10.

Staeheli, P., Puehler, F., Schneider, K., Gobel, T. W., and Kaspers, B. (2001). Cytokines of birds: conserved functions a largely different look. *J Interferon Cytokine Res* **21**, 993-1010.

Sun, J., Wang, Y., Zong, Y., and Qu, S. (2003). Cloning of chicken anemia virus vp3 gene and apoptosis inductive effect of vp3 gene in vitro. *J Huazhong Univ Sci Technolog Med Sci* **23**, 329-31, 334.

Todd, D., Creelan, J. L., Connor, T. J., Ball, N. W., Scott, A. N., Meehan, B. M., McKenna, G. F., and McNulty, M. S. (2003). Investigation of the unstable attenuation exhibited by a chicken anaemia virus isolate. *Avian Pathol* **32**, 375-82.

Wallace, J.C., Korth, M.J., Paeper, B., Proll, S.C., Thomas, M.J., Magness, C.L., Iadonato, S.P., Nelson, C., Katze, M.G. (2007). High-density rhesus macaque oligonucleotide microarray design using early-stage rhesus genome sequence information and human genome annotations. *BMC Genomics*, **8**, 28.

Wang, J. F., Olson, M. E., Winkfein, R. J., Kulyk, W. M., Wright, J. B., and Hart, D. A. (2002). Molecular and cell biology of porcine HSP47 during wound healing: complete cDNA sequence and regulation of gene expression. *Wound Repair Regen* **10**, 230-40.

Weigend, S., Matthes, S., Solkner, J., and Lamont, S. J. (2001). Resistance to Marek's disease virus in White Leghorn chickens: effects of avian leukosis virus infection genotype, reciprocal mating, and major histocompatibility complex. *Poult Sci* **80**, 1064-72.

Weining, K. C., Sick, C., Kaspers, B., and Staeheli, P. (1998). A chicken homolog of mammalian interleukin-1 beta: cDNA cloning and purification of active recombinant protein. *Eur J Biochem* **258**, 994-1000.

Wu, Y. F., Liu, H. J., Chiou, S. H., and Lee, L. H. (2007). Sequence and phylogenetic analysis of interleukin (IL)-1beta-encoding genes of five avian species and structural and functional homology among these IL-1beta proteins. *Vet Immunol Immunopathol* **116**, 37-46.

Xu, H., Wang, X., Du, Z., and Li, N. (2006). Identification of microRNAs from different tissues of

chicken embryo and adult chicken. *FEBS Lett* **580**, 3610-6.

Zanella, A., Dall'Ara, P., Lavazza, A., Marchi, R., Morena, M. A., Rampin, T., Sironi, G., and Poli, G. (2001). Interaction between Marek's disease and chicken infectious anemia virus. *In* "Current progress on Marek's disease research" (K. A. Schat, R. M. Morgan, P. M., and J. L. Spencer, Eds.), pp. 11-19. American Association of Avian Pathologists, Kennett Square, PA, Montreal.

Zhuang, S. M., Landegent, J. E., Verschueren, C. A., Falkenburg, J. H., van Ormondt, H., van der Eb, A. J., and Noteborn, M. H. (1995). Apoptin, a protein encoded by chicken anemia virus, induces cell death in various human hematologic malignant cells in vitro. **9 Suppl 1**, S118-20.