Insights into the molecular basis of wooden breast based on comparative analysis of fast- and

slow-growth broilers

Authors: Shawna Marie Hubert[1], Travis J. Williams[1], Giridhar Athrey*[1,2]

1. Department of Poultry Science, Texas A&M University, College Station, Texas 77845

2. Faculty of Ecology & Evolutionary Biology, Texas A&M University, College Station, Texas 77845

Corresponding author contact: giri.athrey@tamu.edu

Abstract

Wooden breast has emerged as an important condition in the poultry industry and affects the breast muscle of commercial meat-type chickens (broilers). Thus far, the condition has been classified as a myopathy, confirmed to not have an infectious origin, with molecular data showing the muscle to be under oxidative stress. The objective in this study was to query the functional origins of wooden breast and reveal its molecular similarities to known conditions based on pathway analyses. To carry out an in-depth comparative analysis, we generated RNAseq data from wooden breast affected birds and incorporated breast specific transcriptomic data from previously published studies. The comparative datasets were constructed from a range of commercial fast-growth and slow-growth varieties. Analysis of high-impact variants identified from transcriptome data provided a list of genes important in cell signaling, cell proliferation, cytoskeletal development, and calcium metabolism as being affected by nucleotide changes. Overlaying the lists of significantly differentially expressed genes with the list of high-impact variants produced a list of twenty genes that suggest an association of mechanistic and functional causes for woody breast; supporting a polygenic basis for this condition. Our study demonstrates that wooden breast shows an age-dependent gene expression pattern, with pathway analysis showing enrichment of glycolysis, cell differentiation, tumor suppression and inhibition, further indicating a complex condition with few similarities to myopathies. In summary, our results indicate the existence of a mechanistic, heritable basis for wooden breast, the drivers of which deserve more in-depth investigation. Additionally, they also suggest wooden breast to be a more complex condition than previously reported, potentially involving other organ systems.

Keywords: chicken, wooden breast, pathway analysis, comparative expression analyses, myopathies, neoplasms

Introduction

The domestic chicken (*Gallus gallus domesticus*) is a major agricultural species and is arguably the most popular source of animal protein around the world. In the United States, chicken breast is the most consumed meat – per capita consumption surpassed 41kgs in 2015 (source: US Poultry) – and the broiler industry has a substantial economic footprint ($30 billion/year). Consumption of poultry has increased in step with human population growth as well as changes in consumption habits [1]. While demand continues to grow, production is under enormous stress due to a variety of disorders (ascites, fatty liver disease) and meat quality issues such as green muscle disease and wooden breast [2–6]. Of these, wooden breast (WB) is the most recent problem that is negatively impacting breast meat quality. WB is a muscle condition categorized as a breast myopathy, causing increased hardness of tissue and reducing meat quality. The frequency of WB has risen steadily over the last five years, being reported globally with reduced consumer acceptability [7] linked to economic losses [8, 9]. WB has become prominent within the last decade and has been reported to affect over 50% of commercial flocks [10, 11], but accurate estimates of global incidence are not known.

Pectoral myopathies are not new in broiler poultry species, and broiler chicken particularly has a well-documented history of dystrophies and myopathies, including pectoral myopathies induced by physical or nutritional stress [12–14]. For example, Siller *et al.* [15] reported deep pectoral myopathy in both turkeys and broiler chicken induced by exercise.

3

However, WB is different from previously studied pectoral myopathies in some meaningful ways; the hallmarks of WB appear to be moderate to severe degenerative necrosis, with varying degrees of interstitial fibrosis. While some of these features are found in other myopathies, the co-occurrence of localized pectoral myopathy with fibrosis and striations has not been observed previously in broilers. WB is often observed in conjunction with white striping (WS), which is characterized by white striations that run parallel to the muscle fibers in the breast [11]. These white striations can resemble marbling and are associated with increased fat content [16]. In the past, some pectoral myopathies have been assigned etiologies ranging from nutritional deficiency (e.g., selenium) to hypoxia or ionophore toxicity [14, 17]. However, these etiologies are not supported in WB. Despite intensive studies of WB, including histopathological analyses, serological studies, dietary interventions, gene expression, and metabolomics studies [10, 18–22], the causative factors of WB remain unknown.

Both WB and WS demonstrate varying degrees of severity and have been identified in multiple commercial varieties. Recent studies have described WB as a polyphasic myodegeneration [11], presenting lymphocytic phlebitis [11, 20]. The only common explanatory factor appears to be the **growth rate of commercial broilers**, with the most severe cases found in the heaviest male birds [19, 23]. Kuttappan *et al*. [3] reported that the rapid growth rate and high-energy diets both increase the incidence of WS. A similar trend is observed with WB. Since 2000, average broiler weights have increased by 3 kg (6.5 lbs), representing a 55% increase (Source: U.S. Poultry). Due to the high value of breast meat in proportion to the total carcass, increasing incidence and severity of WB translates into more significant economic losses [24].

While selective breeding for performance traits (e.g. growth rate, feed efficiency) and advances in nutrition, are primarily responsible for growth rate improvements in broilers, it is not

4

known whether WB, which is associated with growth rate, has a genetic basis. Recent studies have used gene expression (RNAseq) and metabolomics analyses to characterize WB [9, 10, 18], but these investigations have not been informative about the underlying cause(s) of WB. Whereas a previous report suggested low heritability for WB [24], a recent report by Pampouille *et al.* [25] describes the identification of quantitative traits loci (QTL) for WS in high-yield broilers, and further concludes that WS is a polygenic condition, supporting the hypothesis for a strong genetic basis for WB and WS.

In this study, the objective was to utilize comparative analyses to illuminate the basis of WB and to characterize its similarity to other known conditions. This study also evaluated the evidence for the supposition that WB is a muscle myopathy. We addressed this objective by comparative transcriptomic analyses of WB samples against various genotypes/phenotypes, followed by pathway analyses and tests for enrichment of canonical pathways. This study did not specifically focus on molecular features of WB/WS co-occurrence, and hence we do not draw inferences regarding WS. Altogether, our study indicates that a) WB is an age-dependent disorder driven by transcriptional dysregulation in fast-growth broilers, and b) that WB molecular profiles suggest a complex syndrome potentially involving multiple organ systems. These results suggest a genetic basis to WB and emphasize the importance of more in-depth studies of the mechanistic basis of WB. These findings also suggest that WB is a condition with potential consequences for whole organism health.

Materials and Methods

Study design and source of data

In this study, transcriptome data was generated from birds exhibiting WB, which was then analyzed comparatively with data generated for previous broiler gene expression studies and published in the peer-reviewed literature. The details of these samples, with links to original studies, are provided in Table 1. It is important to note that these studies focused on breast tissue-specific gene expression and to our knowledge, WB was not the explicitly stated subject of investigation. However, a subset of these studies used modern commercial broilers. Therefore we cannot be sure that they were unaffected by WB; however based on the high frequency of WB in commercial broiler flocks [21] it is likely that these samples include WB affected individuals.

To determine the common genetic basis for WB as a distinct signal, aside from the breed specific growth and molecular profiles, it is necessary to compare expression profiles across ages and genetic strains. WB has been reported in fast-growth broilers as early as two weeks of age, but with the most dramatic changes in severity occurring in the final three weeks before slaughter [16, 22]. Therefore, we compared gene expression of pectoralis major muscle tissue from fast- and slow-growth broilers. While WB has been reported from all major commercial broiler strains, WB has not been reported in slow-growth and heritage broilers to date.

Secondly, as WB severity has been reported to increase with age and weight, we compared expression profiles among pectoralis muscle tissues from different age categories of both fast- and slow growth. To answer these questions, we used a combination of data generated in-house (reported above) in addition to reusing publicly available sequence data (NCBI Short Read Archive) that matched our analysis criteria. In total, nine publicly available datasets from six previously published studies were included in these comparisons (Table 1). In each instance, we selected studies that generated RNAseq data from the breast tissue, were sequenced on the Illumina platform, and were not from a pathogen challenge experiment. Three datasets were

6

from an environmental ammonia challenge study, using the 42-day old broilers of the Arbor Acres strain (Aviagen, sample prefix ARAC). While these treatments influence their gene expression profiles, our analyses showed that these three sample groups cluster together with other 42-day old fast-growth broilers. The inclusion of data from this experiment did not change the hierarchical clustering and separation of sample transcriptome profiles by performance and age profile; hence we retained all three datasets in further analyses.

Sample Collection and RNA Extraction

Live animal studies and euthanasia procedures performed in-house were approved by the Texas A&M University's Animal Care and Use Committee (assurance number 2016-0065). Breast tissue samples were collected from eight, 42-day old chickens of a high-yielding commercial broiler strain. The birds in the study were from an all-male flock and raised on a three-phase industry standard diet. Two birds were randomly sampled from each of four replicate pens, containing 40 birds each. The breast muscle was palpated to examine birds for hardness, before euthanasia. This approach has been used as a diagnostic method in several recently published studies [10, 26, 27]. Birds were then euthanized by cervical dislocation and dissected for the collection of tissues for genetic analysis. The pectoralis major and pectoralis minor muscles were then examined for gross lesions and hardness of the muscle. Individual samples were classified as either WB+ or WB- based on the observed hardness of breast tissue and the absence of other visible abnormalities. While histological analyses have also been used for WB classification, they are perhaps more applicable for resolution of severity, rather than a diagnostic for presence-absence of WB. Furthermore, histological classification of breast tissue as 'normal' has not been found to be diagnostic of WB at the molecular level [11, 28]. While it

7

has been noted that WB and WS co-occur frequently, this study was focused on WB, and therefore did not specifically classify tissue for WS presence or severity. Of the eight individual birds sampled this way, six birds were classified as severe (WB+) based on palpation and gross lesions, whereas two other samples were less severe cases (WB-). Owing to the weak correlation between physical/histological and molecular markers of WB, all birds sampled in this study were classified into the WB group. Moreover, as the goal was to compare expression patterns across genotypic backgrounds, this grouping allowed better resolution through increased biological replication of the WB group.

Tissue of size approximately $1cm^3$ was excised from the distal portion of the pectoralis major with a scalpel and immediately stored in RNAlater (Ambion Inc). After 24 hours of incubation at 4^0C, the excess RNAlater was removed, and samples were stored at -80^0C until further processing. Total RNA was extracted from about 30mg of the tissue using the RNEasy Mini Kit (Qiagen Inc). Samples were checked for RNA quality and concentration on a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific).

RNA Sequencing and Transcriptome Analysis

Total RNA isolates were submitted for library preparation and RNA-sequencing at the AgriLife Genomics and Bioinformatics Center (Texas A&M University). Sample libraries were prepared by performing DNAse digest, followed by poly-A selection for mRNA molecules. Individual mRNA isolates were then pooled into three sample libraries – namely one library comprising two WB- samples, and two libraries each comprising three WB+ samples. These three sets of pooled samples were used for strand-specific library preparation, and the libraries were sequenced with 125bp single end sequencing.

The 11 datasets, including the in-house generated and downloaded datasets, were then processed identically. In brief, the raw RNAseq data was filtered for adapter contamination and quality trimmed using the program Trimmomatic [29]. Reads with average quality scores less than Q30 and shorter than 20bp in length were discarded. High-quality reads were mapped to the *Gallus gallus* genome (Version 4.8, Ensembl Release 85, July 2016) using the short-read de-novo splice mapper STAR [30, 31], followed by counting of transcripts mapped to the 'exon' features using the tool HTSeq [32]. The counts data for each sample were then compared for statistical significance using the EdgeR package on the R statistical platform [33]. First, low expressed genes across all libraries (CPM <2) were filtered out. Next, normalization factors were calculated for differences in library sizes, followed by estimation of common and then tagwise dispersion (GLM). We used the package COMBAT to check and correct for batch effects [34, 35]. The quasi-likelihood based 'glmQLFTest' function was used to perform tests for significance between expression values among treatments. The QLF approach is known to provide greater protection against Type I error and can handle unbalanced designs better than exact tests. A total of 10 pairwise contrasts were performed between WB data (generated in-house) and downloaded broiler transcriptome data. Following the analyses in EdgeR, topTag tables were used in interpretation and pathway analyses. For each of the ten differential expression results, pathway analyses were performed using the Ingenuity Pathway Analysis platform (Qiagen Inc.). Only genes significant at FDR <0.05, and with Log2FoldChange smaller than -0.5 or greater than 0.5 were included in pathway analyses. Finally, the results from pathway core analyses (10 datasets) were all included in a 'Comparison Analyses' on the IPA platform, to characterize similarity of expression, shared canonical pathways, upstream regulators, and diseases.

Differential expression between Fast-Growth versus Slow-Growth broilers

A secondary differential gene expression analysis was performed by grouping all the modern commercial broilers as the Fast-Growth Commercial Broiler (FGCB), and by grouping the Weichang (WC) and White Recessive Rock (WRR) samples as the Slow-Growth Heritage Broiler (SGHB). The Illinois strain and the hybrid WRR-XH crosses were left out of this comparison as they are neither a heritage breed nor a commercial variety. Differential expression analysis and pathway analysis was performed in the same way as described above.

Variant analysis with RNAseq data

A total of 54 sequence libraries (.fastq), including the eight generated for this study, were used to generate variant calls and to identify shared and unique SNP variants among the breeds included. The Genome Analysis Toolkit [36] best practices pipeline for variant calling from RNAseq data was used to generate a set of high-quality variants for each sample using hard filtering. Briefly, STAR aligned reads (same used for differential expression analysis) were first processed to add read information and to mark duplicates using the tool Picard [37]. Binary alignment files were then fed into HaplotypeCaller, followed by the selection of SNP variants, and variant filtration. SNPs occurring in clusters within 35bp were filtered out, as were variant calls with Qual By Depth (QD) score <5, and Fisher strand bias > 35. The resultant set of high-quality variants obtained this way were passed into the variant effect prediction software SnpEff [38]. Variants annotated as having a 'High' impact modifier by SnpEff were used for comparison among samples. Due to the variability in sequencing library size and depth of coverage, differences in the number of variants were expected. Therefore, to account for potential bias

10

arising from differences in sequencing depth, high-impact variants were compared only between the FGCB group, and the SGHB group. The combined high-impact variant list was generated by pooling all variants by ENSEMBL gene ID and removing duplicates.

Results

Growth rate and age explain global gene expression patterns

Across the ten datasets, a total of 12,202 genes were expressed above threshold (CPM=>2) and were included in both the differential expression and pathway analyses. An ordination analysis using Non-metric Multidimensional Scaling (NMDS, Figure 1) showed that fast-growth breeds (ARAC, ROSS and WR-XH Cross) overlapped each other, with younger (6 and 21-days old) fast-growth broilers being less proximate to WB samples, compared to 42-day old broilers (ARAC), indicating a clear age based expression similarity. The Illinois and Ross breeds (6 and 21-days old) both clustered by age and also by breed, showing clear age-based segregation from 42-day old commercial broilers. The slow-growth breeds (WC and WRR, 120 days and older) formed clusters distinct from the fast-growth broilers. Furthermore, all WB samples formed a tight cluster, validating the observation that birds without obviously visible WB symptoms are, nonetheless, not different at the molecular level. Therefore, birds from the same genetic background may not be suitable as a negative control (Figure 2).

Comparison of pairwise differential expression analyses

Results from EdgeR for each pairwise comparison against WB samples are summarized in Table 2, and mean-average plots for each comparison is shown in Figure 3. The complete list of differentially expressed genes with P-values for all pairwise comparisons is available in

11

supplementary materials. Overall, the WC120+ and WRR120+ slow-growth varieties were most different from WB, based on the total number of differentially expressed genes (~2200). This number was approximately twice as high as any of the other comparisons. After the WC and WRR breeds, the Illinois 21D was most different (1702 genes differentially expressed), and Ross 21D being most similar (693 genes differentially expressed), with other comparisons falling in between the extremes. The three different datasets of 42-day ARAC (3 treatments in the original study), were very similar to each other in their differences to WB (total of 1243, 1167, and 1330 DEG respectively). The top canonical pathways identified were also highly similar, with T-cell receptor signaling in all three comparisons and IL-8 signaling in comparison to both A and C groups of ARAC. Canonical pathways explained by these DE genes showed that IL-8 signaling and T-cell receptor signaling were recurrent terms, but all three comparisons shared TGFB1 and TNF as the upstream regulators.

Comparing WB to Ross 708 (6 and 21-days old) showed more significant differences at six days than at 21-days, with main pathway terms being CD28 Signaling in T-helper Cells and T-cell receptor signaling. Lipopolysaccharide and beta-estradiol were shared, upstream regulators. Unlike the Ross strain, the Illinois 6 and 21-day old birds were distinct in their global expression profile, with more significant differences to WB compared to the younger Ross strain birds. This observation would fit the known biology of the Illinois strain, which is a broiler line with a performance profile from the 1950's. However, pathway terms for these datasets still included CD28 signaling in T-helper Cells and T-cell receptor signaling, which were shared with the Ross strain. Both WC120+ and WRR120+ slow-growth breeds were most different from WB samples in the extent of differential expression, and with no overlap in the top three pathway terms.

Pathway analysis of genes differentially expressed in slow growth varieties also yielded pathway terms that were not shared with other comparisons. Pathway analyses yielded hepatic fibrosis/hepatic satellite cell activation, calcium signaling, eNOS signaling, molecular mechanisms of cancer, NRF2 mediated oxidative stress, ERK/MAPK signaling, signaling by Rho family GTPases, Tec kinase signaling, and PI3K signaling in B lymphocytes as the top terms in ARAC (3 libraries), WRR-XH cross, WC120D+, WRR120D+, Ross 6D and 21D, and Illinois 6D and 21D respectively.. Interestingly, despite the differences in the top canonical pathways identified when comparing WB gene expression to that of slow growth varieties rather than fast-growth varieties, the upstream regulators suggested by this differential gene expression include the same terms. Diseases identified by the differential gene expression of each comparison evaluated included the same terms: Cancer, organismal injury, and gastrointestinal disease, with only one comparison indicating neurological disease. Finally, The frequent occurrence of T-cell receptor signaling and IL-8 Signaling suggest that molecules activated in these pathways may be suitable as biomarkers for detecting WB.

Multisample Comparison Analysis

The multisample comparison analysis allows identification of similarities and trends occurring across multiple datasets, specifically, identification of functions overrepresented across datasets. Pathways are considered significant if a higher number of molecules associated with a pathway are expressed than expected by chance. Based on the pathways with highest -log(P-values) and activation Z-scores, the top shared canonical pathways were T-cell receptor signaling, CD28 signaling in T-helper cells, and signaling by Rho family of GTPases. The top 50 pathway and disease terms are shown in Table 3. The top upstream regulators were TGFb1,

TNF and TP53 and beta-estradiol. The comparison feature also generated a list of the top diseases and disorders; the top three disease terms that emerged from the consensus of the multisample comparison were Cancer, Abdominal Neoplasm, and Solid Malignant Tumor. None of these top 100 disease terms pointed to muscle myopathies or other musculoskeletal disorders. Finally, the top disease signaling pathways were Cancer, MAPK and the P53 pathways.

Fast-Growth versus slow-growth differential expression

For this analysis, commercial broiler strains (ARAC, ROSS, WB) were included in the fast-growth commercial broiler (FGCB) group, and the WC120+ and WRR120+ strains were grouped into the slow-growth heritage broilers (SGHB) group. This particular analysis was designed to separate out those genes that are upregulated or downregulated in FGCB irrespective of age, with the supposition that genes associated with the onset and progression of WB in fast-growth strains would be found across age categories (6-days to 42-days). A total of 11,766 genes were expressed above a threshold (CPM=>2), of which 6406 were differentially expressed (FDR<0.05, LogFC < -0.05 and >0.05). Of the total differentially expressed genes, 6168 genes were significantly downregulated in FGCB. These differentially expressed genes were then analyzed in IPA to identify canonical pathways and diseases/disorders. The top three canonical pathways based on -log(P) values were Mitotic Roles of Polo-like Kinase, ILK Signaling, and ERK/MAPK signaling. As observed with other pairwise comparisons, the main disease terms were Cancer, Solid Malignant Tumors, and Gastrointestinal Disease.

Variant discovery from RNAseq data

14

Variant calling and filtering of the datasets yielded average SNP calls of 79,539, and 83,902 for the FGCB and SGHB groups respectively. The individual variant calls were merged using GATK CombineVariants to generated a consolidated (merged) set of 709,959 and 279,339 high-quality SNPs in FGCB and SGHB respectively. This difference in total variants was driven mainly by differences in the number of samples included in each group - namely 25 and eight for FGCB and SGHB respectively. The multisample VCFs annotated with snpEff to identify effects of the variants and to categorize impacts yielded 395 and 158 high-impact variants in FGCB and SGHB respectively. Overall, the proportions of effected features and functional impacts were evenly matched (Table 4) except where noted. Modifier effects (changes outside coding regions) was the most frequent effect, which was higher in FGCB compared to SGHB. High-impact variants which signify impacts within the coding sequences (e.g., frameshift, stop gained, stop lost), were of equal proportion in both groups, whereas both moderate, and low-impact variants were more frequent in SGHB.

Of the total high-impact variants, 37 were shared among all three fast-growth breeds (ARAC, Ross, WB), whereas 35 were shared among the two slow-growth breeds (WC120+ and WRR120+). Of these 72 total high-impact variants, 14 were found in both the slow and fast-growth breeds, leaving 23 unique high-impact variants in FGCB, and 21 unique high-impact variants in SGHB (Table 5). The membership of both lists is rich in genes involved in cell signaling, cell proliferation, and cellular response to stress (including hypoxia). Particularly notable genes with high-impact variants unique to the FGCB group are SPEG, NPEPPS, and THYN1. These genes are involved in myocyte cytoskeletal development, linked to the cellular response to hypoxia, and associated with the induction of apoptosis, respectively. Notable genes with high-impact variants in the SGHB group were two myosin heavy chain genes (MYH1A,

15

and MYH1B) and dystonin (DST). These genes are involved in motor activity and actin filament binding, and the assembly of collagen fibrils, respectively.

Overlap of high expression and high-impact variants

Genes with high-impact variants in either FGCB or SGHB were cross-referenced against the list of significantly differentially expressed genes between the two groups. Twenty-three of the total 45 genes were also found to be significantly differentially expressed (Table 5). Genes that were up- or down-regulated in FGCB were found in both high-impact variant lists. Interestingly, 17 of the 21 genes with high-impact variants in SGHB were also significantly differentially expressed, suggesting both a mechanistic and functional role for these genes.

Discussion

The hierarchical clustering of the pairwise differential expression analysis and the 100 top pathway terms shared across the 10 comparisons showed that there is a definite age based clustering pattern; all 42-day old broilers (ARAC) cluster together, and based on the Z-score and P-values, are more similar to WB+ tissue in their gene expression profiles and the pathways they activate. On the other hand, younger birds (6 and 21-days old) of slow and fast-growth breeds, and older birds of slow-growth breeds cluster together and separately from the 42-day old broilers. Interestingly, these results show that young broilers (Ross and Illinois 6 and 21-days old) and the slow-growth varieties (WC, WRR strains) appear to have a similar gene expression pattern, in contrast to 42-day old broilers. In summary, the comparison of pathways from multiple pairwise comparisons show that age, first, and then growth rate (broiler strain) are the primary functional differentiators of WB tissue. The differentiation of the 120+ day old slow-

16

growth broilers and 6 and 21-day old fast-growth broilers is especially notable, as they show that a) molecular signatures associated with WB are unique to older, fast-growth broilers, and b) that 21-day old modern broilers (Ross 708) are less similar to 21-days old Illinois breed than to 42-day old commercial broilers. These two points suggest an age-dependent transcriptome dysregulation in WB, which progresses with age, somewhere between the first and third week of life. This conclusion is similar to that reached by Griffin *et al.* [19].

While gene expression and ontology analyses show which specific genes and molecular functions are involved in WB, the design of appropriate remediation strategies requires a better resolution of the similarity of WB to known diseases. A clear understanding of diseases and conditions explained by gene expression patterns is necessary to narrow down specific endogenous as well as environmental factors driving WB. Specifically, we wanted to answer whether WB tissue expression patterns point to muscle myopathies, or whether such patterns are indicative of other conditions. While genes essential in muscle growth and cell differentiation are up-regulated in WB tissue, the totality of expressed genes and pathways show little support for myopathy as the underlying condition.

One important cause for concern is the repeated occurrence of regulators and pathways that suggest neoplastic disorders. Upregulation of glycolysis, which was observed as the primary pathway classifier in the same-background comparison is considered a "near-universal property" of primary and metastatic cancers [39–42]. Oxidative stress and impaired glycolysis can both arise due to mitochondrial dysfunction. Multiple studies have confirmed the transcriptomic [9, 18] and metabolomic signatures of oxidative stress in WB [10], and have also suggested mitochondrial dysfunction [43]. Furthermore, genes essential to glycolysis, angiogenesis, and

17

apoptosis (up-regulated in WB) are transcriptionally regulated in hypoxic conditions [44, 45], that are typical of tissue under oxidative stress.

The individual and comparison pathway analyses provided many of the same terms as being important among comparisons. The top pathways based on activation Z-score were T-cell receptor signaling, CD28 signaling in T-helper cells, and signaling by Rho family of GTPases. CD28 is involved in stimulation of T-cell activation and 23 molecules involved in this pathway were identified in the transcriptome data. CD28 signaling is involved in glucose metabolism, activation of T-cells, and costimulation of immune responses [46]. T-cell receptors bind to antigenic peptides presented by antigen presenting cells and are known to respond to various signal transduction pathways. They may be invoked to regulate cell proliferation, apoptosis, and cytotoxic killing [47, 48]. These processes may be activated in wooden breast in response to apoptosis and necrosis occurring in breast tissue. Finally, the Rho family of GTPases are involved in regulating various processes, including reorganization of the actin cytoskeleton in response to growth factors and cytokines [49–51]. These pathway terms are indicating abnormal expression patterns that together affect cell signaling, cytoskeletal organization, and inflammation have all been identified as features of WB. As it has been previously shown through histological and enzymatic assay that WS/WB does not have an infectious origin [4], these immune responses are likely directed against endogenous cell proliferation and apoptotic processes (resembling neoplasms).

The disease terms from the multigroup comparison analysis in IPA also found the same top conditions like those found in pairwise comparisons. Many of these terms also invoke the digestive system (intestine, colon, liver, abdomen, etc.), which is surprising considering all the analyses were based on differentially expressed genes in the pectoralis major tissue. While

18

pathway analysis relies on over-representation or functional class scoring, they still rely on accurate annotations, cell-specific information, and well-described pathways for the accuracy of results [52, 53]. Therefore, while it is possible that other organs may be involved, validation of that question will depend on additional sampling and wet-lab based approaches. Other locations of organismal injury notwithstanding, these terms still suggest some important syndromes that can be considered very concerning. The major diseases and disorders explained by expression patterns also suggest abnormal cell proliferation and cell signaling mechanisms. The comparative analysis showed that over 80% of the top diseases identified by the pathway analyses indicate tumors, cancers, and neoplastic conditions. Even as some caution is necessary for interpreting pathway analysis for chicken datasets, due to the majority of pathways described being from mammalian models, it has repeatedly been shown that pathway signatures do predict disease outcomes accurately based on shared molecular features [54–57]. The reasons and basis for this similarity of WB to neoplastic disorders deserve further investigation.

Top canonical pathways emerging from the comparison between FGCB and SGHB groups indicated altered activity of multiple serine/threonine kinases including polo-like kinase (Plk), integrin-linked kinase (Ilk), and extracellular signal-regulated kinase/mitogen-activated protein kinase (Erk/Mapk). All three pathways are implicated in the regulation of the cell cycle and cell survival. Specifically, Plk is induced by mitogens and is most abundant during metaphase of mitosis with activities including chromosome segregation, centrosome maturation and spindle assembly [58, 59]. Plk also functions in other stages of mitosis including inactivation of the anaphase-promoting complex and regulation of nuclear envelope breakdown during prophase [58–60]. It has also been shown that depletion of Plk in cancer cells induces apoptosis and it is now considered a high potential target for intervention [58, 60]. Ilk impacts the cell

19

cycle and survival through activation of critical signaling pathways and stimulation of downstream effector proteins, while Erk/Mapk is similarly involved by activating many growth factors, cytokines and transcription factors [61–64]. Erk/Mapk further promotes cell survival by phosphorylating and thus inhibiting the pro-apoptotic protein BAD (Bcl2 associated agonist of cell death) while also inducing the expression of cell survival genes [63]. Finally, of considerable interest is the ability of Ilk to anchor actin filaments to cell-matrix contact sites, regulating changes in cell shape, cell migration, cell adhesion, as well as the ability to phosphorylate myosin in smooth muscle cells [61, 62]. Each of these pathways plays an essential role in the regulation of cell proliferation, maintenance, and death which are all physiological activities that have been identified as perturbed in the WB condition and thus this analysis provides a focused framework for further investigating the underlying mechanisms of the condition.

Molecular Basis of WB

In studies of WB, it has been observed that few live bird or carcass quality variables are accurately predictive of the presence of WB (Athrey et al., unpublished). For example, in replicate flocks of broilers of the same breed raised under identical conditions, no combination of rearing or dietary variables has been found to prevent WB occurrence [22, 65]. Furthermore, the severity of WB varies within the same flock under identical conditions [4]. Based on these observations, it appears likely that nutritional interventions, while perhaps useful for ameliorating severity, may be of limited utility in eliminating wooden breast. On the other hand, these data suggest a genetic basis underlying WB. This hypothesis has received recent support with the identification of QTL for WS [25]. While the relationships between WB and WS are not

fully resolved, the co-occurrence of these two conditions makes it more likely that WB has a genetic basis.

In this study, we identified highly expressed genes that contained high-impact variants. This small subset of genes is involved in critical cell proliferation and signaling functions. The identification of genes with high-impact variants that are also differentially expressed point towards a genetic basis that links the changes at the DNA level to functional expression. Such DNA variants that are associated with expression differences are called Cis-acting regulatory variants and are known to explain a substantial amount of phenotypic variation, as well as having a role in disease etiology [66–68]. In this study, we identified a total of 44 high-impact variants, of which 21 were also significantly differentially expressed genes. While not all of these 21 genes may be directly affecting WB occurrence or severity, their expression patterns and the type of SNP modification make their involvement in WB highly probable. Particularly noteworthy genes identified in this analysis were MYH1A, MYH1B (high-impact variants in SGHB), the pair of which are myosin heavy chain genes. Both of these genes had 'splice acceptor variants' that may result in a different mature mRNA product and protein. These genes are members of a larger group of myosin genes which regulate development and function of avian skeletal muscle [69]. Interestingly, MYH1B was also identified as a candidate gene for white striping in a QTL mapping study by Pampouille *et al.* [25]. Another pair of notable genes with high-impact variants were DST (dystonin) and TMEM108 (Transmembrane protein). DST was modified in SGHB, whereas TMEM108 was modified in FGCB. DST is a cytoskeletal linker protein, which is involved in collagen trimerization and formation of scar tissue following injury [70]. DST is known to interact with TMEM108, a gene that regulates the stability of microtubules, by recruiting TMEM108 for the transport of endosomal vesicles [71]. Mutations in the DST gene

21

have been identified as being responsible for hereditary neuropathy and dystonia (abnormal muscle tone) in mouse models [72]. The high-impact variants and differential expression of these genes may be associated with the rigidity, collagen content, and microscopic features observed in WB.

While the short list of genes identified in this study have a high likelihood of being explanatory of WB, and even as potential diagnostic biomarkers for the condition, we stop short of calling these candidate genes; population-level analyses such as association testing (GWAS) would be necessary to confirm if variants in these genes are causative of WB. However, these findings do lend support for a polygenic basis for WB, but perhaps occurring in conjunction with regulatory mechanisms that are yet to be identified and confirmed. Whether a putative genetic basis for wooden breast can be traced to linked loci under selection for growth traits or is a result of *de novo* mutations and structural variants, remains to be established. The high frequency of WB within flocks is suggestive of underlying causes that are not highly variable among individuals of a flock. Any variation in severity could, therefore, be a result of particular genotypes, and the resulting allele-specific growth traits and nutritional interactions. Such a pattern is consistent with multi-genic traits, alleles for which may segregate in populations as heterozygotes, and the occurrence and severity of the condition may be driven by allele-specific expression patterns [73, 74]. De novo mutations would also explain a *similar* pattern.

The results from our multisample comparison analysis show that WB has an age-dependent expression pattern, with molecular signatures and phenotypic markers becoming more evident in older birds. Such a phenomenon, called age-dependent penetrance, has frequently been observed in various heritable diseases. Recent studies have shown that some features of WB are observable as early as two weeks of age [19, 20], and therefore would explain our

22

observation of 21-day old Ross birds being more similar to the older WB affected the group. The inherited human neurodegenerative disorder, Huntington's Disease, is known to manifest in middle to later life, and gene expression studies show age-dependent expression and dysregulation of various signaling genes [75]. Age-dependent disorders may also include heritable genetic mechanisms such *de novo* mutations, or as somatic mutations [76–78] that may affect genome organization, or repair mechanisms and increase the penetrance of diseases in later life [79, 80]. Further investigations of genome organization, the frequency of *de novo* mutations, or breakdown of repair mechanisms in fast-growth broilers are necessary to illuminate whether and how these processes may be important in WB.

Conclusions

Our study used transcriptomic datasets to compare pectoralis tissue from commercial broilers with wooden breast against multiple genotypic backgrounds, and confirmed the previously reported molecular signatures in addition to previously unreported molecules and pathways. The comparison of tissue from fast-growth genetic backgrounds to those from slow-growth genetic backgrounds and different age classes suggests a genetic basis for WB that elicits age-dependent expression patterns in fast-growth broiler strains. The functional analyses of pathways from comparative data suggest that WB is a potentially polygenic, complex syndrome, with molecular similarities to neoplastic disorders. Through analysis of high-impact variants among the studied breeds, we identified a short list of genes with high-impact variants that are also significantly differentially expressed, suggesting Cis-Regulatory processes involving essential developmental and cytoskeletal genes. This result underscores the need for more in-

23

depth analyses to investigate the role of these genes, basis of these disease pathways and similarities to complex disorders.

Author Contributions: GA conceived, designed experiment design and carried out tissue sampling. SP performed wet lab work in preparation for sequencing. SP, TW, and GA conducted analyses, interpreted results, and generated figures. All authors reviewed the manuscript.

Competing financial interests: The authors declare no competing financial interests

Literature Cited

1. OECD, FAO. OECD-FAO Agricultural Outlook 2017-2026. OECD Publishing; 2017. doi:10.1787/agr_outlook-2017-en.

2. Ayamchirakkunnel Kuttappan V. Characterization and consumer acceptance of white striping in broiler breast fillets. Doctoral dissertation. University of Arkansas; 2012.

3. Kuttappan VA, Brewer VB, Apple JK, Waldroup PW, Owens CM. Influence of growth rate on the occurrence of white striping in broiler breast fillets. Poult Sci. 2012;91:2677–85. doi:10.3382/ps.2012-02259.

4. Kuttappan VA, Brewer VB, Mauromoustakos A, McKee SR, Emmert JL, Meullenet JF, et al. Estimation of factors associated with the occurrence of white striping in broiler breast fillets. Poult Sci. 2013;92:811–9. doi:10.3382/ps.2012-02506.

5. Scheele CW. Pathological changes in the metabolism of poultry related to increasing production levels. Vet Q. 1997;19:127–30. doi:10.1080/01652176.1997.9694756.

6. Siller WG. Deep pectoral myopathy: a penalty of successful selection for muscle growth. Poult Sci. 1985;64:1591–5.

7. Tasoniero G, Cullere M, Cecchinato M, Puolanne E, Dalle Zotte A. Technological quality, mineral profile, and sensory attributes of broiler chicken breasts affected by White Striping and Wooden Breast myopathies. Poult Sci. 2016;95:2707–14. doi:10.3382/ps/pew215.

8. Tijare VV, Yang FL, Kuttappan VA, Alvarado CZ, Coon CN, Owens CM. Meat quality of broiler breast fillets with white striping and woody breast muscle myopathies. Poult Sci. 2016;95:2167–73. doi:10.3382/ps/pew129.

9. Zambonelli P, Zappaterra M, Soglia F, Petracci M, Sirri F, Cavani C, et al. Detection of differentially expressed genes in broiler pectoralis major muscle affected by White Striping - Wooden Breast myopathies. Poult Sci. 2016;95:2771–85. doi:10.3382/ps/pew268.

10. Abasht B, Mutryn MF, Michalek RD, Lee WR. Oxidative stress and metabolic perturbations in wooden breast disorder in chickens. PLoS ONE. 2016;11:e0153750. doi:10.1371/journal.pone.0153750.

11. Sihvo HK, Lindén J, Airas N, Immonen K, Valaja J, Puolanne E. Wooden breast myodegeneration of pectoralis major muscle over the growth period in broilers. Vet Pathol. 2017;54:119–28. doi:10.1177/0300985816658099.

12. Randall CJ. Acute pectoral myopathy in broiler breeders. Avian Pathol. 1982;11:245–52. doi:10.1080/03079458208436098.

13. Wight PA, Siller WG. Pathology of deep pectoral myopathy of broilers. Vet Pathol. 1980;17:29–39.

14. Van Vleet J, Valentine B. Exertional Myopathies. In: Pathology of Domestic Animals. 4th edition. Academic Press; 2007.

15. Siller WG, Wight PAL, Martindale L. Exercise-induced deep pectoral myopathy in broiler fowls and turkeys. Vet Res Commun. 1978;2:331–6. doi:10.1007/BF02291463.

16. Kuttappan VA, Hargis BM, Owens CM. White striping and woody breast myopathies in the modern poultry industry: a review. Poult Sci. 2016;95:2724–33. doi:10.3382/ps/pew216.

17. Van Vleet JF, Valentine BA. Muscle and tendon. In: Pathology of Domestic Animals. 4th edition. Jubb KVF, Kennedy PC, Palmer N, editors. Academic Press; 2007.

18. Mutryn MF, Brannick EM, Fu W, Lee WR, Abasht B. Characterization of a novel chicken muscle disorder through differential gene expression and pathway analysis using RNA-sequencing. BMC Genomics. 2015;16:399. doi:10.1186/s12864-015-1623-0.

19. Griffin JR, Moraes L, Wick M, Lilburn MS. Onset of white striping and progression into wooden breast as defined by myopathic changes underlying Pectoralis major growth. Estimation of growth parameters as predictors for stage of myopathy progression. Avian Pathol. 2018;47:2–13. doi:10.1080/03079457.2017.1356908.

20. Papah MB, Brannick EM, Schmidt CJ, Abasht B. Evidence and role of phlebitis and lipid infiltration in the onset and pathogenesis of Wooden Breast Disease in modern broiler chickens. Avian Pathol. 2017;46:623–43. doi:10.1080/03079457.2017.1339346.

21. Cruz RFA, Vieira SL, Kindlein L, Kipper M, Cemin HS, Rauber SM. Occurrence of white striping and wooden breast in broilers fed grower and finisher diets with increasing lysine levels. Poult Sci. 2017;96:501–10. doi:10.3382/ps/pew310.

22. Radaelli G, Piccirillo A, Birolo M, Bertotto D, Gratta F, Ballarin C, et al. Effect of age on the occurrence of muscle fiber degeneration associated with myopathies in broiler chickens submitted to feed restriction. Poult Sci. 2017;96:309–19. doi:10.3382/ps/pew270.

23. Velleman SG. Relationship of skeletal muscle development and growth to breast muscle myopathies: A review. Avian Dis. 2015;59:525–31. doi:10.1637/11223-063015-Review.1.

24. Bailey RA, Watson KA, Bilgili SF, Avendano S. The genetic basis of pectoralis major myopathies in modern broiler chicken lines. Poult Sci. 2015;94:2870–9. doi:10.3382/ps/pev304.

25. Pampouille E, Berri C, Boitard S, Hennequet-Antier C, Beauclercq SA, Godet E, et al. Mapping QTL for white striping in relation to breast muscle yield and meat quality traits in broiler chickens. BMC Genomics. 2018;19:202. doi:10.1186/s12864-018-4598-9.

26. Mutryn MF. RNA-sequencing analysis of Wooden Breast disease: characterization of a novel muscle disease in chickens through differential gene expression and pathway analysis. University of Delaware; 2015.

27. Clark DL, Velleman SG. Spatial influence on breast muscle morphological structure, myofiber size, and gene expression associated with the wooden breast myopathy in broilers. Poult Sci. 2016;95:2930–45. doi:10.3382/ps/pew243.

28. Kuttappan VA, Bottje W, Ramnathan R, Hartson SD, Coon CN, Kong B-W, et al. Proteomic

analysis reveals changes in carbohydrate and protein metabolism associated with broiler breast myopathy. Poult Sci. 2017;96:2992–9. doi:10.3382/ps/pex069.

29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20. doi:10.1093/bioinformatics/btu170.

30. Dobin A, Gingeras TR. Optimizing RNA-Seq Mapping with STAR. Methods Mol Biol. 2016;1415:245–62. doi:10.1007/978-1-4939-3572-7_13.

31. Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. Curr Protoc Bioinformatics. 2015;51:11.14.1-19. doi:10.1002/0471250953.bi1114s51.

32. Anders S, Pyl PT, Huber W. HTSeq — a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–9. doi:10.1093/bioinformatics/btu638.

33. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40. doi:10.1093/bioinformatics/btp616.

34. Nygaard V, Rødland EA, Hovig E. Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. Biostatistics. 2016;17:29–39. doi:10.1093/biostatistics/kxv027.

35. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8:118–27. doi:10.1093/biostatistics/kxj037.

36. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013;11:11.10.1-11.10.33. doi:10.1002/0471250953.bi1110s43.

37. Horner DS, Pavesi G, Castrignanò T, De Meo PD, Liuni S, Sammeth M, et al. Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing. Brief Bioinformatics. 2010;11:181–97. doi:10.1093/bib/bbp046.

38. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012;6:80–92. doi:10.4161/fly.19695.

39. Ganapathy-Kanniappan S, Geschwind J-FH. Tumor glycolysis as a target for cancer therapy: progress and prospects. Mol Cancer. 2013;12:152. doi:10.1186/1476-4598-12-152.

40. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? Nat Rev Cancer. 2004;4:891–9. doi:10.1038/nrc1478.

41. Han J, Li J, Ho JC, Chia GS, Kato H, Jha S, et al. Hypoxia is a key driver of alternative splicing in human breast cancer cells. Sci Rep. 2017;7:4108. doi:10.1038/s41598-017-04333-0.

42. Sajnani K, Islam F, Smith RA, Gopalan V, Lam AK-Y. Genetic alterations in Krebs cycle

and its impact on cancer pathogenesis. Biochimie. 2017;135:164–72. doi:10.1016/j.biochi.2017.02.008.

43. Kong B-W, Hudson N, Seo D, Lee S, Khatri B, Lassiter K, et al. RNA sequencing for global gene expression associated with muscle growth in a single male modern broiler line compared to a foundational Barred Plymouth Rock chicken line. BMC Genomics. 2017;18:82. doi:10.1186/s12864-016-3471-y.

44. Brown JM, Wilson WR. Exploiting tumour hypoxia in cancer treatment. Nat Rev Cancer. 2004;4:437–47. doi:10.1038/nrc1367.

45. Pouysségur J, Dayan F, Mazure NM. Hypoxia signalling in cancer and approaches to enforce tumour regression. Nature. 2006;441:437–43. doi:10.1038/nature04871.

46. Boomer JS, Green JM. An enigmatic tail of CD28 signaling. Cold Spring Harb Perspect Biol. 2010;2:a002436. doi:10.1101/cshperspect.a002436.

47. Zhou L, Ivanov II, Spolski R, Min R, Shenderov K, Egawa T, et al. IL-6 programs T(H)-17 cell differentiation by promoting sequential engagement of the IL-21 and IL-23 pathways. Nat Immunol. 2007;8:967–74. doi:10.1038/ni1488.

48. Mitsiades CS, Mitsiades N, Poulaki V, Schlossman R, Akiyama M, Chauhan D, et al. Activation of NF-kappaB and upregulation of intracellular anti-apoptotic proteins via the IGF-1/Akt signaling in human multiple myeloma cells: therapeutic implications. Oncogene. 2002;21:5673–83. doi:10.1038/sj.onc.1205664.

49. Schmitz AA, Govek EE, Böttner B, Van Aelst L. Rho GTPases: signaling, migration, and invasion. Exp Cell Res. 2000;261:1–12. doi:10.1006/excr.2000.5049.

50. Kjoller L, Hall A. Signaling to Rho GTPases. Exp Cell Res. 1999;253:166–79. doi:10.1006/excr.1999.4674.

51. Moorman JP, Luu D, Wickham J, Bobak DA, Hahn CS. A balance of signaling by Rho family small GTPases RhoA, Rac1 and Cdc42 coordinates cytoskeletal morphology but not cell survival. Oncogene. 1999;18:47–57. doi:10.1038/sj.onc.1202262.

52. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol. 2012;8:e1002375. doi:10.1371/journal.pcbi.1002375.

53. García-Campos MA, Espinal-Enríquez J, Hernández-Lemus E. Pathway analysis: state of the art. Front Physiol. 2015;6:383. doi:10.3389/fphys.2015.00383.

54. Ma S, Kosorok MR. Detection of gene pathways with predictive power for breast cancer prognosis. BMC Bioinformatics. 2010;11:1. doi:10.1186/1471-2105-11-1.

55. Yu JX, Sieuwerts AM, Zhang Y, Martens JWM, Smid M, Klijn JGM, et al. Pathway analysis of gene signatures predicting metastasis of node-negative primary breast cancer. BMC Cancer. 2007;7:182. doi:10.1186/1471-2407-7-182.

56. Shchetynsky K, Diaz-Gallo L-M, Folkersen L, Hensvold AH, Catrina AI, Berg L, et al. Discovery of new candidate genes for rheumatoid arthritis through integration of genetic association data with expression pathway analysis. Arthritis Res Ther. 2017;19:19. doi:10.1186/s13075-017-1220-5.

57. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. Nat Commun. 2015;6:8971. doi:10.1038/ncomms9971.

58. Liu X, Erikson RL. Polo-like kinase (Plk)1 depletion induces apoptosis in cancer cells. Proc Natl Acad Sci USA. 2003;100:5789–94. doi:10.1073/pnas.1031523100.

59. Nigg EA. Polo-like kinases: positive regulators of cell division from start to finish. Curr Opin Cell Biol. 1998;10:776–83. doi:10.1016/S0955-0674(98)80121-X.

60. Strebhardt K, Ullrich A. Targeting polo-like kinase 1 for cancer therapy. Nat Rev Cancer. 2006;6:321–30. doi:10.1038/nrc1841.

61. Wu C, Dedhar S. Integrin-linked kinase (ILK) and its interactors: a new paradigm for the coupling of extracellular matrix to actin cytoskeleton and signaling complexes. J Cell Biol. 2001;155:505–10. doi:10.1083/jcb.200108077.

62. Qian Y, Zhong X, Flynn DC, Zheng JZ, Qiao M, Wu C, et al. ILK mediates actin filament rearrangements and cell migration and invasion through PI3K/Akt/Rac1 signaling. Oncogene. 2005;24:3154–65. doi:10.1038/sj.onc.1208525.

63. Bonni A, Brunet A, West AE, Datta SR, Takasu MA, Greenberg ME. Cell survival promoted by the Ras-MAPK signaling pathway by transcription-dependent and -independent mechanisms. Science. 1999;286:1358–62.

64. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res. 2002;12:9–18. doi:10.1038/sj.cr.7290105.

65. Trocino A, Piccirillo A, Birolo M, Radaelli G, Bertotto D, Filiou E, et al. Effect of genotype, gender and feed restriction on growth, meat quality and the occurrence of white striping and wooden breast in broiler chickens. Poult Sci. 2015;94:2996–3004. doi:10.3382/ps/pev296.

66. Pastinen T, Hudson TJ. Cis-acting regulatory variation in the human genome. Science. 2004;306:647–50. doi:10.1126/science.1101659.

67. Ongen H, Andersen CL, Bramsen JB, Oster B, Rasmussen MH, Ferreira PG, et al. Putative cis-regulatory drivers in colorectal cancer. Nature. 2014;512:87–90. doi:10.1038/nature13602.

68. Brandler WM, Antaki D, Gujral M, Kleiber ML, Whitney J, Maile MS, et al. Paternally inherited cis-regulatory structural variants are associated with autism. Science. 2018;360:327–31. doi:10.1126/science.aan2261.

69. Camoretti-Mercado B, Dizon E, Jakovcic S, Zak R. Differential expression of ventricular-like myosin heavy chain mRNA in developing and regenerating avian skeletal muscles. Cell Mol Biol Res. 1993;39:425–37.

70. Dalpé G, Leclerc N, Vallée A, Messer A, Mathieu M, De Repentigny Y, et al. Dystonin is essential for maintaining neuronal cytoskeleton organization. Mol Cell Neurosci. 1998;10 5/6:243–57. doi:10.1006/mcne.1997.0660.

71. Liu J-J, Ding J, Wu C, Bhagavatula P, Cui B, Chu S, et al. Retrolinkin, a membrane protein, plays an important role in retrograde axonal transport. Proc Natl Acad Sci USA. 2007;104:2223–8. doi:10.1073/pnas.0602222104.

72. Ferrier A, De Repentigny Y, Lynch-Godrei A, Gibeault S, Eid W, Kuo D, et al. Disruption in the autophagic process underlies the sensory neuropathy in dystonia musculorum mice. Autophagy. 2015;11:1025–36. doi:10.1080/15548627.2015.1052207.

73. Dermitzakis ET. From gene expression to disease risk. Nat Genet. 2008;40:492–3. doi:10.1038/ng0508-492.

74. Lappalainen T. Functional genomics bridges the gap between quantitative genetics and molecular biology. Genome Res. 2015;25:1427–31. doi:10.1101/gr.190983.115.

75. Nguyen HP, Metzger S, Holzmann C, Koczan D, Thiesen H-J, von Hörsten S, et al. Age-dependent gene expression profile and protein expression in a transgenic rat model of Huntington's disease. Proteomics Clin Appl. 2008;2:1638–50. doi:10.1002/prca.200800018.

76. Acuna-Hidalgo R, Veltman JA, Hoischen A. New insights into the generation and role of de novo mutations in health and disease. Genome Biol. 2016;17:241. doi:10.1186/s13059-016-1110-1.

77. Goldmann JM, Wong WSW, Pinelli M, Farrah T, Bodian D, Stittrich AB, et al. Parent-of-origin-specific signatures of de novo mutations. Nat Genet. 2016;48:935–9. doi:10.1038/ng.3597.

78. Veltman JA, Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet. 2012;13:565–75. doi:10.1038/nrg3241.

79. Win AK, Jenkins MA, Dowty JG, Antoniou AC, Lee A, Giles GG, et al. Prevalence and penetrance of major genes and polygenes for colorectal cancer. Cancer Epidemiol Biomarkers Prev. 2017;26:404–12. doi:10.1158/1055-9965.EPI-16-0693.

80. Zane L, Sharma V, Misteli T. Common features of chromatin in aging and cancer: cause or coincidence? Trends Cell Biol. 2014;24:686–94. doi:10.1016/j.tcb.2014.07.001.

81. Yi B, Chen L, Sa R, Zhong R, Xing H, Zhang H. Transcriptome Profile Analysis of Breast Muscle Tissues from High or Low Levels of Atmospheric Ammonia Exposed Broilers (Gallus gallus). PLoS ONE. 2016;11:e0162631. doi:10.1371/journal.pone.0162631.

82. Qiu F, Xie L, Ma J-E, Luo W, Zhang L, Chao Z, et al. Lower Expression of SLC27A1 Enhances Intramuscular Fat Deposition in Chicken via Down-Regulated Fatty Acid Oxidation Mediated by CPT1A. Front Physiol. 2017;8:449. doi:10.3389/fphys.2017.00449.

83. Chen B, Xu J, He X, Xu H, Li G, Du H, et al. A Genome-Wide mRNA Screen and

Functional Analysis Reveal FOXO3 as a Candidate Gene for Chicken Growth. PLoS ONE. 2015;10:e0137087. doi:10.1371/journal.pone.0137087.

84. Davis RVN, Lamont SJ, Rothschild MF, Persia ME, Ashwell CM, Schmidt CJ. Transcriptome analysis of post-hatch breast muscle in legacy and modern broiler chickens reveals enrichment of several regulators of myogenic growth. PLoS ONE. 2015;10:e0122525. doi:10.1371/journal.pone.0122525.

85. Ouyang H, He X, Li G, Xu H, Jia X, Nie Q, et al. Deep Sequencing Analysis of miRNA Expression in Breast Muscle of Fast-Growing and Slow-Growing Broilers. Int J Mol Sci. 2015;16:16242–62. doi:10.3390/ijms160716242.

Figure Legends

Figure 1: Non-Metric Multidimensional Scaling plot of the RNAseq datasets in the multi-sample comparison study. The samples from each age and growth-rate group cluster clearly within groups. Commercial fast-growth broilers also appear more proximate to each other and are arranged from top to bottom in order of increasing age, while youngest and slowest growth breeds are furthest away from older, fast-growth breeds.

Figure 2: Hierarchical clustering heatmap of the top 100 most expressed genes, based on logCPM values, across samples. Clustering shows that woody breast samples fall among Ross 21-day old and ARAC 42 old birds, with slow-growth heritage birds (WC and WRR) forming a distinct cluster.

Figure 3: Mean-Average plots for analysis of differentially expressed genes for each pairwise comparison performed against the woody breast sample set. Points in red show the genes that were expressed at <-2 logFC or > 2 logFC, with a FDR <0.05.

Tables

Table 1: Summary of data used in the comparative analysis, including information about the chicken breed, tissue type, the age of birds at sampling, SRA accession info, and authors of the original study.

SRA accession ID	Chicken Breed/variety	Breed Type	Tissue Type	Age of bird at sampling	Authors
PRJNA339392	Arbor Acres	Commercial Broiler	Pectoralis major	42-days	Yi et al., 2016 [81]
PRJNA342997	WC and WRR	Heritage Broilers	Pectoralis major	120 and 180 days	Qui et al., 2017 [82]
PRJNA294010	Recessive White Rock (WRR) & Xinhua	Heritage & Indigenous	Pectoralis major	42-days	Chen et al. [83]
PRJNA273416	Ross 708 & Illinois Chickens	Commercial Broiler & Legacy Breed	Pectoralis major	6 and 21-days	Davis et al., 2015 [84]
PRJNA266323	WRR & XH	Heritage & Indigenous	Pectoralis major	42-days	Ouyang et al., 2015 [85]

WRR: White Recessive Rock, WC: Wenchang Chicken, XH: Xinghua Chicken

Table 2: Summary of results from pairwise contrasts performed among RNAseq data using edgeR. The table shows the information on differentially expressed genes, summary of pathway analysis using IPA, upstream regulators for each dataset, and top disease terms.

Against	Arbor acres 42Day	Arbor acres 42Day	Arbor acres 42Day	WRR-XHCross	WC120D+	WRR120D+	ROSS-21D	ROSS-6D	ILL-21D	ILL-6D
Group	1	2	3	4	5	6	7	8	9	10
Up in WB	912	855	1001	372	1435	1227	438	686	1131	473
Down in WB	331	312	329	249	849	894	255	482	571	789
Total DE	1243	1167	1330	621	2284	2121	693	1168	1702	1262
Top Canonical PW	T Cell Receptor Signaling	CD28 Signaling in T Helper Cells	Axonal Guidance Signaling	eNOS signaling	Axonal Guidance Signaling	NRF2 Mediated Oxidative Stress	T-Cell Receptor Signaling	CD28 Signaling in T Helper Cells	Signaling by Rho Family of GTPases	CD28 Signaling in T Helper Cells
	IL-8 Signaling	T Cell Receptor Signaling	IL-8 Signaling	Calcium Signaling	Molecular Mechanisms of Cancer	IL-8 Signaling	Production of Nitric Oxide	IL-12 Signaling and Production of Macrophages	CD28 Signaling in T Helper Cells	T-cell Receptor Signaling
	Signaling by Rho Family of GTPases	Phospholipase C Signaling	T-cell Receptor Signaling	Hepatic Fibrosis / Hepatic Stellate Cell Activation	T -cell Receptor Signaling	ERK/MAPK Signaling	CD28 Signaling in T Helper Cells	T-Cell Receptor Signaling	Tec Kinase Signaling	PI3K Signaling in B Lympocytes

Upstream Regs	TGFB1 TNF beta-estradiol	TGFB1 TNF beta-estradiol	TGFB1 TNF Lipopolysaccharide	TNF beta-estradiol TGFB1	TGFB1 beta-estradiol TP53	TGFB1 beta-estradiol TP53	Lipopolysaccharide beta-estradiol TGFB1	TP53 Lipopolysaccharide beta-estradiol	TNF beta-estradiol TGFB1	Lipopolysaccharide beta-estradiol TNF
Diseases	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Neurological Disease	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Gastrointestinal Disease	Cancer Organismal Injury Gastrointestinal Disease

Table 3: List of top 50 canonical pathway terms, and diseases and disorders identified from comparative analysis. In this analysis, multiple DEG datasets are compared to identify common pathway terms that are found more often than expected by chance.

Ranked by -log(P)	Diseases and Disorders	Canonnical Pathways
1	T Cell Receptor Signaling	Cancer
2	CD28 Signaling in T Helper Cells	abdominal neoplasm
3	Signaling by Rho Family GTPases	cancer
4	Phospholipase C Signaling	solid tumor
5	Axonal Guidance Signaling	malignant solid tumor
6	Tec Kinase Signaling	adenocarcinoma
7	FcÎ³ Receptor-mediated Phagocytosis in Macrophages and Monocytes	digestive system cancer
8	FXR/RXR Activation	digestive organ tumor
9	Germ Cell-Sertoli Cell Junction Signaling	abdominal cancer
10	CTLA4 Signaling in Cytotoxic T Lymphocytes	non-melanoma solid tumor
11	Semaphorin Signaling in Neurons	tumorigenesis of tissue
12	Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	neoplasia of epithelial tissue
13	Phagosome Formation	carcinoma
14	GP6 Signaling Pathway	malignant neoplasm of large intestine
15	Role of Tissue Factor in Cancer	intestinal cancer
16	B Cell Receptor Signaling	large intestine neoplasm
17	Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	intestinal tumor
18	Calcium-induced T Lymphocyte Apoptosis	large intestine adenocarcinoma
19	Primary Immunodeficiency Signaling	large intestine carcinoma
20	Colorectal Cancer Metastasis Signaling	intestinal carcinoma
21	Leukocyte Extravasation Signaling	gastrointestinal carcinoma
22	PI3K Signaling in B Lymphocytes	gastrointestinal tract cancer
23	LXR/RXR Activation	gastrointestinal neoplasia

24	Ephrin B Signaling	necrosis
25	Cardiac β-adrenergic Signaling	cell death
26	RhoGDI Signaling	apoptosis
27	CXCR4 Signaling	organization of cytoskeleton
28	Protein Kinase A Signaling	organization of cytoplasm
29	α-Adrenergic Signaling	organismal death
30	Opioid Signaling Pathway	morbidity or mortality
31	Breast Cancer Regulation by Stathmin1	liver lesion
32	Integrin Signaling	abdominal carcinoma
33	Acute Phase Response Signaling	tumorigenesis of epithelial neoplasm
34	iCOS-iCOSL Signaling in T Helper Cells	abdominal adenocarcinoma
35	PKCθ Signaling in T Lymphocytes	tumorigenesis of malignant tumor
36	Glioma Invasiveness Signaling	tumorigenesis of carcinoma
37	Paxillin Signaling	morphology of cells
38	ILK Signaling	morphology of body cavity
39	Th1 and Th2 Activation Pathway	migration of cells
40	Role of JAK1 and JAK3 in γc Cytokine Signaling	cell movement
41	G-Protein Coupled Receptor Signaling	liver carcinoma
42	Th2 Pathway	liver tumor
43	Hepatic Fibrosis / Hepatic Stellate Cell Activation	liver cancer
44	Sirtuin Signaling Pathway	hepatobiliary system cancer
45	Hereditary Breast Cancer Signaling	breast or colorectal cancer
46	Superoxide Radicals Degradation	genital tract cancer
47	Fc Epsilon RI Signaling	pelvic tumor
48	Gαi Signaling	pelvic cancer
49	Complement System	genitourinary tumor
50	Virus Entry via Endocytic Pathways	urogenital cancer

Table 4: Comparison of variant effect prediction between the slow-growth heritage broilers (SGHB), and the fast-growth commercial broilers (FGCB). Results from snpEff based on high-quality SNP variants are shown. Colored boxes highlight notable differences in predicted effects for any category. Green colored boxes show elevated frequency of effects that are less likely to cause adverse impacts, whereas red shaded boxes show elevated frequency of phenotype-changing effects.

Group		SGHB		FGCB	
	Type	Count	Percent	Count	Percent
Effects by impact	HIGH	158	0.03%	395	0.03%
	LOW	60566	11.67%	94998	8.01%
	MODERATE	17001	3.28%	30856	2.60%
	MODIFIER	441262	85.02%	1059364	89.35%
Effects by functional class	MISSENSE	17021	22.22%	30907	25.17%
	NONSENSE	50	0.07%	97	0.08%
	SILENT	59514	77.71%	91795	74.75%
	Missense - Silent ratio	0.286		0.336696	
Count by effects	3 prime UTR variant	53496	10.29%	82491	6.94%
	5 prime UTR premature start codon gain variant	544	0.10%	1302	0.11%
	5 prime UTR variant	3529	0.68%	8050	0.68%
	downstream gene variant	143715	27.65%	292411	24.60%
	initiator codon variant	1	0.00%	6	0.00%
	intergenic region	109334	21.03%	255700	21.51%
	intron variant	70624	13.59%	268782	22.61%
	missense variant	17001	3.27%	30856	2.60%

	non canonical start codon	0	0.00%	1	0.00%
	non coding exon variant	169	0.03%	369	0.03%
	splice acceptor variant	36	0.01%	107	0.01%
	splice donor variant	57	0.01%	152	0.01%
	splice region variant	837	0.16%	3099	0.26%
	start lost	9	0.00%	21	0.00%
	stop gained	50	0.01%	97	0.01%
	stop lost	10	0.00%	25	0.00%
	stop retained variant	24	0.00%	36	0.00%
	synonymous variant	59490	11.44%	91758	7.72%
	upstream gene variant	60901	11.72%	153487	12.91%
	DOWNSTREAM	143715	27.69%	292411	24.66%
	EXON	76508	14.74%	122280	10.31%
	INTERGENIC	109334	21.07%	255700	21.57%
	INTRON	70124	13.51%	266869	22.51%
Count by genomic region	SPLICE SITE ACCEPTOR	36	0.01%	107	0.01%
	SPLICE SITE DONOR	53	0.01%	145	0.01%
	SPLICE SITE REGION	747	0.14%	2771	0.23%
	UPSTREAM	60901	11.73%	153487	12.95%
	UTR 3 PRIME	53496	10.31%	82491	6.96%
	UTR 5 PRIME	4073	0.78%	9352	0.79%

Table 5: Lists of top high-impact variants unique to the fast-growth commercial broilers and slow-growth heritage broiler groups. Highlighted genes are those that were also significantly differentially expressed in comparison of Fast versus Slow groups. The directionality of regulation is also given for these differentially expressed genes.

HI Variants Unique to FGCB	Expression & Direction	Description	Function
ATAD5		ATPase family AAA domain-containing protein 5	Immunoglobulin prouction
C10orf71		chromosome 10 open reading frame 71	Unknown
CELSR1		Cadherin EGF LAG seven-pass G-type receptor 1	Calcium Ion Binding
CUEDC1		CUE domain containing protein	Downregulation of Estrogen Receptor 1
DECR1		2,4-dienoyl-CoA reductase 1	fatty acid beta-oxidation
ENSGALG00000004746		Uncharacterized	Unknown
ENSGALG00000022316	Down in FGCB	C-type lectin family member	Fucose/Mannose Binding
ENSGALG00000023351	Down in FGCB	Small Integral Membrane Protein	Unknown
ENSGALG00000023846		Uncharacterized	Unknown
ENSGALG00000026688		Uncharacterized	Unknown
KIFC1		Kinesin-like protein KIFC1	ATPase activity
MYCBPAP		MYCPG-Associated Protein	Cell differentiation
NES		NEST Protein	intermediate filament binding
NPEPPS		Puromycin-sensitive aminopeptidase	Cellular response to hypoxia
OGFR		Opioid growth factor receptor	Regulation of cell growth
PKD1		polycystin 1, transient receptor potential channel interacting	Regulation of calcium channels
RFC4		replication factor C subunit 4	DNA-dependent ATP-ase activity

	Direction	Description	Function
SPEG	Up in FGCB	SPEG Complex locus	Myocyte cytoskeletal development
STAT1	Up in FGCB	Signal transducer and transcription activator	Mediates cell viability in response to stress
THYN1	Up in FGCB	Thymocyte Nuclear Protein	Induction of apoptosis
TMEM108	Down in FGCB	Transmembrane Protein	cellular response to brain-derived neurotrophic factor stimulus
tvb		TNF-related apoptosis inducing ligand	Tumor necrosis factor-activated receptor activity
WIPI1			

HI Variants Unique to SGHB	Direction	Description	Function
BFIV21		MHC Class 1 component	Metal ion binding
Blec2	Up in FGCB	C-type Lectin Like Receptor 2	Inhibition of natural killer cytotoxicity
C2CD5	Up in FGCB	C2 Calcium Dependent Domaning Containg 5	Calcium bion binding and calcium-dependent phospholipid binding
CASP10	Up in FGCB	Cystein-Aspartic Acid protease family	Regulation of apoptosis
CCNL1	Up in FGCB	Cyclin gene family member	pre-mRNA splicing and regulation of RNA polymerase II
CIRH1A		Ribosome biogenesis factor	nucleolar processing of pre-18S rRNA
DST	Up in FGCB	Dystonin	Cytoskeletal linker protein
ENSGALG00000002067	Down in FGCB	Mannosyltranferase activity	Mannosylation of lipid-linked oligosaccharides
ENSGALG00000021835	Up in FGCB	Uncharacterized	Unknown
ENSGALG00000028551	Up in FGCB	Glutathione transferase	Cellular defense against toxic compounds
Lpin1		Lipin 1 phospatidate phosphatase	Fatty acid metabolism, transcription regulation
MYH1A	Down in FGCB	myosin, heavy chain 1A, skeletal muscle	Microtubule motor activity, actin filament binding
MYH1B	Down in FGCB	myosin, heavy chain 1B, skeletal muscle	Microtubule motor activity, actin filament binding

RPL12		ribosomal protein L12	Ribosomal large subunit assembly
SLC9A2	Up in FGCB	solute carrier family 9 member A2	Proton extrusion, regulation of pH, sodium absorption
STAMBP	Down in FGCB	STAM binding protein	Zinc metalloprotease activtity
STRN3	Up in FGCB	striatin 3	Calcium dependent calmodulin binding
SULT1E1		Sulfotransferase	Sulfate conjugation of estradiol and estrone
TLN1	Down in FGCB	talin-1	actin filament binding
WRAP73	Down in FGCB	WD repeat containing, antisense to TP73	Regulation of mitotic spindle assembly
ZDHHC5		palmitoyltransferase ZDHHC5	Palmitoyltransferase activity
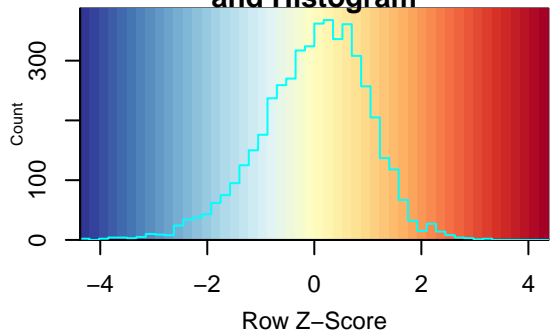
Top 100 most expressed genes across samples
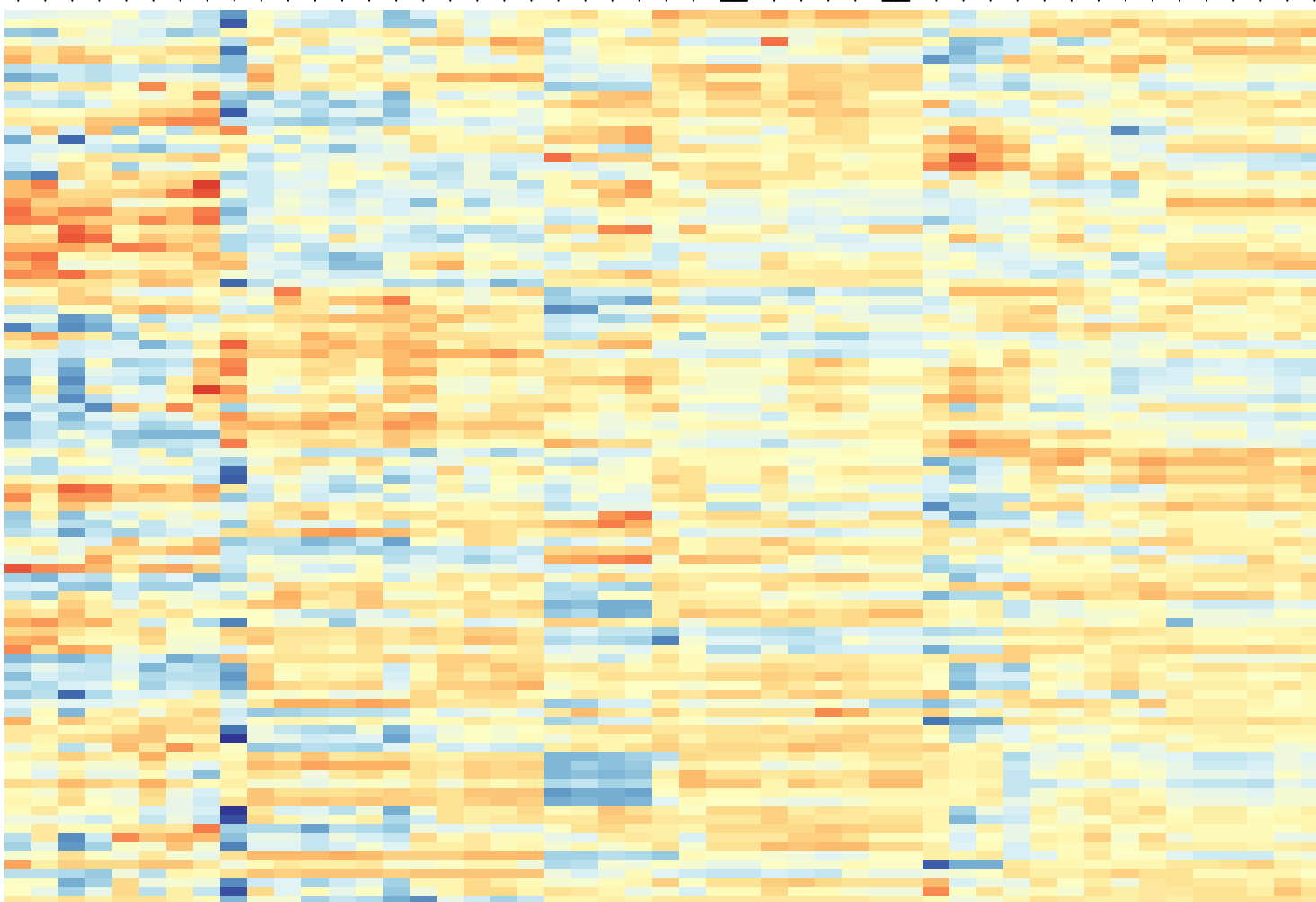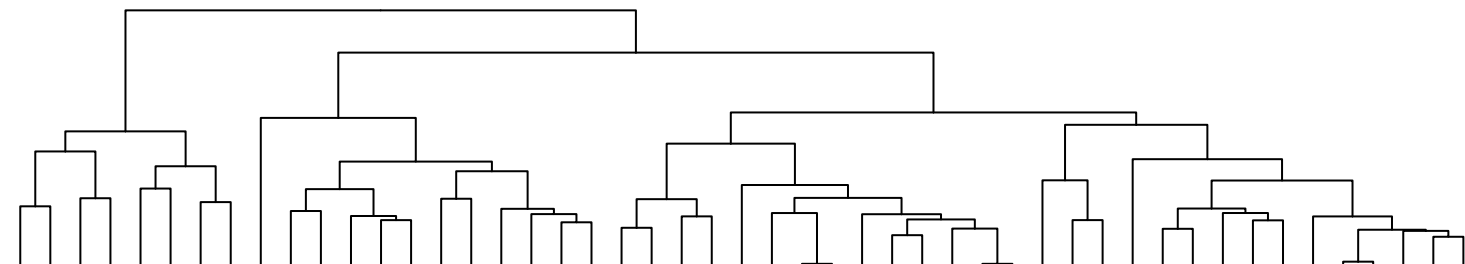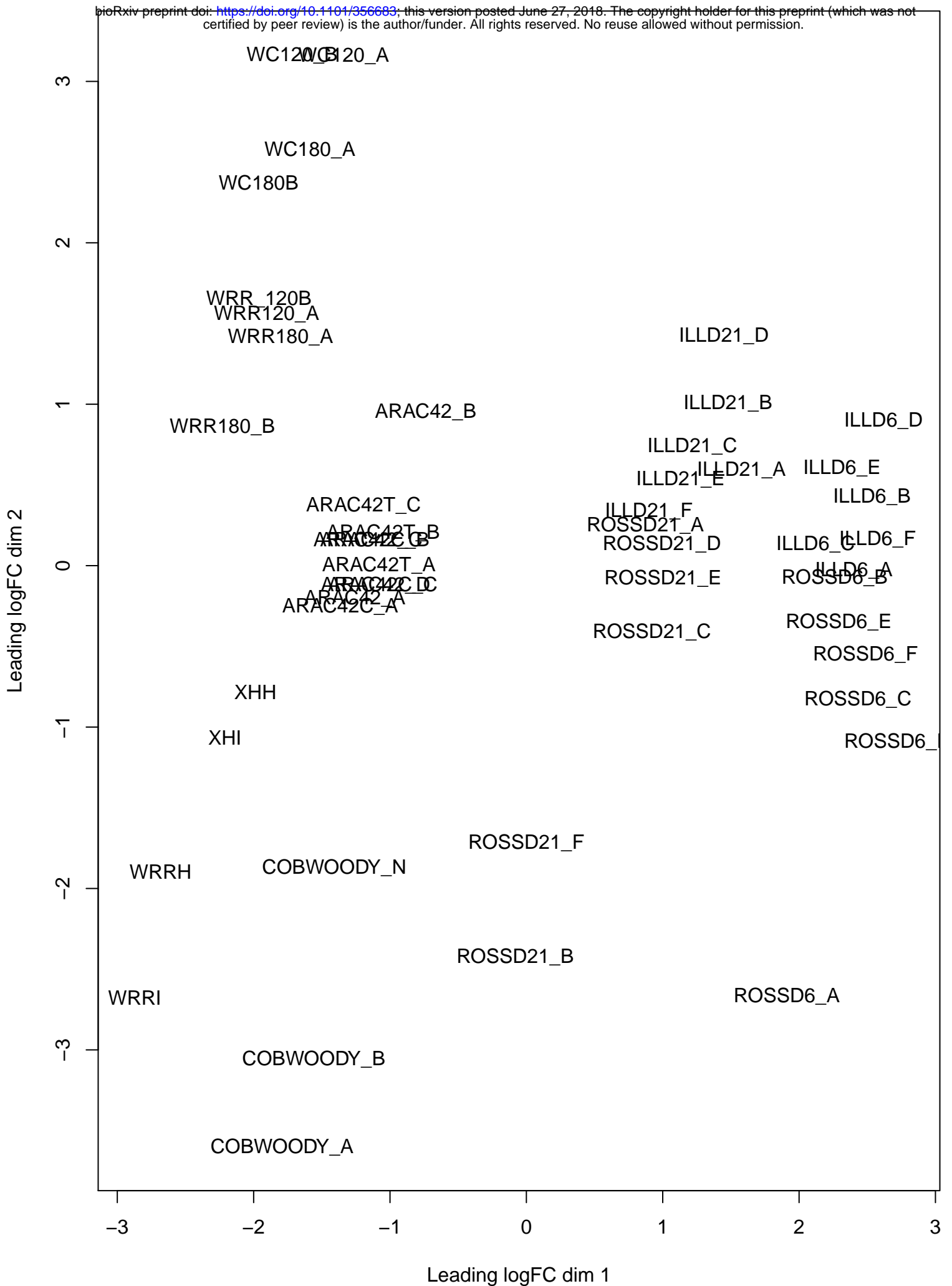
WB vs. ARAC
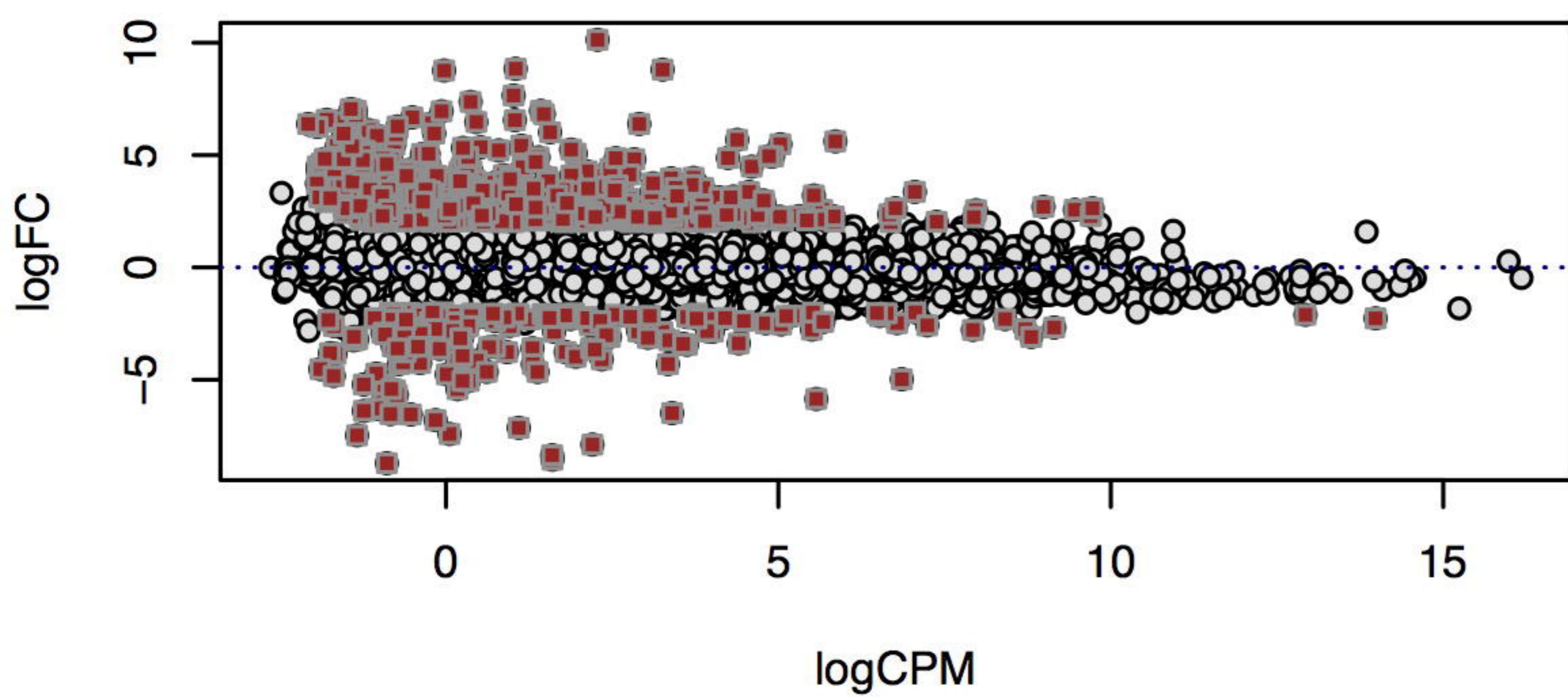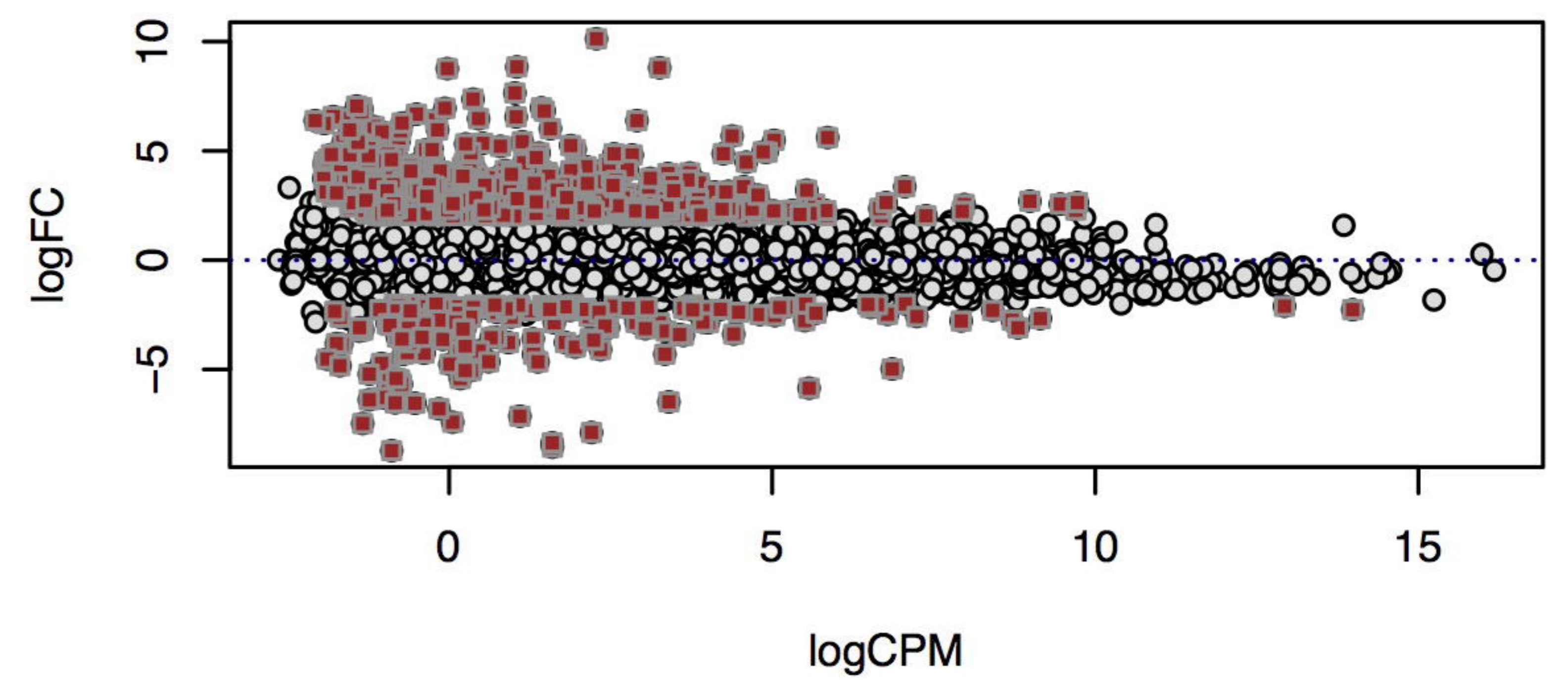
WB vs. ARAC

WB vs. ARAC
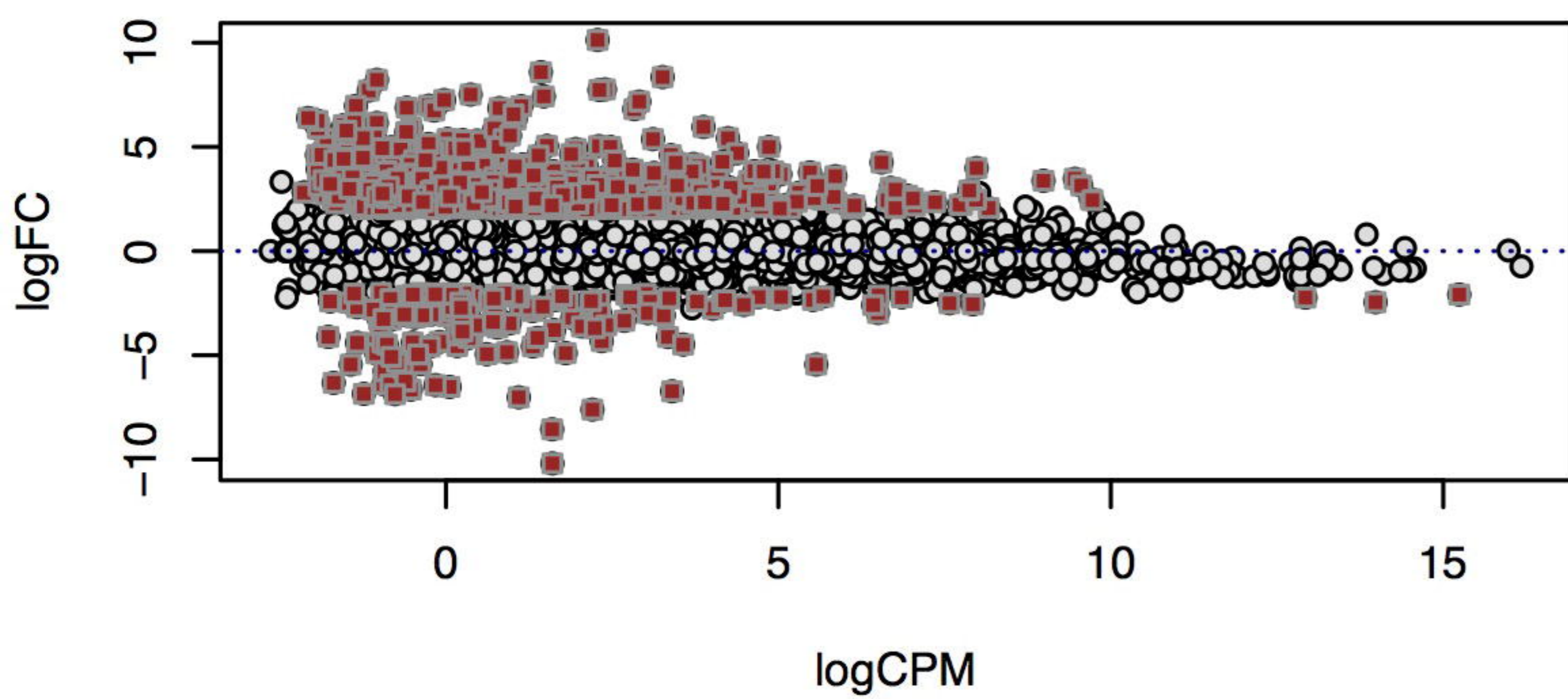
WB vs. WRR–XH Cross
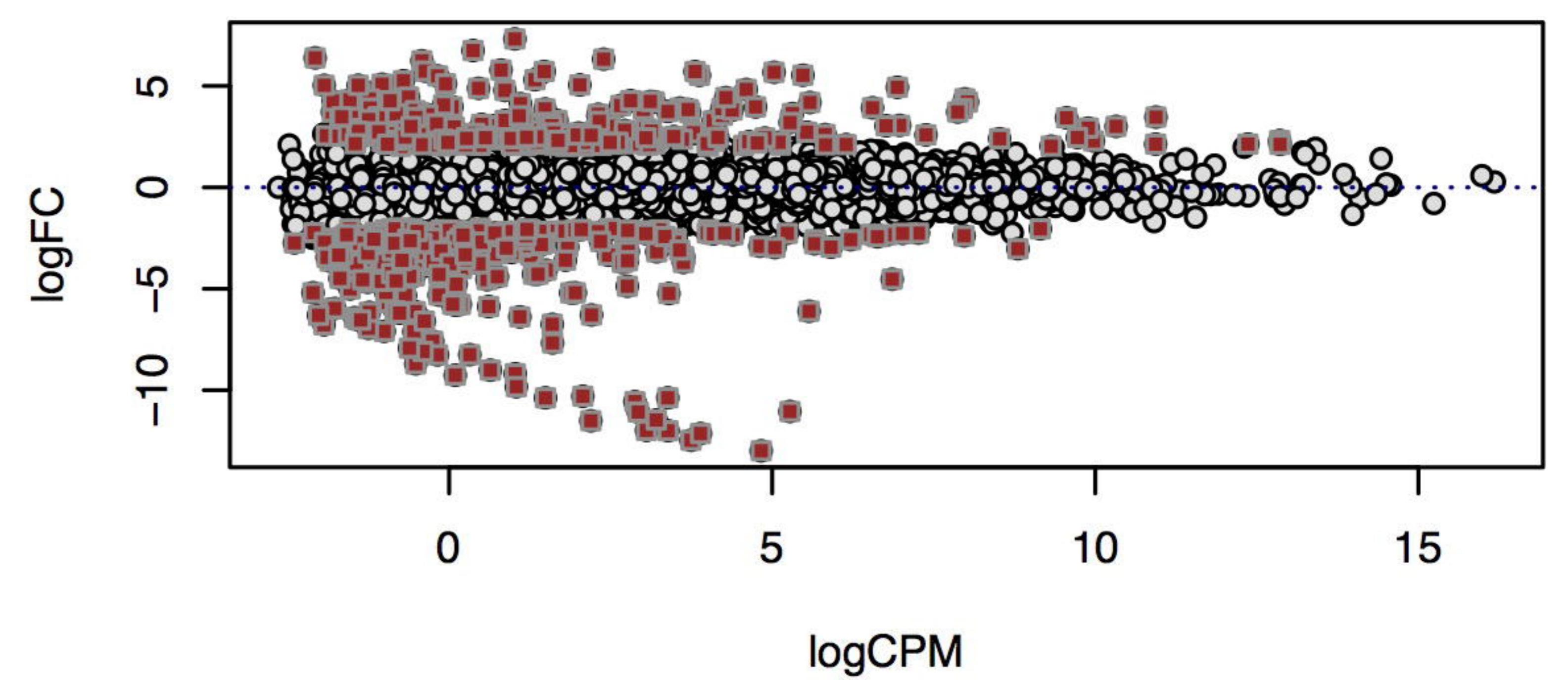
WB vs. WC120 Day+
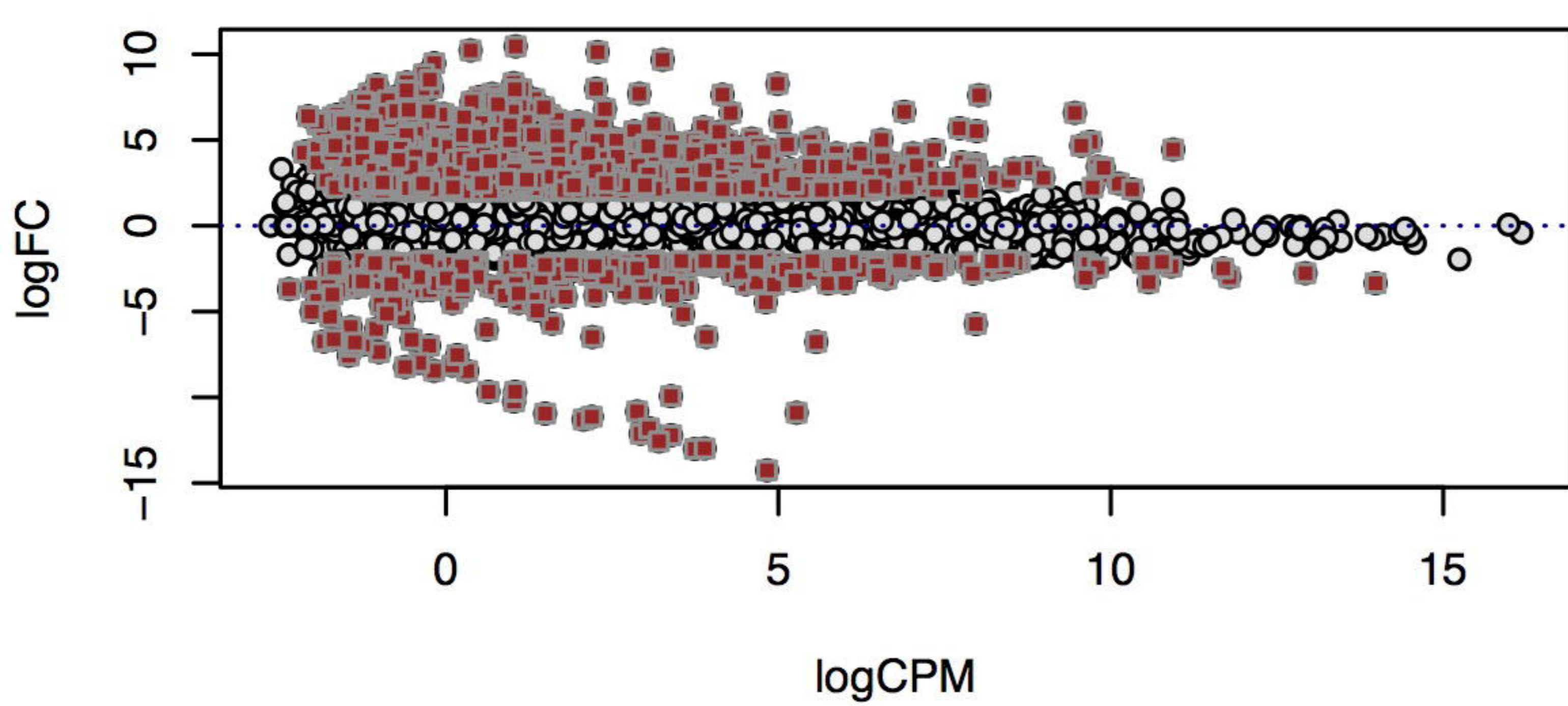
WB vs. WRR 120 Day+
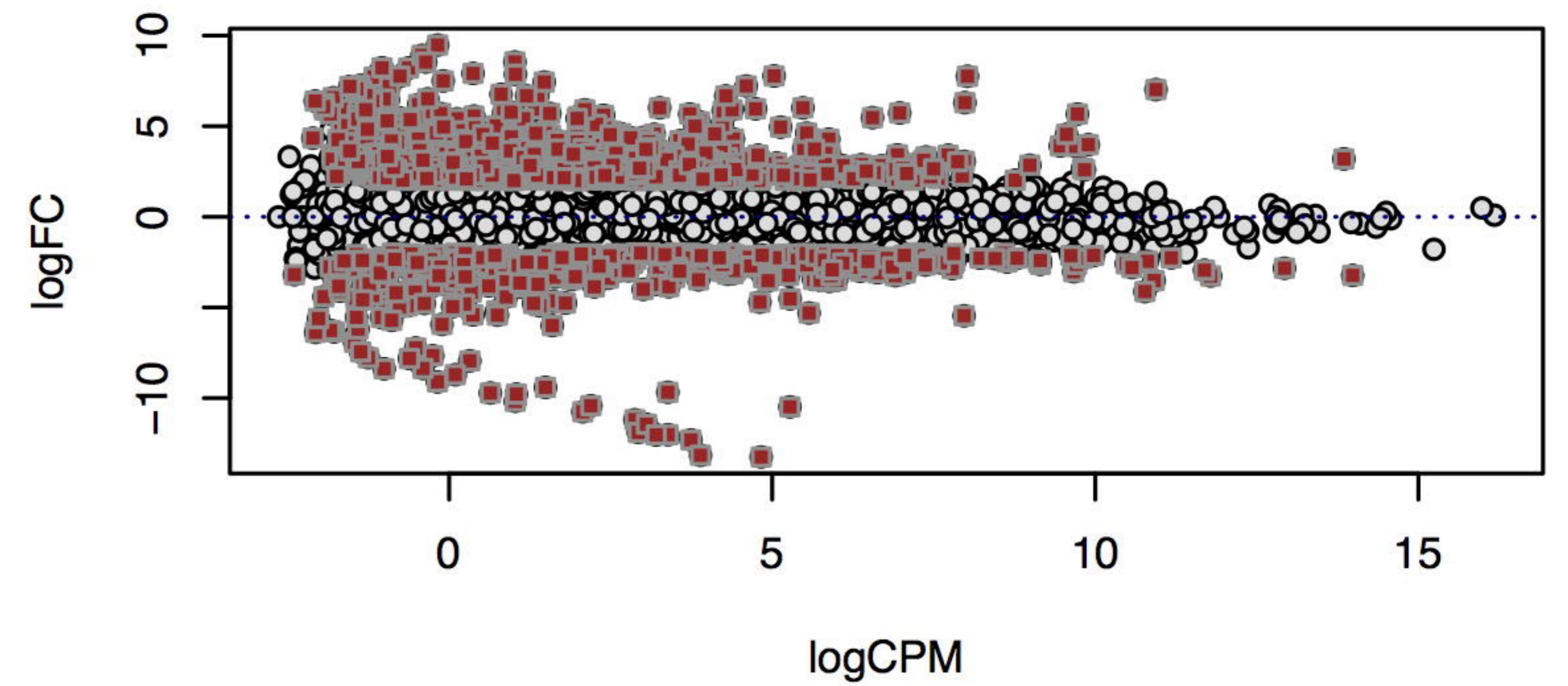
WB vs. Ross 708 21 Day old
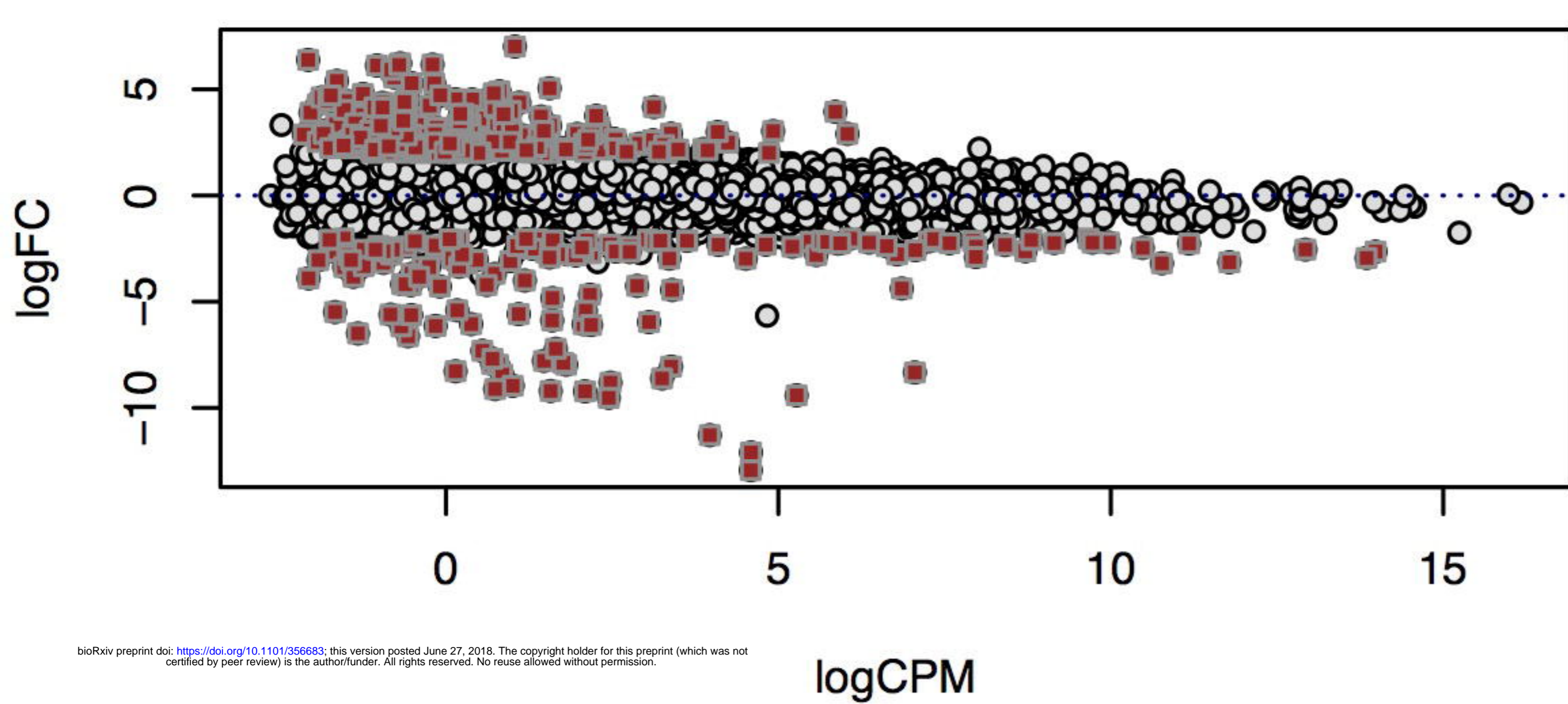
WB vs. Ross 708 6 Day old
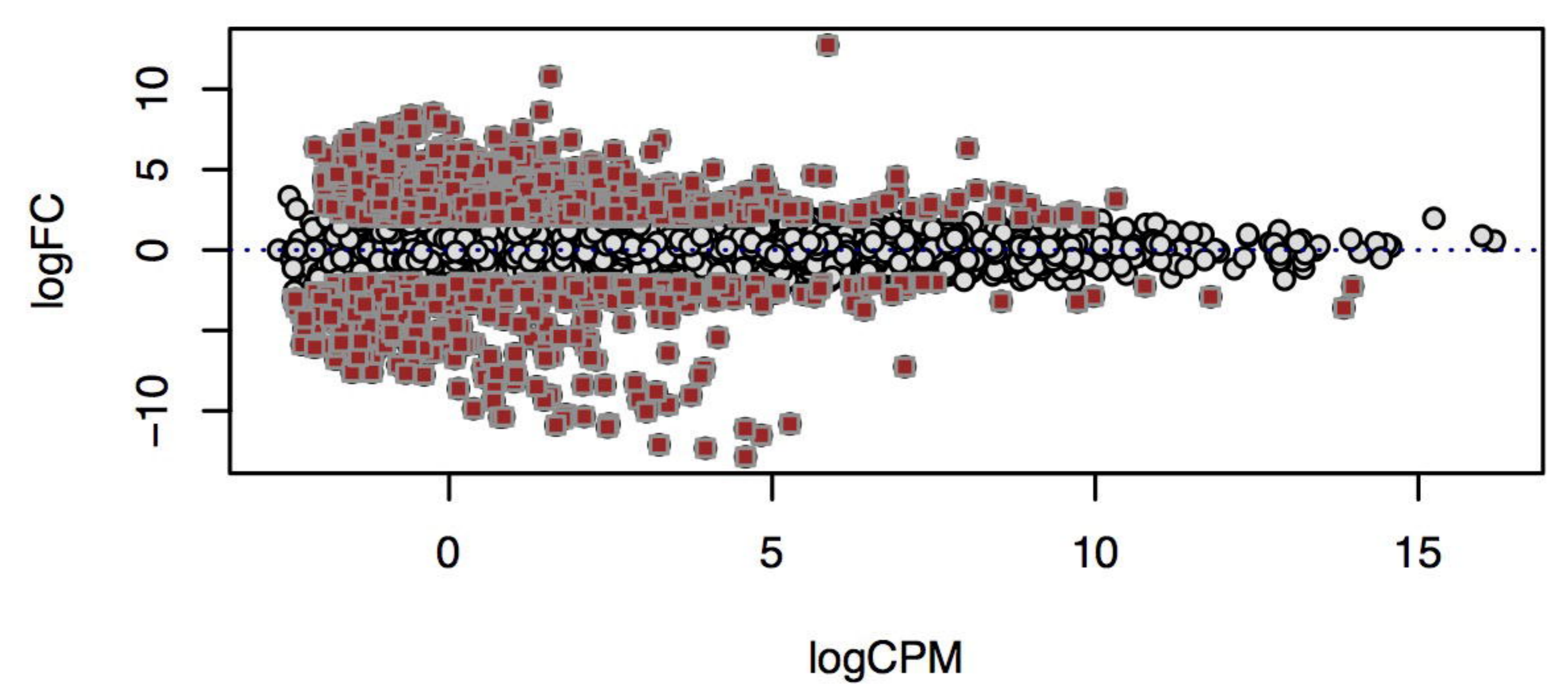
WB vs. Illinois 21 Day old

WB vs. Illinois 6 Day old