MULTIPLEX META-ANALYSIS OF MEDULLOBLASTOMA EXPRESSION STUDIES WITH EXTERNAL CONTROLS

ALEXANDER A. MORGAN
Stanford University School of Medicine
Stanford, CA 94305, USA
Email: alexmo@stanford.edu

MATTHEW D. LI
Stanford University School of Medicine
Stanford, CA 94305, USA
Email: mdli@stanford.edu

ACHAL S. ACHROL
Neurosurgery
Stanford University School of Medicine
Stanford, CA 94305, USA
Email: achrol@stanford.edu

PURVESH J. KHATRI
Institute for Immunity, Transplant and Infection
Stanford University School of Medicine
Stanford, CA 94305, USA
Email: pkhatri@stanford.edu

SAMUEL H. CHESHIER
Neurosurgery, Stanford University School of Medicine
Stanford, CA 94305, USA
Email: cheshier@stanford.edu

We propose and discuss a method for doing gene expression meta-analysis (multiple datasets) across multiplex measurement modalities measuring the expression of many genes simultaneously (e.g. microarrays and RNAseq) using external control samples and a method of heterogeneity detection to identify and filter on comparable gene expression measurements. We demonstrate this approach on publicly available gene expression datasets from samples of medulloblastoma and normal cerebellar tissue and identify some potential new targets in the treatment of medulloblastoma.

1. Background

Highly multiplex gene expression studies using microarrays or RNAseq are very useful for probing the functional genomics of a wide range of biological processes. The analysis of gene expression data typically involves some sort of comparison between samples. Often this comparison is between samples drawn from different conditions. Possible comparisons include samples from tissue treated with different pharmaceuticals, samples drawn from different tissue types or different developmental stages, or samples taken from diseased tissue compared with samples taken from healthy tissue. Many cancer types have been the focus of extensive gene expression analysis, both to identify new molecular subtypes of cancer by comparing different cancer samples, one to another, but also to compare the gene expression differences between healthy tissue and cancerous tissue to help elucidate the molecular processes in different forms of neoplasia. Comparing gene expression levels across thousands of genes in healthy tissue and cancer is a powerful tool in investigating cancer pathogenicity and the development of new pharmacological agents to treat cancer. In many types of cancer such as breast or prostate cancer, it is standard practice during therapeutic surgical removal of a tumor to remove an accompanying portion of nearby healthy tissue surrounding the tumor (i.e. the margin). This provides material from which paired mRNA can be extracted for comparison between healthy and neoplastic tissue.

However, for many types of cancer this is not possible. For primary brain tumors, surgical resection of the tumor is often a balanced tradeoff between removing as much neoplastic material as possible, while leaving as much essential (eloquent) tissue structures as possible to maintain as much function as possible. In aggressive brain tumors, the border of the malignancy and the healthy tissue may not be distinct or clearly separable. For obvious ethical reasons, it is not possible to obtain brain biopsies of healthy tissue from volunteers, unlike tissue types such as skin or blood. This makes having samples for multiplex comparison of gene expression between tumor and healthy brain tissue very difficult.

Medulloblastoma is a type of highly malignant primary brain tumor that typically originates in the cerebellum below the tentorium cerebelli in the posterior fossa. Gene expression studies of samples taken from medulloblastoma solid tumor tissue have focused on identifying different genomic subtypes of medulloblastoma that might lead to new targeted therapies or stratify prognosis [1,2]. Although it might be possible to do a post-mortem analysis of gene expression changes between samples drawn from tumor tissue and nearby brain in those unfortunate individuals who succumb to the disease, most victims of medulloblastoma are treated with radiation, chemotherapy or both, which can cause dramatic gene expression changes in both tumor and non-neoplastic tissue, making a true comparison of tumor with "normal" tissue difficult. Some of the only gene expression datasets of healthy normal brain tissue come from samples taken from freshly deceased cadavers, often from individuals tragically killed in accidents who pre-arranged to donate biological samples to research or whose families do so on their behalf.

Recent developments in techniques of multiplex meta-analysis have led to techniques that synthesize multiple highly multiplex gene expression studies (e.g. microarray or RNAseq) to help remove batch effects, increase statistical power, and identify differences more likely to biologically relevant and to be reproduced in subsequent studies [3–6]. In short, these approaches typically involve two steps, one is to identify if the measurements of gene expression across studies are even comparable, or if there is too much variation. The second step is to develop some overall estimate of the relative variation in gene expression across the studies and its statistical significance, against the typical null hypothesis of no difference in underlying expression between conditions.

One possible way to address this problem of gene expression samples without matched controls is to find a way to identify genes expression profiles which look the same within datasets studying a particular condition (e.g. medulloblastoma and healthy cerebellar tissue), and then look for genes that then vary between datasets. To make this intuition more formal, we propose using a statistical measure of heterogeneity across datasets for medulloblastoma and healthy cerebellum respectively to identify genes that are consistently expressed at an equivalent level within the datasets studying each condition (i.e., low heterogeneity implies homogeneity of expression). At the same time, we compute a meta-estimate of effect (expression level) with an appropriate meta-estimate of a confidence interval in that expression level across datasets and compare these two differences between conditions. Figure 2 shows some contrasting patterns of expression across datasets that demonstrate these concepts pictorially.

In order to investigate this concept further, we searched through the Gene Expression Omnibus (GEO) [7] to identify publicly available datasets of gene expression of medulloblastoma and

healthy, normal cerebellar tissue. To make the best comparison possible, we focused on control brain samples from the cerebellum. Gene expression samples were excluded if they were associated with a particular diagnosis (e.g. Huntington's disease) or a drug treatment. We obtained a total of 191 cerebellar control microarrays, and a total of 414 medulloblastoma microarrays for a total of 605 microarrays. We also collected a dataset of 20 microarrays on brain aging to compare differences in gene expression in the tumor samples with normal brain aging. The datasets we collected are summarized in Table 1. With any large meta-analysis, not all datasets are completely consistent in their methodology or content. The Fiaschetti20011 mRNA is from medulloblastoma tissue culture, not primary tumor tissue, and the Remke2011 and Northcott2012 datasets share some overlap in the tumor source for 15 samples (~5% of the Northcott2012 dataset), but these were processed at different times on different microarray platforms, and we consider them as

Table 1. Gene expression datasets used in this paper. For the multiplex meta-analysis of gene expression in medulloblastoma, four studies of medulloblastoma and four studies of healthy cerebellar cortex were synthesized. An additional dataset of gene expression in the brain as a function of age (individuals age 26-73 were used) was also used to compare gene expression changes in aging against gene expression changes found in medulloblastoma. Note that the GPL570 platform (Affymetrix U133 Plus 2.0) has been used for both some control datasets and some medulloblastoma datasets, setting a point of relative comparison between conditions.

Dataset	Number of Arrays	Sample Type	Gene Expression Series	Pubmed ID	Publication	GEO Platform
Gibbs2010	146	Cerebellar Control	GSE15745	20485568	JR Gibbs, PLoS Genetics, 2010	GPL6104
Hodges2006	27	Cerebellar Control	GSE3790	16467349	A Hodges, Hum Mol Genet, 2006	GPL96
Roth2006	9	Cerebellar Control	GSE3526	16572319	RB Roth, Neurogenetics 2006	GPL570
Roth2007	9	Cerebellar Control	GSE7307		Unpublished	GPL570
Fiaschetti2011	3	Medulloblastoma	GSE22139	21317922	G Fiaschetti, Oncogene, 2011	GPL570
Kool2008	62	Medulloblastoma	GSE10327	18769486	M Kool, PLoS One, 2008	GPL570
Northcott2012	285	Medulloblastoma	GSE37382	22832581	PA Northcott, Nature, 2012	GPL11532
Remke2011	64	Medulloblastoma	GSE28245	21911727	M Remke, J Clin Oncol, 2011	GPL6480
AgingCortex	20	Frontal Cortex	GSE1572	15190254	T Lu, Nature 2004	GPL8300

independent datasets.

2. Results of Analysis

For each microarray dataset, the expression data was obtained from the Gene Expression Omnibus (GEO) [7] and quantile normalized. The probe identifiers for each sample were mapped to Entrez Gene identifiers using AILUN [8]; probes that mapped to multiple identifiers were excluded. If multiple probes mapped to a single gene in a study, the median expression of all probes was taken for that gene. The expression levels of 7724 different genes were measured in all medulloblastoma and cerebellar datasets, but there was also some missing expression levels in individual microarrays, leaving us with 7015 genes with sufficient data to compare expression across

datasets. The genes were quantile normalized the genes across all microarrays together to get a normalized expression level across datasets.

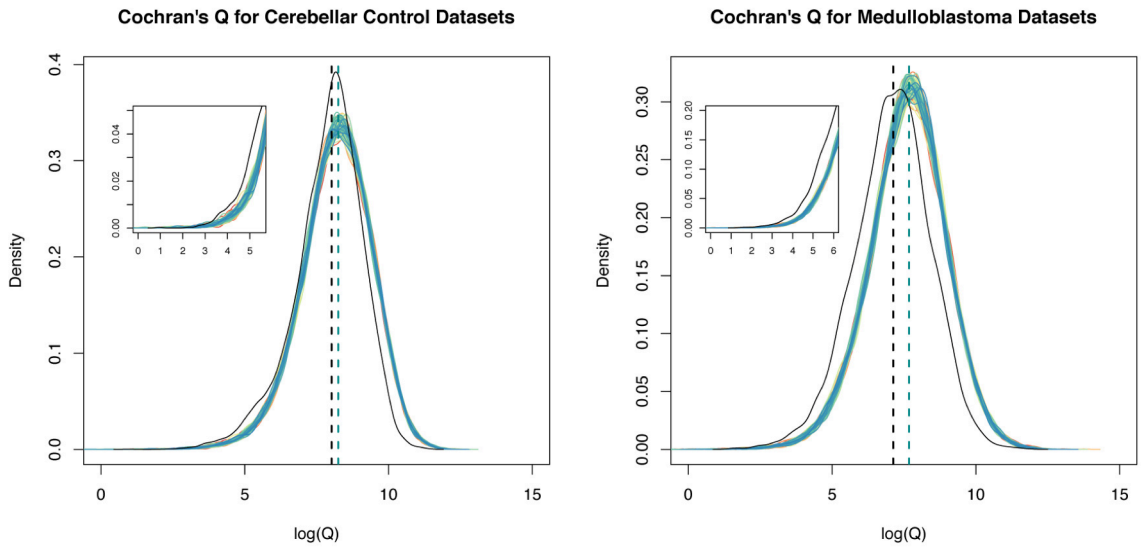

Figure 1: Smoothed histograms of the distribution of Cochran's Q across 100 randomizations (spectrum of colors, visible in blue) compared with the distribution of the measure of heterogeneity in the actual samples in black. The median Q for the actual sample labels is shown in the vertical dashed black line, and then the median for the 100 randomization tests is shown in the blue vertical line to the right. The inset in each panel highlights that at the lower levels of heterogeneity there is substantially more genes showing expression homogeneity in the real data compared with the randomized samples (black line lies above the collection of colored lines, one for each of the 100 randomization).

We then performed a meta-analysis for each gene in the cerebellar and medulloblastoma datasets separately. For each gene in each dataset, we computed the mean expression rank and the standard error of that mean. We used the meta-analysis method proposed by Hedges, et al. [9] which creates a meta-effect estimate based on a random effects linear model, weighting the contribution of the effect (rank expression level) estimate from each included dataset inversely with the standard error of that estimate. This method has been widely used for microarray meta-analysis [5,10,11]. We computed a meta-effect size estimate and we computed a measure of heterogeneity, Cochrane's Q [12] for each gene across the cerebellar and medulloblastoma datasets, respectively. This gave us a consensus measure of relative expression of each gene across the cerebellar studies, a confidence interval around that estimate, and a measure of how heterogeneous/homogenous expression of that gene was across studies. We created the corresponding meta-statistics for expression across the medulloblastoma studies.

By identifying the genes with the lowest 20% of heterogeneity in the cerebellar datasets and the genes with the lowest 20% of lowest heterogeneity in the medulloblastoma datasets, and then taking the intersection, we were left with 318 genes. These represent 318 genes that are consistently expressed at about the same level across all the cerebellar datasets and consistently expressed at about the same level across all the medulloblastoma datasets, but may differ in expression between the two conditions. To test the robustness of this result, we performed 100

random reshufflings of the dataset labels and repeated this analysis. Figure 1 shows that there was more heterogeneity in the randomized samples compared to the actual datasets. The median heterogeneity was always less in the actual data compared to the randomized samples. This suggests that it is possible to find a highly specific set of genes that are more homogenous in expression across datasets for each of the two conditions than random chance.

Of the 318 genes homogeneous in the datasets for both conditions (lower 20% of homogeneity in cerebellar and medulloblastoma datasets), 20 varied in meta-expression difference between medulloblastoma datasets and cerebellar controls greater than the computed meta-confidence interval (Equation 4). In the 100 randomized reshufflings of the dataset labels, there were a median of 8 genes (mean 7.97) genes which met the set criteria for heterogeneity and significant different, suggesting a false discovery rate of 40%.

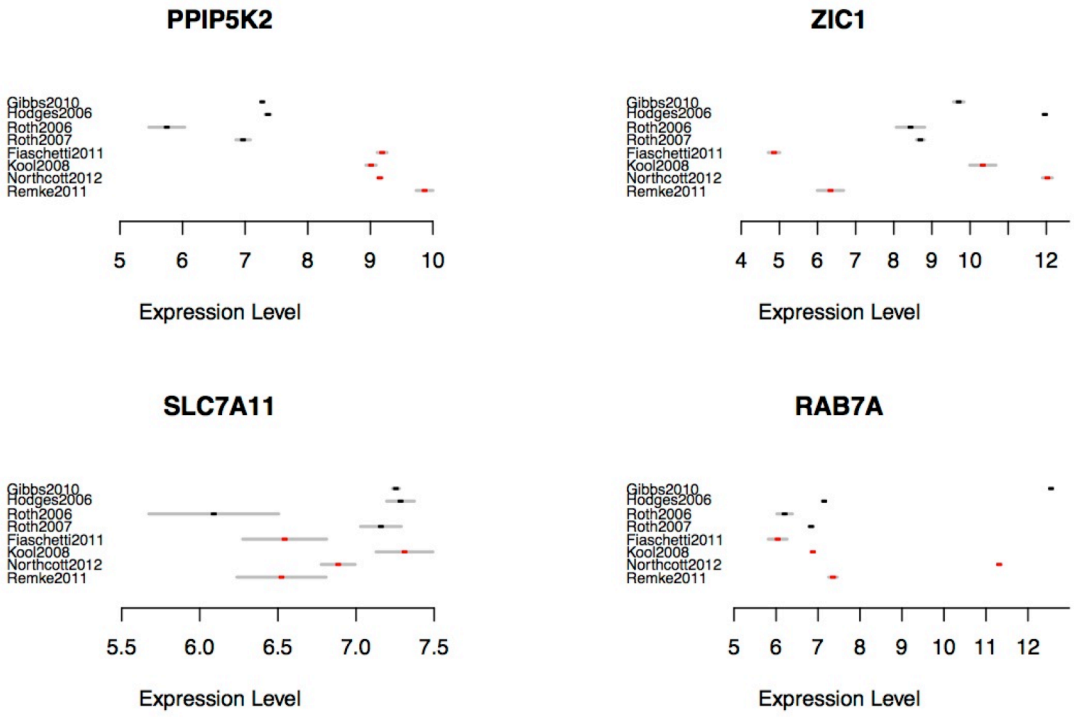

Figure 2: Four forest plots of gene expression across datasets. The four cerebellar controls are shown in black, the medulloblastoma datasets in red. The expression pattern of PPIP5K2 shows narrow confidence intervals and also low heterogeneity, as the expression values across medulloblastoma datasets are similar to one another and the same is true for the controls. PPIP5K2 also shows a pronounced difference in expression between these two conditions; this is the type of expression pattern across datasets that shows strong evidence of increased expression in medulloblastoma. In contrast ZIC1, shows high heterogeneity and thus would be filtered out even though the meta-estimate would suggest substantially lowered expression in medulloblastoma, as the expression levels within datasets studying medulloblastoma vary widely, while the confidence intervals around the expression measurements are also narrow. SLC7A11 shows low heterogeneity, as the expression levels within datasets have broad confidence intervals that nearly overlap, but there is also no significant difference between medulloblastoma and cerebellar controls and thus would be filtered out. RAB7A both has high heterogeneity and low meta-estimate of the difference between medulloblastoma and cerebellar controls and would then be filtered out for both reasons.

Of the 20 genes found to be differentially expressed, two were located on chromosome X or & and may reflect a gender imbalance in samples and were removed, the remaining 18 are shown in Table 2. Figure 3 shows the expression pattern of ENC1 which encodes Ectodermal-Neural Cortex 1, a gene induced by P53 and which interacts with the Retinoblastoma protein. The PRKAR2B gene shows a similar pattern of expression (Figure 3); it is greatly increased in expression in medulloblastoma compared to the healthy cerebellar tissue. PRKAR2B encodes a regulatory subunit that plays a role in regulation of energy metabolism in the cell.

The filtering by heterogeneity is intended to limit the false positive rate, but we might want to focus on sensitivity and expand the coverage of our meta-analysis. Ignoring the filtering by heterogeneity, we can focus on the top genes whose meta-effect estimate significantly differs

Table 2. Top differentially expressed genes with consistent homogeneous expression levels across studies within each condition (medulloblastoma or cerebellar control).

Symbol	Map Location	Description	Cerebellar Control Q	Medullo Q	Cerebellar Expression Level	Medullo Expression Level	Differential Expression
B3GALNT1	3q25	beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	501.1	54.5	6.89 ± 0.22	8.00 ± 0.13	1.10 ± 0.25
DNAJC1	10p12.31	DnaJ (Hsp40) homolog, subfamily C, member 1	533.1	108.7	7.07 ± 0.11	8.25 ± 0.10	1.18 ± 0.15
ENC1	5q13	ectodermal-neural cortex 1 (with BTB-like domain)	559.7	4.3	7.27 ± 0.32	8.92 ± 0.11	1.64 ± 0.34
FAM115A	7q35	family with sequence similarity 115, member A	37.8	84.7	7.50 ± 0.07	9.03 ± 0.15	1.52 ± 0.16
FZD7	2q33	frizzled family receptor 7	190.9	159.1	8.52 ± 0.27	7.42 ± 0.47	-1.10 ± 0.54
LBH	2p23.1	limb bud and heart development homolog (mouse)	384.0	66.2	6.82 ± 0.17	8.06 ± 0.21	1.24 ± 0.27
LMNB1	5q23.2	lamin B1	251.9	155.2	6.86 ± 0.14	8.94 ± 0.24	2.08 ± 0.27
LRIF1	1p13.3	ligand dependent nuclear receptor interacting factor 1	169.6	128.1	6.91 ± 0.09	8.36 ± 0.17	1.45 ± 0.19
MORC3	21q22.13	MORC family CW-type zinc finger 3	511.0	50.8	7.35 ± 0.33	8.52 ± 0.09	1.17 ± 0.34
OSBPL8	12q14	oxysterol binding protein-like 8	561.9	38.6	7.88 ± 0.31	9.42 ± 0.06	1.54 ± 0.32
PAX6	11p13	paired box 6	451.4	144.4	9.52 ± 0.32	7.98 ± 0.51	-1.54 ± 0.60
PODXL	7q32-q33	podocalyxin-like	201.8	63.7	7.16 ± 0.17	8.97 ± 0.15	1.82 ± 0.23
PPIP5K2	5q21.1	diphosphoinositol pentakisphosphate kinase 2	153.7	120.2	6.95 ± 0.10	9.30 ± 0.12	2.34 ± 0.15
PRKAR2B	7q22	protein kinase, cAMP-dependent, regulatory, type II, beta	165.8	32.3	7.09 ± 0.09	9.34 ± 0.17	2.25 ± 0.19
SACS	13q12	spastic ataxia of Charlevoix-Saguenay (sacsin)	145.4	128.6	7.08 ± 0.13	9.04 ± 0.29	1.96 ± 0.32
STMN1	1p36.11	stathmin 1	188.0	177.3	7.53 ± 0.10	8.77 ± 0.12	1.23 ± 0.15
TRMT11	6q11.1-q22.33	tRNA methyltransferase 11 homolog (S. cerevisiae)	183.4	103.4	7.24 ± 0.17	8.40 ± 0.17	1.17 ± 0.24
ZFP36	19q13.1	zinc finger protein 36, C3H type, homolog (mouse)	306.7	121.1	8.00 ± 0.64	6.94 ± 0.33	-1.06 ± 0.72

between cerebellar and medulloblastoma datasets (354 genes were found to be increased in medulloblastoma compared to the controls when ignoring heterogeneity), and do an analysis for functional enrichment of DAVID [13]. This analysis shows that these genes are highly over-enriched relative to the background of the genes measured across all datasets in such functional annotations such as cell cycle (10^{-19}, Benjamini corrected p-value for multiple hypothesis testing), M phase of mitosis (10^{-15}), cell division (10^{-11}) and being involved with cancer (10^{-5}), as might be expected. The 483 genes with lowered expression in medulloblastoma compared to healthy cerebellum (again ignoring the heterogeneity criterion) were highly enriched for genes annotated to be involved in the synapse (10^{-11}), transmission of nerve impulses (10^{-9}), synaptic transmission (10^{-9}), the transport of neurotransmitters (10^{-7}), psychiatric disorders (10^{-6}), and the regulation of nerve impulse transmission (10^{-5}). All this is clearly in line with our understanding of medulloblastoma replacing cells essential to the neurological functioning of the brain with cells focused on rapid replication and suggests that this multiplex meta-analysis approach for using external controls is producing differentially expressed genes with biological relevance to our understanding of medulloblastoma.

To address one of the larger potential biases in our datasets, we also investigated the relationship to differential expression of genes due to normal aging. Although we don't have age information for all of our samples, medulloblastoma is a type of neurological cancer that preferentially targets younger individuals. At the same time, most of the healthy cerebellar brain samples are likely from recently deceased older adults, so there may be a bias toward discovering genes which vary in expression in the cerebellum due to development and aging. We do not have access to expression datasets from healthy cerebellar tissue in children of different ages; however, we do have some expression data on tissue from aging brains in adults. If we look at the dataset on aging of brain samples taken from the frontal cortex of samples taken from recently deceased adults, we can look to see if there is any evidence that the gene expression differences between medulloblastoma and healthy cerebellum could be attributable to simple differences in age. This is not a perfect comparison, but simple compromise based on what data we have available.

Using a dataset from Lu, et al. [14] obtained from the Gene Expression Omnibus [7], we obtained gene expression levels from microarrays made from samples from the frontal cortex from twenty individuals aged 26-73. The original dataset contains additional expression measurements from older brain samples, but we wanted to focus our analysis on gene expression changes in younger adults, and our exploratory analysis found that when using all the data the changes we identified were substantially driven by the samples drawn from much older individuals. Data was again quantile normalized, and we simply looked at the significance test for the Pearson correlation between age and gene expression level. The significance estimates were adjusted using the Benjamini-Hochberg method for addressing multiple hypothesis testing. Examples are shown in Figure 4.

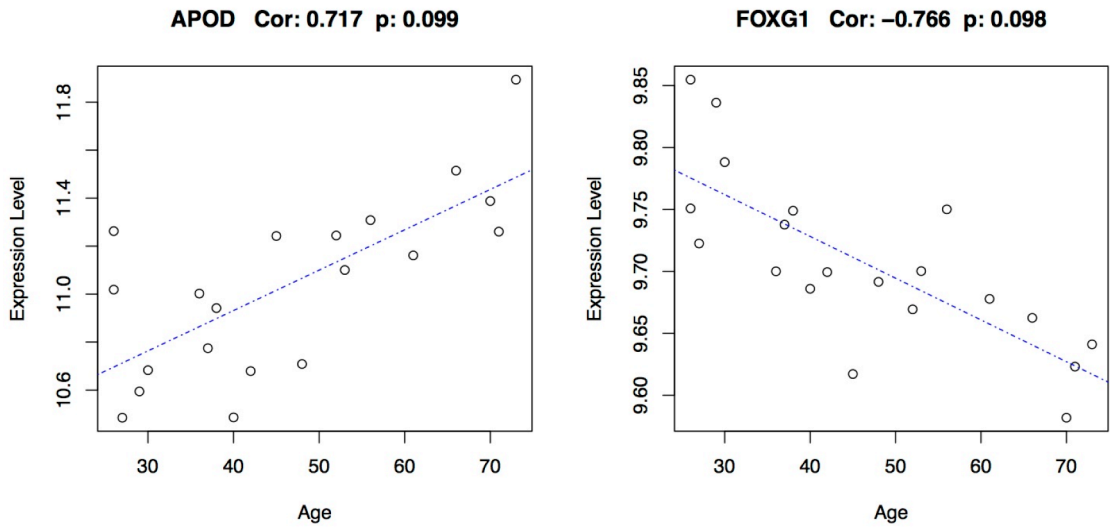

Figure 4: Two examples of genes found to vary in expression with increasing age in the frontal cortex. The expression level is plotted on the vertical axis and the age of the individual on the horizontal. The correlation coefficient and the corrected p-value appear at the top. APOD codes for apolipoprotein D, and FOXG1 codes for a member of the forkhead transcription factors that plays a role in brain development.

None of the same genes were found to be significantly (adjusted $p < 0.1$) differentially expressed with age that were identified in the multiplex meta-analysis of the medulloblastoma and cerebellar controls. Although this does not prove that gene expression changes that we identified in medulloblastoma are not due to differences in the age of individuals sampled, it does suggest that we are not identifying changes in gene expression solely based on the most dramatic age-related changes.

3. Statistical Methods

Computations were done using the 'meta' package in R [15]. For each gene with an average expression level f_i in each dataset, i, the meta-expression estimate of that gene within all the datasets studying a given condition was estimated by taking a weighted average of the expression levels across the gene level median of the probeset expression levels, where the weighting is the inverse of the sum of the within study variance and the estimate of the variance in expression between datasets. For a given gene with expression of f_i and a within dataset variance of v_i in each of k datasets, the estimate of meta fold-change, M, for that given gene is shown in Eq. (1) and as described by Hedges & Olkin [9]. This analysis was done using quantile normalized gene expression levels.

$$M = \frac{\sum_{i=1}^{k} w_i f_i}{\sum_{i=1}^{k} w_i} \tag{1}$$

The weight for the contribution from each dataset, i, is given by adding the estimate of the variance between each dataset and within each dataset and inverting, as shown in Eq. (2).

$$w_i = T^2 + v_i \tag{2}$$

We estimate the between study variance, T^2, using the method of moments, Eq. (3)

$$T^2 = \frac{\sum_{i=1}^{k} \frac{f_i^2}{v_i} - \frac{\left(\sum_{i=1}^{k} \frac{f_i}{v_i}\right)^2}{\sum_{i=1}^{k} \frac{1}{v_i}} - k + 1}{\sum_{i=1}^{k} \frac{1}{v_i} - \frac{\sum_{i=1}^{k} \frac{1}{v_i^2}}{\sum_{i=1}^{k} \frac{1}{v_i}}} \tag{3}$$

The confidence intervals and then p-values for the meta fold-change, M, are computed from the estimate of the variance, v_M, which is computed from the inverse weights, Eq. (4).

$$v_M = \frac{1}{\sum_{i=1}^{k} w_i} \tag{4}$$

The homogeneity test statistic, Q, is computed by Eq. (5)

$$Q = \sum_{i=1}^{k} w_i f_i^2 - \frac{\left(\sum_{i=1}^{k} w_i f_i\right)^2}{\sum_{i=1}^{k} w_i} \tag{5}$$

4. Discussion

We have presented a possible method for the multiplex meta-analysis of gene expression with external controls amenable for use in gene expression studies of some types of cancer, and have presented a set of genes differentially expressed in medulloblastoma compared to cerebellar control tissue. There are considerable batch effect differences that usually make directly comparing two gene expression datasets for differences in expression challenging or impossible. There are also significant differences between gene expression platforms that make a single cross platform analysis impossible. However, the power of looking at multiple studies enables investigation of shared features across datasets to identify commonalities of expression that enable comparison of differences between collections of datasets. We have only begun to scratch the surface of what is possible using the vast resources of the constantly expanding publicly available

data on gene expression. Additional strategies and tools for merging data across varying datasets will be crucial for leveraging the full power of all this data.

Multiplex gene expression measurement modalities are not the only datasets in need of such approaches. For example, the analysis of data drawn from sequencing often is based on a comparison against shared or pooled controls that has its own biases as a semi-external control set. Other highly multivariate (multiplex) measurement modalities will have similar problems. Often when doing such analyses we are interested in an analysis with very high specificity, such as identifying new biomarkers or drug targets, and it is acceptable to filter aggressively, such as by requiring very high levels of homogeneity within datasets of a particular condition, making such an approach tenable. A further investigation of approaches would include a greater examination of non-parametric, rank based approaches, as have been previously investigated for comparing against external controls [16]. It is also possible to use much larger datasets with existing included controls (such as other forms of cancer) to demonstrate accuracy and consistency of results across a variety of cancers and other pathologies or use information about heterogeneity of expression across large numbers of datasets [17].

Although a false discovery rate estimated at 40% may seem unimpressive, it is also highly context dependent. To go from a list of tens of thousands of potential genes, down to a few dozen, with only half of them potentially being false positives may have use in many applications, including biomarker development. Also, the general approach of filtering for genes of within sample class heterogeneity could be used with RNAseq data, which should have substantially less platform variability, but still has experimental and technical biases which would confound direct sample to sample comparison of expression [18].

Another important avenue of future investigation is to look closer at molecular subtypes of medulloblastoma separately. The original research work that provided these medulloblastoma expression datasets identified several clinically different subtypes with characteristic gene expression profiles [1]. Our analysis grouped all the medulloblastoma samples together, looking only at shared properties in expression patterns, but this opens up exciting new possible avenues of hierarchical meta-analytic methodology development.

We hope this work can lead to greater insight into the genomic and molecular pathogenicity of aggressive primary brain tumors like medulloblastoma, and although it will be only one part of future large scale data integration across experimental modalities, it will facilitate further methods of investigations in the absence of custom made control data. The full data from these analyses is available on request from the authors (alexmo@stanford.edu).

5. Acknowledgments

This work was supported by Stanford University School of Medicine's Medical Scholars Program.

References

[1] P. A. Northcott, D. J. H. Shih, M. Remke, Y.-J. Cho, M. Kool, C. Hawkins, C. G. Eberhart, A. Dubuc, T. Guettouche, Y. Cardentey, E. Bouffet, S. L. Pomeroy, M. Marra, D. Malkin, J. T. Rutka, A. Korshunov, S. Pfister, and M. D. Taylor, Acta Neuropathologica **123**, 615 (2012).

[2] M. D. Taylor, P. A. Northcott, A. Korshunov, M. Remke, Y.-J. Cho, S. C. Clifford, C. G. Eberhart, D. W. Parsons, S. Rutkowski, A. Gajjar, D. W. Ellison, P. Lichter, R. J. Gilbertson, S. L. Pomeroy, M. Kool, and S. M. Pfister, Acta Neuropathologica **123**, 465 (2012).

[3] A. A. Morgan, P. Khatri, R. H. Jones, M. M. Sarwal, and A. J. Butte, BMC Bioinformatics **11 Suppl 9**, S6 (2010).

[4] A. Campain and Y. H. Yang, BMC Bioinformatics **11**, 408 (2010).

[5] A. Ramasamy, A. Mondry, C. C. Holmes, and D. G. Altman, PLoS Med **5**, e184 (2008).

[6] F. Hong and R. Breitling, Bioinformatics **24**, 374 (2008).

[7] T. Barrett, D. B. Troup, S. E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I. F. Kim, A. Soboleva, M. Tomashevsky, and R. Edgar, Nucleic Acids Res **35**, D760 (2007).

[8] R. Chen, L. Li, and A. J. Butte, Nat Methods **4**, 879 (2007).

[9] L. V Hedges and I. Olkin, *Statistical Methods for Meta-analysis* (1985).

[10] J. K. Choi, U. Yu, O. J. Yoo, and S. Kim, Bioinformatics **21**, 4348 (2005).

[11] A. A. Morgan, P. Khatri, R. H. Jones, M. M. Sarwal, and A. Butte, in *AMIA Summit on Translational Bioinformatics* (San Francisco, 2010).

[12] Cochrane Collaboration, *Cochrane Handbook for Systematic Reviews of Interventions* (2008).

[13] G. Dennis Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, and R. A. Lempicki, Genome Biol **4**, P3 (2003).

[14] T. Lu, Y. Pan, S. Y. Kao, C. Li, I. Kohane, J. Chan, and B. A. Yankner, Nature **429**, 883 (2004).

[15] G. Schwarzer, *Package "Meta'* (University of Freiburg, 2013).

[16] H. Haeberle, J. T. Dudley, J. T. C. Liu, A. J. Butte, and C. H. Contag, Neoplasia (New York, N.Y.) **14**, 666 (2012).

[17] A. A. Morgan, J. T. Dudley, T. Deshpande, and A. J. Butte, Physiol Genomics **40**, 128 (2010).

[18] K. D. Hansen, R. A. Irizarry, and Z. Wu, Biostatistics (Oxford, England) **13**, 204 (2012).