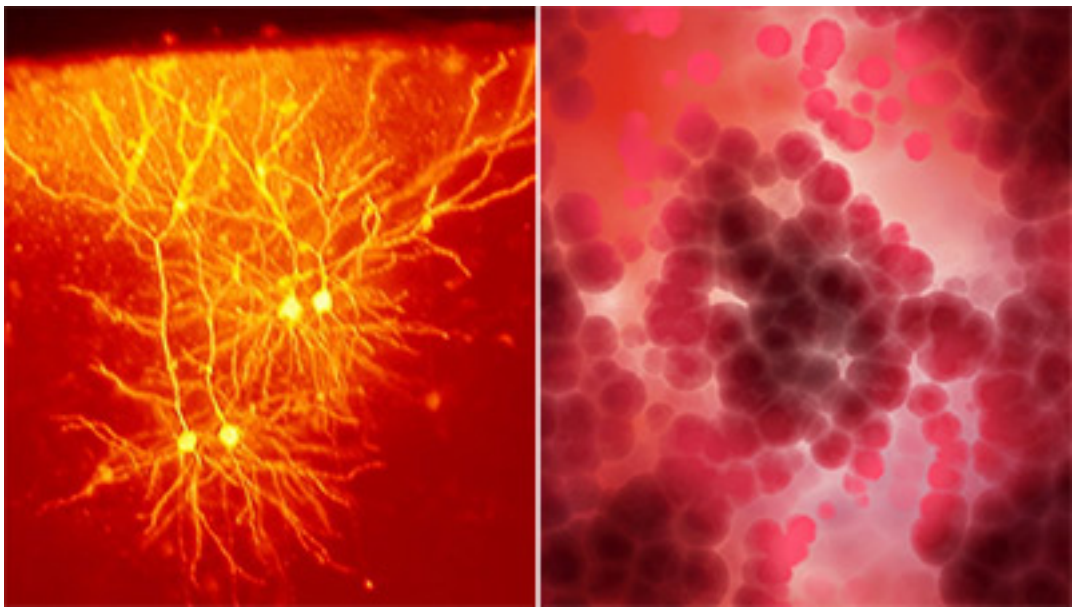
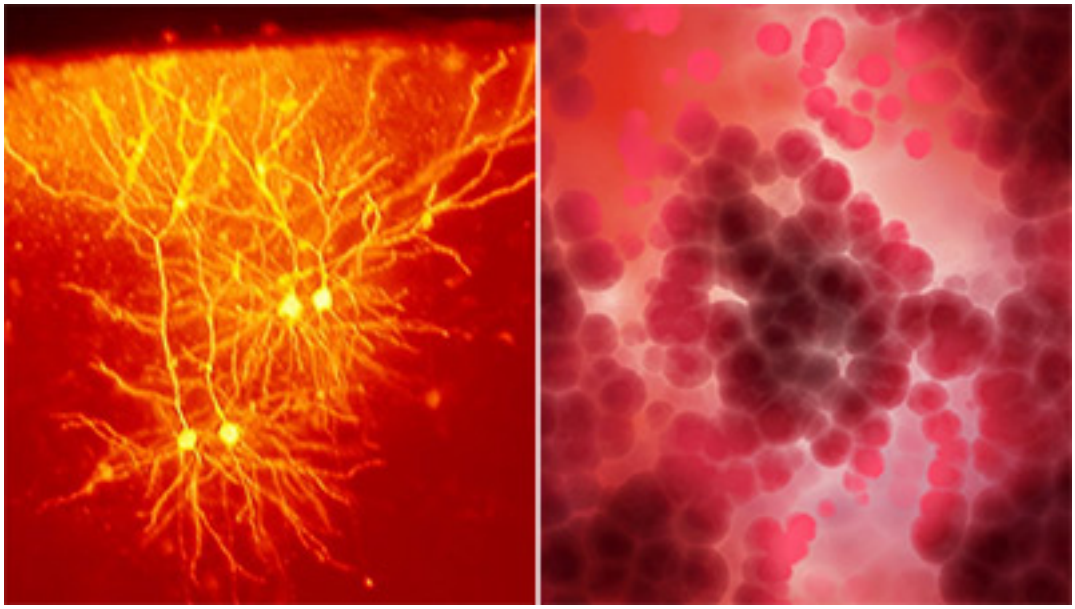
VIEWPOINT

No 'ideal' tissue for gene expression studies of autism

BY **MICHAEL TALKOWSKI, JAMES GUSELLA**

15 APRIL 2014





Model tissue: Nearly 90 percent of autism genes are expressed by mouse neurons, and only about 60 percent in human blood cells.

Measuring RNA expression from individual genes provides one route to novel insights in understanding human disorders such as autism. For example, it may reveal the impact of a genetic mutation on the gene itself or its indirect consequences on the expression of other genes.

Patterns of RNA expression can also help to define genetic networks and to reveal the effects of perturbing these systems. Applying network approaches over time can also inform us about the disorder and how it progresses.

The most appropriate tissue in which to study gene expression and networks is a frequently debated topic. As powerful new technologies generate seemingly endless streams of data, it has become clear that there is no single answer. The best source from which to draw meaningful conclusions about gene expression depends on the specific question being asked. And generalizing from one tissue to another inevitably involves appropriate caveats that must temper interpretation.

The primary debate among investigators studying autism and other psychiatric disorders pits the **use of peripheral tissue** such as blood cells **against brain tissue**. The former is easily accessible from people and can support large population-based collections. On the other hand, we expect the latter to be more directly connected to the neurological and behavioral features of interest.

However, brain tissue is limited to collection after death and can be **greatly influenced by tissue treatment** and handling.

The majority of studies have focused on blood-derived cells, usually lymphoblastoid cell lines, or on postmortem brain tissues. Another option developed in the past few yearsis to create **induced pluripotent stem (iPS) cells**, which can be prompted to form neurons, from individuals with particular disorders. Fibroblasts, or skin cells, are the most frequent source of iPS cells but are certainly not the only option.

Full picture:

The more we learn about gene expression patterns, tissue specificity and genetic alterations in autism, the clearer it becomes that each of these tissues can provide critical and complementary insights. Given the complex genetic architecture of autism, understanding the gene expression changes associated with the disorder will require looking at many different genetic lesions in a diverse set of tissues from human and animal models.

It is becoming increasingly clear that autism is a multi-system disorder that affects much more than just the brain. Because of this, we can't draw all-encompassing conclusions from any one tissue. For example, lymphoblastoid cells are no better proxies for local expression in the brain than postmortem brain tissue is for monitoring a cell's response to environmental stimuli.

Lymphoblastoid cells are useful for screening the impact of an autism mutation directly on gene function and for discovering genome-wide interactions. They capture natural genetic backgrounds and provide a measure of the interactions between gene networks. The major caveat, however, is that not all genes of interest in a disorder such as autism are expressed in these cells. Some network connections seen in these cells may not even exist in other tissues.

Researchers are conducting ongoing **RNA-sequencing studies** to directly address this caveat. For example, the Genotype-Tissue Expression project is creating a comprehensive public atlas of gene expression across multiple human tissues from genetically characterized individuals[1]. From these data, we see that individual gene expression varies dramatically between tissues.

Brain-derived tissues express the major proportion of known genetic contributors to autism.

One example is PPRT2, which is a gene located in the **16p11.2 region linked to autism**. It codes for a protein expressed during early developmental stages in the brain and the spinal cord. PPRT2 is expressed at levels more than 100 times higher in brain cells than it is in blood cells. In most cases, however, lymphoblastoid cells and neural tissues are more similar than this example suggests.

In one ongoing study, we evaluated gene expression patterns in three types of tissue that researchers use regularly in autism studies: lymphoblastoid cells, neural progenitor cells derived from iPS cells, and brain tissue from mice.

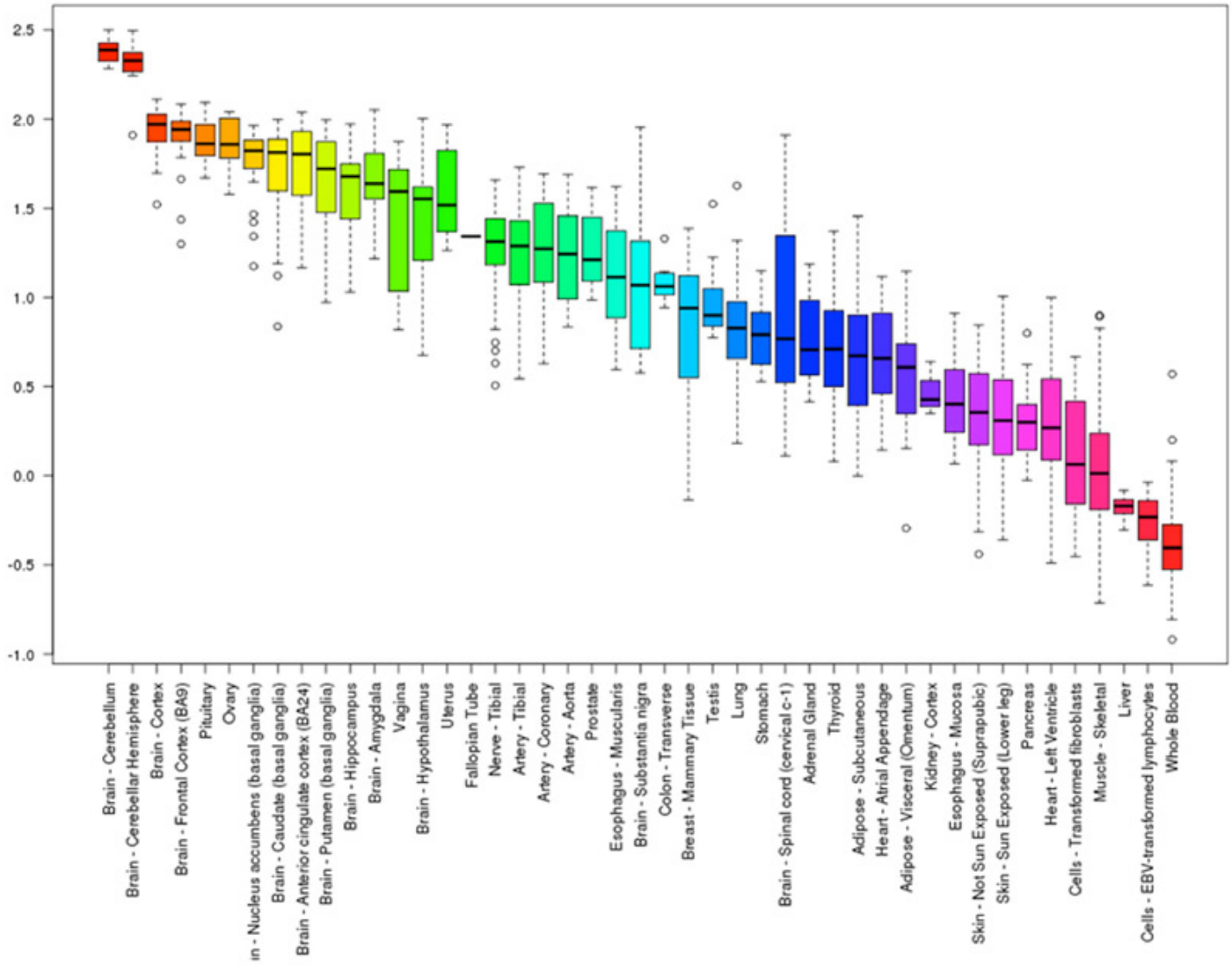
Using sequencing methods and an approach that can detect low levels of RNA (called RNA-Seq), we found that 57 percent of measurable genes are expressed in all three tissues. The two human tissues are as similar to each other as the human neural progenitors are to the mouse cortex. These meaningful similarities across tissues suggest that one can exploit easy-to-access-tissues such as lymphoblastoid cells to understand basic functions of some autism genes and to predict at least some of their potential effects in other tissues.

Autism model:

A critical question is how much of the picture of gene expression each type of system is able to capture. We directly compared the expression of 604 genes previously implicated in autism, as described by **SFARI Gene**[2] and AutismKB[3], in these three systems.

We found that 53.1 percent of autism genes are captured in all three tissues; 364 (60.3 percent) are measurable in lymphoblastoid cells, 474 (78.4 percent) in neural progenitors and the most, 543 (or nearly 90 percent) are expressed in the mouse cortex.

Collectively, 209 autism-associated genes measured in neurons from mice or iPS cells are not detectable in lymphoblastoid cells, whereas only 9 autism genes are detectable in these cells alone.



Enlarge image >>Brain skew: PPRT2, a protein important for brain development, is expressed 100 times more in brain cells (left) than it is in blood cells (right).

These data suggest that brain-derived tissues express the major proportion of known genetic contributors to autism and, if readily accessible, would be the preferred tissue for most measurements. However, even in these nervous-system tissues, RNA levels may not capture all genetic factors predisposing to autism.

Unfortunately, at present there are pragmatic barriers to large-scale studies with such samples. **Human autism postmortem tissue is limited**, and large-scale generation of neural progenitor

cells from iPS cells has not yet been achieved, although it is rapidly becoming more feasible. Nonetheless, the combination of hundreds, or even thousands, of autism risk factors and the biological and technical variance inherent in any gene expression study dictate that large sample sizes will be needed.

The 'all-or-nothing' tissue-specific expression pattern described above is but one limitation to interpretation from a single tissue. Another is isoform-specific expression,meaning cases in which more than one version of a gene product is expressed from a single gene. For example, **TCF4** — which regulates the expression of other genes, which is mutated in Pitt-Hopkins syndrome[4] and which has been **implicated in autism**[5] and schizophrenia[6] — is present in different isoforms in lymphoblastoid cells than in neural progenitors.

For such genes, looking at gene expression in only one type of tissue could lead to misinterpreting the changes that result from a certain mutation. A mutation affecting one of the longer forms of TCF4 that are not expressed in lymphoblastoid cells may be harmful only in the brain, for example, and not in other cell types. This sort of diversity strongly emphasizes the point that we need detailed studies in each specific tissue of interest. What's more, one cannot assume that all autism genes act directly in the brain.

There is hope that iPS cells will offer a technical leap toward understanding autism-specific gene expression, as they can be derived from people who have autism risk factors and retain the influences on gene expression inherent in the genetic background of each individual. Alternatively, researchers can control for the effects of varying genetic backgrounds by using **genome engineering technology**, either correcting mutations or introducing them into control lines.

iPS cell-derived neurons may also offer the opportunity to compare neurons between people with the same disorder or with mutations in the same gene**.** Moving toward large-scale studies in this system is not a trivial challenge, but these offer a promising way forward.

Another major consideration is the degree to which any RNA-based study, irrespective of tissue, can model a disorder. The amount of RNA in a cell does not necessarily reflect the level of the corresponding protein. Association-based gene expression networks are a small but critical component of the overall cause-and-effect networks that confer risk for the disorder.

We will ultimately need to integrate data on all aspects of gene expression and its regulation. This will allow us rely less on correlation and to begin to address relationships between genetic perturbations and the disorder's natural progression.

Indeed, as some studies have elegantly demonstrated, the **stage of development being tested** can be as important as the tissue itself[7, 8].

From these data, we conclude that no study design provides a panacea for all of the caveats that

accompany any gene expression study in a complex disorder such as autism. Ultimately, we will need to investigate a broad spectrum of human- and animal-derived tissues to accurately reflect the biological complexity of the disorder and to use this knowledge to develop rational therapeutic interventions.

Michael E. Talkowski is assistant professor of neurology at Harvard Medical School. James F. Gusella is the Bullard professor of neurogenetics at Harvard Medical School.

References:

1: GTEx Consortium *Nat. Genet.* **45**, 580-585 (2013) **PubMed**

2: Abrahams B.S. *et al. Mol. Autism* **4**, 36 (2013) **PubMed**

3: Xu L.M. *et al. Nucleic Acids Res.* **40**, D1016-1022 (2012) **PubMed**

4: Rosenfeld J.A. *et al. Genet. Med.* **11**, 797-805 (2009) **PubMed**

5: Talkowski M.E. *et al. Cell* **149**, 525-537 (2012) **PubMed**

6: Ripke S. *et al. Nat. Genet.* **43**, 969-976 (2011**) PubMed**

7: Willsey A.J. *et al. Cell* **155**, 997-1007 (2013) **PubMed**

8: Parikshak N.N. *et al. Cell* **155**, 1008-1021 (2013) **PubMed**