Not All Autism Genes Are Created Equal: A Response to Myers et al.

To the Editor: In our recent Autism Sequencing Consortium (ASC) study,[1] we show that cohorts ascertained for autism spectrum disorder (ASD) and those ascertained for a broader array of neurodevelopmental disorders differ both in the rate and in the relative distribution of mutations across genes. In commenting on our study, Myers and colleagues[2] appear to misinterpret our analyses and mischaracterize our conclusions. Here, we set the record straight on the substance of their concerns and consider how their focus on the clinical/diagnostic relevance of rare mutations might have contributed to these misunderstandings.

Myers and colleagues[2] (henceforth Myers) are concerned about trends in the genetic testing world to generate what some claim to be "autism panels" and to market them. We see and agree with their concern. Included among the authors in Myers are leading medical geneticists, and this is their area of expertise. However, the ASC report[1] does not present evidence for "ASD-specific" genes; the notion of ASD-specific genes is Myers' construction, not ours. In fact, in the ASC manuscript[1] (see also Castellani and Arking[3]), we observe that, with current sample sizes, we cannot *yet* identify genes that have significantly more mutations in our ASD-ascertained cohort than in a more generally ascertained developmental disabilities cohort. We do, however, clearly show that these cohorts differ in the rate, and in the relative distribution, of mutations across genes. The Myers commentary appears to misconstrue our analysis of heterogeneity across these cohorts and its purpose, not distinguishing between clinically useful findings, their area, and scientifically interesting ones, our area.

Given the synopsis of our work in Myers, it is useful to review what we did do in the ASC study. In short, our work covered the following topics: generation and compilation of data on ultra-rare mutations in the protein-encoding portions of genes in a large number of individuals with ASD; interpreting those mutations, ranging from the general observations about whether the overall observed mutational counts are more than expected by chance, how they disrupt the function of genes, how they are distributed over male versus female individuals with ASD, and whether they accumulate in certain genes far more than expected by chance, in which case they are dubbed "ASD genes;" relating these ASD genes to the far more common genetic variation in the population and to structural variation known to affect ASD and other neurodevelopmental disorders; given the strong genetic overlap of ASD and other neurodevelopmental disorders, studying whether there is any way to tease out which ASD genes have a larger impact on the core features of ASD and which have more general effects on neurodevelopment, including elevating the chance of diagnosis of ASD; and finally, investigating where in neurobiological development atypical development is likely to arise and how, given the identified set of ASD genes.

Within this range of topics, Myers seizes upon the relationship between ASD and other neurodevelopmental disorders. The kernel of their argument arises from this question we asked: is there any way to tease out which ASD genes have a larger impact on the core features of ASD—social deficits, restrictive interests, and repetitive behaviors—than those having more general effects on neurodevelopment? This is a thorny problem. It has been known for decades that other forms of neurodevelopmental disorders overlap with ASD; for example, about 50% of males and 20% of females with fragile X syndrome would also receive a diagnosis of ASD if they were assessed for it.[4] The same phenomenon is observed for other genes found to be associated with ASD. Myers gives several such examples and, in the face of this observation, says the problem cannot be solved without data that do not yet exist.

We thought differently; perhaps there would be a path forward via the comparison of the mutational pattern in the ASC sample, ascertained for idiopathic ASD, to a cohort ascertained for more general developmental disorders, most often neurodevelopmental disorders, which we will call NDDs. In other words, was there evidence for genes that were more associated with cohorts ascertained for ASD than with cohorts ascertained for NDDs and, hence, more readily discoverable in an ASD-ascertained cohort? We introduced the analysis with a clear statement of purpose: "Distinguishing genes that, when disrupted, lead to ASD more frequently than NDD could shed new light on how atypical neurodevelopment maps onto the core deficits of ASD."[1]

To address this question, we first had to determine whether there was heterogeneity between the sample sets: for the 102 genes significantly enriched for deleterious variants in ASD, we can evaluate whether the count of disruptive *de novo* events in genes is homogeneous for the ASC versus NDD samples. We report that the data are highly heterogeneous ($p = 5 \times 10^{-12}$). Thus, we conclude that there *is overwhelming statistical evidence* that existing ASD-ascertained cohorts are distinct from an existing NDD-ascertained cohort, not just in rate but in the relative distribution of mutations across genes. All 102 discovered genes in our study are ASD genes by our criteria (i.e., a gene's transmitted and de novo association analysis (TADA) q value < 0.1 for the ASC cohort), while the significant heterogeneity of the observations in ASC and NDD samples means that some genes are more readily identified in ASD-ascertained cohorts.

After, and only after, confirming this heterogeneity, the ASC then carried out *secondary*, exploratory analyses,

making use of a binary classifier to separate the 102 ASD genes into those more often disrupted in individuals ascertained for ASD versus those more often disrupted in individuals ascertained for NDDs. The use of a binary classifier is standard for such an exploratory analysis when a quantitative trait is bimodal or (as in this case) the ratio is of two small observed counts and thus not quantitatively precise. The ASC choice of 1:1, particularly since it split the data roughly in half, is a natural split for such a classifier. Myers objects to our classifier because it has a "scientifically arbitrary threshold used to define ASD-predominance …"[2] It does; classifiers typically do. The key question we sought to answer was, is it useful?

After much discussion among our group, we carefully chose the terms "ASD-predominant" (ASD_P) and "ASD-NDD" (ASD_{NDD}) to label the two groups of genes assigned by the classifier. There is an important reason for what looks like an absurdly complicated nomenclature. All of our participants were ascertained for ASD. Thus, the genes we discover are all "ASD genes." *Some* of those ASD genes have more mutation-based evidence for overlap with intellectual disability and other NDDs, i.e., the ASD_{NDD} genes, and others have less, i.e., the ASD_P genes. (And note that ASD_P and "ASD-specific" should not be conflated because the latter suggests a clearly identifiable group with ASD alone.) The question we then asked was, are there meaningful differences in individuals *with ASD* who carry mutations in ASD_P versus ASD_{NDD} genes? Here are the striking differences between these two groups.

First, when repeating the heterogeneity analyses for the ASD_P and ASD_{NDD} genes by themselves, neither subsample shows significant heterogeneity. This is important to understand. In other words, *as imperfect as the classifier is, it still leads to more homogeneous grouping* when looked at from a very solid perspective of rates of disruptive mutations.

Second, looking at protein-truncating variants (PTVs) in *parents* for the two classes of genes within the ASD-ascertained sample, the ASC observed significantly greater frequency in ASD_P genes (1.17 per gene) than in ASD_{NDD} genes (0.45 per gene) ($p = 6.6 \times 10^{-6}$), whereas the frequency of *de novo* PTVs in the ASC *cases* is not markedly different between the two groups (95 in ASD_P genes, 121 in ASD_{NDD} genes; $p = 0.07$). Similarly, within families, rare PTVs in ASD_P genes showed significant transmission disequilibrium to affected children, something that was not observed in ASD_{NDD} genes. The ASC then notes that, "The paucity of inherited PTVs in ASD_{NDD} genes is consistent with greater selective pressure acting against disruptive variants in these genes and highlights fundamental differences between these two classes."[1]

Third, the ASC looked at comorbid phenotypes in those *ASD-ascertained samples* that had this information available and found significant differences based on class of mutation, even though the samples were ascertained for an ASD diagnosis. We observed that individuals with ASD who carry a disruptive *de novo* variant in ASD_{NDD} genes walk 2.6 ± 1.2 months later than those with disruptive *de novo* variants in ASD_P genes ($p = 2.3 \times 10^{-5}$). Moreover, individuals with ASD who carry a disruptive *de novo* variant in ASD_{NDD} genes have an IQ 11.9 ± 6.0 points lower than those with disruptive *de novo* variants in ASD_P genes ($p = 1.1 \times 10^{-4}$).

On the basis of this evidence, we find our classifier for ASD_P and ASD_{NDD} to be useful scientifically. Yet, we *completely agree* with Myers that this binary classifier has, as they put it, "dubious clinical significance."[2] However, we made no such claim, nor did we even suggest that all genes are correctly and perfectly classified into the ASD_P and ASD_{NDD} groups. Rather, we demonstrate, by several analyses, that the *classification is broadly meaningful, genetically, developmentally, and behaviorally.* In other words, what Myers deems impossible is possible, although it is a small start on a challenging problem. It is important to note here that it is completely consistent to observe compelling and unequivocal heterogeneity between two groups while being unable to assign individual members to one category or the other unequivocally. Significance testing to classify particular genes "correctly," as proposed by Myers, is not meaningful in the context of the ASC analyses.

This issue of finding scientific utility for imperfect classifiers is not unique to genetics. Imposing a classifier of "people who smoked at least one pack of cigarettes a day" versus "those who never smoked," Hammond and Horn[5] were able to convincingly establish that smoking is associated with death from cancer. Did their work mean to suggest that smoking one cigarette less than a pack was fine but one cigarette more was pathological? No, of course not. Did they suggest that one pack (not more or less) was a clinically actionable amount of smoking? Absolutely not. Did they intend to argue that one pack a day was the best possible classifier of smoking? No, they merely established a link between smoking and cancer death so strong that it formed the rational basis for all the continued research into smoking and death that has happened since. A classifier can be scientifically important and useful without ever reaching the standard of being "ideal" or clinically actionable.

Also, scientists are familiar with the use of arbitrary classifiers. They have generally agreed on a threshold for significance of a single hypothesis test, 0.05. When the p value is 0.049, they declare the null hypothesis rejected in favor of the alternative. When it is 0.051, the evidence is insufficient.

Changing direction slightly, Myers goes on to argue that one can only begin to determine whether there are ASD_P genes—in our meaning, not theirs, we hope—by studying cohorts on which there is complete and consistent phenotyping, which in the case of the ASC and NDD cohorts, there was not. They rightly note that the NDD cohort was never assessed systematically for ASD (nor could it be if some of the affected individuals were too young at enrollment) and it could have high rates of ASD. That is an interesting point; however, considering the implications of incomplete phenotyping in the two cohorts, we note that it actually *strengthens* the conclusions of the

ASC. Misphenotyping or incomplete phenotyping, as they describe, will make the two cohorts more similar to each other than otherwise. If the NDD cohort were a mixture of subjects manifesting NDDs without ASD and others with NDDs and ASD, then the NDD cohort must be more similar to the ASC's ASD cohort than would be a cohort consisting solely of subjects with NDDs and without ASD. This follows from first principles in statistics. Thus, to the extent that one adopts this criticism, one is effectively saying that there is likely to be *greater* evidence of differences between "cleaner" ASD cohorts and NDD cohorts. From the ASC perspective, since the two samples sets are demonstrably different, this limitation, although real, is of limited impact on the ASC conclusions and not in the direction that concerns them.

Among the many points Myers make about teasing apart the core features of ASD from cognitive function, they indicate that no gene could be labeled ASD_P unless, when it is mutated in individuals with ASD, their average IQ = 100. Specifically, they argue "What would be necessary to demonstrate meaningful ASD specificity (or predominance) of large-effect rare variants? If loss of function of a particular gene conferred risk that was purely specific to ASD, the mean IQ associated with *de novo* pathogenic variants in that gene would not be significantly different from the population mean (100), or at least from the familial background mean …"[2] This argument is groundless. That a gene is more readily detectable in ASD cohorts than it is in NDD cohorts does not imply it has no effect on NDDs or "IQ." That a gene might have more effect on one phenotype (ASD) than it does on another (IQ) does not mean that it must have no effect on the other. Genetic pleiotropy can exist such that a variant in a gene has a large effect on one trait while having a small effect on another. For instance, you would discover *de novo* mutations in *FGFR3* in individuals ascertained for extremely short stature far more often than you would in individuals ascertained for spinal stenosis. That is because *FGFR*3 mutations have a much larger effect on adult height than they do on spinal stenosis, but they certainly affect both traits.

Coming full circle, we reiterate that the ASC never claimed to find ASD-specific genes. Instead, our findings are consistent with the view that there is a phenotypic spectrum of impact where mutations in some genes are more deleterious than in other genes, a statement so trivial that we do not expect much disagreement on this point. We show that there is heterogeneity across existing ASD- and NDD-ascertained cohorts, and we expect subsequent studies will seek to identify genes more specifically associated with one or the other end of this spectrum. In fact, we can even use our results to highlight where to look for ASD_P genes. First, they will be found more readily in cohorts without clinically significant intellectual disability. When the ASC partitioned ASD subjects into those with an IQ of 70 or higher (69.4%) versus those with an IQ of less than 70 (30.6%), individuals in the higher-IQ group still carry a greater burden of *de novo* variants relative to expectation, and this remains true when partitioning the IQ at the cohort mean (full-scale IQ 82) or when considering the 102 ASD genes only. In addition, we know that mutations in the ASD_P genes are going to be more likely to be inherited compared to more severe ASD_{NDD} genes, so a focus on inherited rare variation will illuminate more such genes. Our current gene findings in ASD are biased toward more deleterious mutations and to variants of higher effect sizes (by the pragmatic focus on *de novo* variants and variants absent in large control cohorts). Sample sizes will need to be substantially larger to identify ASD_P genes that are harboring rare variants that are more likely to be inherited and of lower effect size.

And this brings us to another important scientific concern: that of study design. If there were no appreciable difference between rare variant discovery in ASD- and NDD-ascertained cohorts, it would be more efficient to study NDD cohorts, where mutations per subject are higher and phenotyping is potentially less costly. Under this scenario, one could ask, what is the justification for ongoing large-scale collection and genetic analysis of ASD-ascertained samples, including multiple efforts led by authors on Myers et al.? Since our earliest analyses, the question of uniformity among the genes in our findings has been raised, and we decided in our most recent work to further examine this question. Having demonstrated clear differences between ASD- and NDD-ascertained cohorts, and having provided evidence for the existence of ASD_P and ASD_{NDD} genes, we can say with confidence that experiments to identify individual ASD_P and ASD_{NDD} genes are worth pursuing, and as noted above, we can begin to define best approaches to the discovery of ASD_P genes. And, to the degree that Myers argues for a large, extensively phenotyped and genetically characterized cohort as the basis for *ideal* study design, who could argue?

In summary, for all the reasons outlined above, we agree with Myers about the lack of clinical utility of "ASD-specific" panels and that there is "currently insufficient evidence to establish ASD-specificity of any genes."[2] The utility of our classification of genes into ASD_P and ASD_{NDD} genes resides in its ability to highlight sets of genes that, when disrupted, alter the core features of ASD while creating lesser or greater perturbations of other features of neurodevelopment. No individual ASD_P gene, when disrupted, is known to affect only core features of ASD; in turn, no ASD_{NDD} gene, when disrupted, is known to have only negligible impact on those core features. At the group level, these sets of genes are useful biologically—to understand developmental processes underlying ASD and NDD—but not clinically, because of the phenotypic differences they evoke.

Our results also show that there *will* be genes that are more clearly associated with the extremes of a clinical phenotypic spectrum. We would further argue that the discovery of genes associated with the extremes of the spectrum will provide further insights into pathways disrupted in ASD and that such discovery is an important direction for future research. We look forward to working with our

colleagues in the ASC and in the broader community, including the Myers team, toward attaining these goals.

Joseph D. Buxbaum,[1,*] David J. Cutler,[2] Mark J. Daly,[3,4] Bernie Devlin,[5] Kathryn Roeder,[6] and Stephan J. Sanders[7], for the Autism Sequencing Consortium
[1]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [2]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA; [3]Broad Institute of MIT and Harvard, Cambridge MA 02142, USA; [4]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA; [5]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, USA; [6]Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, PA 15313, USA; [7]Department of Psychiatry, UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA 94158, USA
*Correspondence: joseph.buxbaum@mssm.edu

Funding

This work was supported by the National Institute of Mental Health (MH111661, MH111660, MH111658, and MH111662).

References

1. Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L., et al.; Autism Sequencing Consortium; and iPSYCH-Broad Consortium (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell *180*, 568–584.e23.
2. Myers, S.M., Challman, T.D., Bernier, R., Bourgeron, T., Chung, W.K., Constantino, J.N., Eichler, E.E., Jacquemont, S., Miller, D.T., Mitchell, K.J., et al. (2020). Insufficient Evidence for "Autism-Specific" Genes. Am. J. Hum. Genet. *106*, 587–595.
3. Castellani, C.A., and Arking, D.E. (2020). High-Risk, High-Reward Genetics in ASD. Neuron *105*, 407–410.
4. Kaufmann, W.E., Kidd, S.A., Andrews, H.F., Budimirovic, D.B., Esler, A., Haas-Givler, B., Stackhouse, T., Riley, C., Peacock, G., Sherman, S.L., et al. (2017). Autism Spectrum Disorder in Fragile X Syndrome: Cooccurring Conditions and Current Treatment. Pediatrics *139* (*Suppl 3*), S194–S206.
5. Hammond, E.C., and Horn, D. (1954). The relationship between human smoking habits and death rates: a follow-up study of 187,766 men. J. Am. Med. Assoc. *155*, 1316–1328.

https://doi.org/10.1016/j.ajhg.2020.09.013.