Not all autism genes are created equal: A response to Myers et al

Joseph D. Buxbaum¹, David J. Cutler², Mark J. Daly³, Bernie Devlin⁴, Kathryn Roeder⁵, and Stephan J. Sanders⁶ for the Autism Sequencing Consortium

¹Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029;

²Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, 30322;

³Broad Institute of MIT and Harvard, Cambridge MA 02142 and Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114; ⁴Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213; ⁵Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, PA 15313; ⁶Department of Psychiatry, UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA 94158.

In our recent Autism Sequencing Consortium (ASC) study (Satterstrom et al 2020), we show that cohorts ascertained for autism spectrum disorder (ASD) and those ascertained for a broader array of neurodevelopmental disorders (NDDs) differ both in the rate and in the relative distribution of mutations across genes. In secondary analyses, we used a binary classifier to separate the 102 ASD genes that we identified into those more often disrupted in individuals ascertained for ASD versus those more often disrupted in individuals ascertained for NDDs; we refer to these as ASD-predominant (ASD_P) and ASD_{NDD} genes, respectively. Protein truncating variants in ASD_P genes are more likely to be found in parents, and are more likely to be inherited, consistent with reduced selective pressure acting against disruptive variants in ASD_P genes. Moreover, in individuals ascertained for ASD, those harboring ASD_{NDD} gene mutations show, on average, greater delays in walking and lower IQ, when compared to those harboring ASD_P gene mutations. The results demonstrate that there will be genes that are more clearly associated with the extremes of a clinical phenotypic spectrum. We suggest 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. Here, in response to a recent commentary by our esteemed colleagues (Myers et al. 2020), we discuss the ASC findings in greater detail and place our results in the context of future research in ASD.

Myers and colleagues (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 their concern. Included amongst the authors in Myers et al are leading medical geneticists and this is their bailiwick. However, to make their argument, Myers suggest that the ASC report presents evidence for “ASD-specific” genes; this is not the case, rather, the very notion of ASD-specific genes is Myers’ construction, not ours. In the ASC manuscript (see also Castellani and Arking 2020), which has a focus entirely different than that of Myers and different than what Myers construes it to be, we observe that, with current sample sizes, we cannot 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. Regrettably, the Myers commentary misconstrues one of our analyses and its purpose, as well as failing to appreciate other aspects of ASD research. That places their work firmly in our bailiwick.

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 these 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 are they 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 is atypical development 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 (Kaufmann et al., 2017). 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, say the problem cannot be solved without data that do not yet exist.

We thought differently, that perhaps there would be a path forward by comparing the mutational pattern in the ASC sample, ascertained for idiopathic ASD, to a cohort ascertained for more general developmental disorders, most often neurodevelopmental disabilities, which we will call NDD. In other words, was there evidence for genes that were more associated with cohorts ascertained for ASD than with cohorts ascertained for NDD, 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.”

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 are ASD genes by our criteria, 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 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 complains our classifier has an arbitrary threshold. It does; classifiers typically do. The key question we sought to answer is, 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 neurodevelopmental disabilities, 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 homogenous 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.”

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 ($p = 1.1 \times 10^{-4}$).

Based on this evidence, we find our classifier for ASD_P and ASD_{NDD} to be useful. Yet, we completely agree with Myers that this binary classifier has ‘dubious clinical significance.’ 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.

A few other issues require clarification. On the arbitrariness of the classifier, we find it surprising that Myers wants to substitute ours for another arbitrary classifier, namely significance testing. Scientists have generally agreed on an arbitrary 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. Thresholds are arguable and have

been debated more than we care to consider. Testing whether particular genes are classified correctly, however, as Myers suggests, is not meaningful in the context of the ASC analyses. 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. Furthermore, we take a dim view of the notion of using statistical significance and clinical significance interchangeably. Clinicians are concerned about false positive and false negative rates of a clinical test and rightly so. Significance tests have awful properties in this context.

Myers argues 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 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. Mis- or incomplete phenotyping, as they describe, will make the two cohorts more similar to each other than otherwise. 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 ID/DD cohorts. From the ASC perspective, since the two samples sets are demonstrably different, this limitation, while real, is of limited impact on the ASC conclusions and not in the direction that concerns them.

Furthermore, Myers claim that no gene could be labeled ASD_P unless, when it is mutated in individuals with ASD, their average IQ = 100. In other words, on average, the gene has no impact on intellectual function. Let's set aside how the ASC defined ASD_P genes, which has nothing to do with IQ, and simply look at the biology of ASD. Consider the Simons Simplex Collection of autism families, where virtually all ASD participants, one per family, had an IQ test. If we remove all participants with damaging mutations in ASD genes, the average IQ of the cohort falls far below 100. Even if we remove participants with any mutation thought to be damaging for any NDD, the average IQ is far less than 100. We would be remiss if we did not note that some individuals with ASD have exceptionally high IQ. Yet, there is something about ASD cohorts that, on average, leads to lower IQ than the population average. Perhaps there is some relationship between the social functions assessed for ASD and cognitive function? We take this subject seriously because, on average, individuals with ASD also harbor genetic variation that is not only common in the population, it occurs more commonly in the portion of the population with above average IQ. These are important observations that are difficult to reconcile and, yet, are very relevant to understanding the pathobiology of autism. They also potentially have broad implications for our society; for example, how might we tap this potentially enhanced cognitive function arising from common variation risk for ASD to optimize education in ASD?

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 towards 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 point, that of rational study design. If there were no difference between ASD- and NDD-ascertained cohorts it would be more efficient to study NDD cohorts, where mutations per subject are higher, and phenotyping potentially less costly. One could ask what is the justification for large-scale collection and genetic analysis of ASD-ascertained samples? Note that, amongst the Myers team, several are focused on developing and genetically analyzing such ASD-ascertained cohorts and it would be interesting to hear their views on this. Early in our analyses, we learned from our funders that the question of ‘specificity’ in our findings had been raised, and we were asked to further examine this question. Having demonstrated clear differences between ASD- and NDD-ascertained cohorts, and providing evidence for the existence of ASD_P and ASD_{NDD} genes, we can say with confidence that experiments to identify the individual genes are worth pursuing and, as noted above, we can begin to define best approaches to the discovery of ASD_P genes.

In summary, for all the reasons outlined above, we agree with Myers *et al.* about the lack of clinical utility of “ASD-specific” panels and that there is “currently insufficient evidence to establish ASD-specificity of any genes.” However, our results 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, towards attaining these goals.

References

Castellani CA, Arking DE (2020) High-Risk, High-Reward Genetics in ASD. *Neuron*. **105**:407-410.

Kaufmann WE, Kidd SA, Andrews HF, Budimirovic DB, Esler A, Haas-Givler B, Stackhouse T, Riley C, Peacock G, Sherman SL, Brown WT, Berry-Kravis E (2017) Autism Spectrum Disorder in Fragile X Syndrome: Cooccurring Conditions and Current Treatment. *Pediatrics* **139**(Suppl 3):S194-S206

Myers SM, Challman TD, Bernier R, Bourgeron T, Chung WK, Constantino JN, Eichler EE, Jacquemont S, Miller DT, Mitchell KJ, Zoghbi HY, Martin CL, Ledbetter DH (2020) Insufficient Evidence for “Autism-Specific” Genes. DOI: 10.31219/osf.io/24gm3

Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L, Stevens C, Reichert J, Mulhern MS, Artomov M, Gerges S, Sheppard B, Xu X, Bhaduri A, Norman U, Brand H, Schwartz G, Nguyen R, Guerrero EE, Dias C; Autism Sequencing Consortium; iPSYCH-Broad Consortium, Betancur C, Cook EH, Gallagher L, Gill M,

Sutcliffe JS, Thurm A, Zwick ME, Børglum AD, State MW, Cicek AE, Talkowski ME, Cutler DJ, Devlin B, Sanders SJ, Roeder K, Daly MJ, Buxbaum JD (2020) Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**:568-584.