Autism spectrum disorders and brain volume link through a set of mTOR-related genes

Arenella, Martina\textsuperscript{1,2,3}, Mota, Nina R.\textsuperscript{1,3}, Teunissen, Mariel W.A.\textsuperscript{1,4}, Brunner, Han G.\textsuperscript{1,5,6}, Bralten, Janita\textsuperscript{1,3}

2. Department of Forensic and Neurodevelopmental Science, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, The United Kingdom
3. Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
5. Department of Clinical Genetics, Maastricht University Medical centre, Maastricht, The Netherlands.
6. GROW School of development and Oncology, and MHENS school of Neuroscience, Maastricht University, the Netherlands

Word count: 5702 + 225 = 5927

The authors declare no conflicts of interest.

Abbreviated title: Autism and brain volume link through mTOR-related genes
Abstract

Background
Larger than average head and brain sizes are often observed in individuals with Autism Spectrum Disorders (ASDs). ASDs and brain volume are both highly heritable, with multiple genetic variants contributing. However, it is unclear whether ASDs and brain volume share any genetic mechanisms. Genes from the mammalian target of rapamycin (mTOR) pathway influence brain volume, and variants are found in rare genetic syndromes that include autism spectrum disorder features. Here we investigated whether variants in mTOR-related genes are also associated with ASDs and if they constitute a genetic link between large brains and ASDs.

Methods
We extended our analyses between large heads (macrocephaly) and rare de novo mTOR-related variants in an intellectual disability cohort (N=2,258). Subsequently using Fisher’s exact tests we investigated the co-occurrence of mTOR-related de novo variants and ASDs in the denovo-db database (N=23,098). We next selected common genetic variants within a set of 96 mTOR-related genes in genome-wide genetic association data of ASDs (N=46,350) to test gene-set association using MAGMA. Lastly, we tested genetic correlation between genome-wide genetic association data of ASDs (N=46,350) and intracranial volume (N=25,974) globally using LD-score regression as well as mTOR-specific by restricting the genetic correlation to the mTOR-related genes using GNOVA.

Results
Our results show that both macrocephaly and ASDs occur above chance-level in individuals carrying rare de novo variants in mTOR-related genes. We found a significant mTOR gene-set association with ASDs (p=0.0029) and an mTOR-stratified positive genetic correlation
between ASDs and intracranial volume ($p=0.027$), despite the absence of a significant genome-wide correlation ($p=0.81$).

Conclusions

This work indicates that both rare and common variants in mTOR-related genes are associated with brain volume and ASDs and genetically correlate them in the expected direction. We demonstrate that genes involved in mTOR signalling are potential mediators of the relationship between having a large brain and having ASDs.

*Keywords*: autism spectrum disorders; genetics; brain volume; mTOR; stratified genetic correlation.
Background

Autism spectrum disorders (ASDs) are a set of common neurodevelopmental conditions that are clinically characterized by atypical social communication and interaction, the presence of rigid, repetitive, and stereotyped behaviors, and atypical sensory processing (American Psychiatric Association, 2013). Heterogeneity at the phenotypic and etiological level in ASDs make it difficult to pinpoint the underlying biological mechanisms and to make steps towards better targeted (i.e., personalized) treatment options. One of phenotypic heterogeneities in ASDs is observed at the level of head size, with a larger than average head size being reported in some individuals with ASDs in the early years of life (Sacco et al., 2015). These individuals may present with macrocephaly and/or head circumference above the normal range (i.e., occipitofrontal circumference (OFC) > two standard deviations above the mean). The occurrence of large head circumference in ASDs is estimated to range between 12-15% and correlates with clinical severity (Fidler et al., 2000). Because head circumference is correlated with brain volume (Lindley et al., 1999), it may reflect a larger than average brain size. Previous studies indicate that individuals with ASDs have larger estimates of intracranial volume, total grey matter volume and cortical thickness in comparison with their matched controls (Van Rooij et al., 2018; Vonder Haar et al., 2017). The molecular mechanisms that may contribute to the observed variation in brain and head size in ASDs are not defined yet.

Previous studies suggest that genetic factors may play a role. Large heads are not only observed in individuals with ASDs but have been reported in ~15% of their first-degree relatives (Butler et al., 2005). Both ASDs and brain volume have moderate-to-high heritability (h²) estimated around ~70-90% (Tick et al., 2016) and ~65% (Zhao et al., 2019) and their genetic architecture is complex and multifactorial – including both common and rare genetic variants. The question remains if these two traits share some degree of genetic liability. A
previous study investigated the genetic correlation between ASDs and brain volume on a genome-wide scale and found no significant genetic correlation (Adams et al., 2016). However, it is possible that these two traits are connected through specific genetic variants and that such effects are diluted in global genetic correlation analyses (Werme et al., 2021).

Genetic studies of brain phenotypes report associations of multiple genes (Adams et al., 2016; Grasby et al., 2020), including genes controlling mTOR signalling (Licausi & Hartman, 2018), (Kim, 2015). Genes belonging to the mTOR signalling pathway are known to mediate general cell growth, proliferation, and survival (Yu & Cui, 2016) and are organized in two distinct branches, the RAS-MAPK sub-branch and the PI3K-AKT sub-branch (Kanehisa & Goto, 2000; Reijnders et al., 2017). During neurodevelopment, these mTOR-related genes also influence the proliferation of neuronal progenitor cells and are key contributors to the genesis of neurons and their reciprocal communication (Licausi & Hartman, 2018). The PI3K-AKT mTOR sub-branch controls the size of the soma and dendrite tree, whereas its interaction with the RAS-MAPK sub-branch allows the arborization of dendrites and the integration of synapses throughout the brain (Costa-Mattioli & Monteggia, 2013), (Kumar et al., 2005). Animal studies showed that dysregulation of mTOR signalling may drive excessive brain growth, that is consistent with a halt of neuronal apoptosis and consequent alterations in normal synaptic pruning (Huber et al., 2015). In a previous study de novo variants in a set of 101 mTOR-related genes were associated with the occurrence of macrocephaly in individuals with intellectual disability (ID). Moreover, common variants in this mTOR gene-set showed an association with intracranial volume (ICV) in the general population (Reijnders et al., 2017).

Some of the mTOR-related genes have also been implicated in ASDs (Onore et al., 2017). For example, animal studies showed that variants in the mTOR-related gene PTEN drive social impairment, a core feature of ASDs, in rodents (Tilot et al., 2016). Variants in PTEN, MTOR
and PIK3CA, which are also mTOR-related genes, have been encountered in individuals with rare syndromes that comprise ASDs, and, in some instances, also macrocephaly (Klein et al., 2013), (Yeung et al., 2017). Besides rare syndromes, the mTOR-related genes have also been linked to idiopathic ASDs. The Simons Foundation Autism Research Initiative (SFARI) defined a list of genes based on both rare and idiopathic ASD studies and showed that many of these ASD-associated genes belong to the mTOR and the RAS-MAPK signalling gene-sets (Wen et al., 2016). Research on common genetic variants further indicated an over-representation of mTOR-related genes among ASD-associated genetic variants (The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium, 2017).

Taken together, there is evidence of the occurrence of increased head and brain size in ASDs and studies that investigated brain volume and ASDs separately both pointed to the potential contribution of mTOR-related genes. There is however no study that investigates whether genetic variants in mTOR-related genes represent a genetic link between ASDs and a tendency towards having a large head and brain. Therefore, the present study aims to investigate the specific role of genes from the mTOR gene-set in the relationship between brain volume and ASDs. We first explore locally collected data and publicly available data to investigate the incidence of macrocephaly and ASDs in individuals carrying rare de-novo variants in the set of 101 mTOR-related genes defined by Reijnders et al. 2017. Next, using common genetic variant data, we investigated whether the same mTOR gene-set, that previously showed an association with ICV, was also associated with ASD. Lastly, using common variant data for both ASD and ICV, we tested with a stratified genetic covariance analysis if the mTOR gene-set constitutes a specific genetic link between ASD and ICV variability and if this correlation was in the expected direction.
Methods

The set of mTOR-related genes was selected based on the mTOR pathway representation in the Kyoto Encyclopaedia of Genes and Genomes (Kanehisa & Goto, 2000), with two sub-pathways: the PI3K-AKT-mTOR pathway and the RAS-MAPK-mTOR pathway (as defined by Reijnders et al., 2017). The list of mTOR-related genes was derived from three authoritative reviews of the mTOR regulators (Laplante & Sabatini, 2012; Shimobayashi & Hall, 2014; Sun & Hevner, 2014). Through literature search, protein complexes were mapped to single proteins and genes. The final list contains 101 mTOR-related genes: 96 genes map on autosomes, and 5 genes belong to the X-chromosome, with 60 genes belonging to the PI3K-AKT pathway and 76 genes to the RAS-MAPK pathway (Table S1 for further details).

Following-up on previous work by Reijnders et al., 2017, we acquired additional exome data to reach a total of 2,258 patients with intellectual disability (ID). We aimed to confirm the enrichment of macrocephaly in patients carrying mTOR-related de novo variants. The participants had undergone diagnostic trio whole exome sequencing (WES) at the Radboud university medical center and the Maastricht University Medical Centre. Level of ID ranged between mild (ID 50-70) and severe (IQ<30). Patients were classified as macrocephalic (OFC > 2.5 SD), normocephalic (2.5 SD < OFC < 2.5 SD) or microcephalic (OFC < 2.5 SD). Diagnostic WES was approved by the local medical ethics committee. Written informed consent was obtained from all the families. Patient-parent trio exomes were sequenced, using DNA isolated from blood at the Beijing Genomics Institute in Copenhagen. Exome capture was performed using Agilent SureSelect v4 and v5 and samples were sequenced on an Illumina HiSeq 4000 platform with paired-end reads to a median target coverage of 112x. Sequence reads were aligned to the hg19 reference genome using BWA (v0.7.12). Variants were subsequently called by the GATK haplotypecaller (v3.4-46). For further details, see Supplementary Information. We performed a Fisher’s exact test in R to establish if the
occurrence of macrocephaly was significantly higher in mTOR-related variant carriers compared to chance, setting the significance threshold at 0.05. Under- or over-enrichment of macrocephaly among mTOR-related variant carriers was tested based on the cumulative distribution function of the hypergeometric distribution using the webtool of the Graeber Lab, https://systems.crump.ucla.edu/hypergeometric/. Similarly, we tested specific enrichment for de novo variants in genes of the RAS-MAPK and PIK3-AKT gene sub-branches separately.

To understand whether de novo variants in the mTOR gene-set have been reported in ASDs, we exploited an online database of de novo variants, the de-novo-db database (http://denovo-db.gs.washington.edu/denovo-db/). This database includes 40 independent studies on psychiatric and neural phenotypes providing a list of genome-wide de novo variants detected in 23,098 individuals. All participants in each study provided written informed consents and agreed with the use of their data for research purposes. We filtered the de-novo-db data on participants diagnosed with “autism” (the definition used in de-novo-db) as primary phenotype and controls. We used the publicly available data that did not include data derived from the Simon Complex Sample Collection. For further details about diagnostic criteria and cohort descriptives in the de-novo-db see (Turner et al., 2017) and http://denovo-db.gs.washington.edu/denovo-db/. We counted de novo variants in the mTOR gene-set and in genes outside the set in both individuals with autism and in controls. As background we used genes throughout the genome and subsequently genes expressed in the brain (based on annotations from eMAGMA(Gerring et al., 2021). We built a 2x2 contingency table (Table S2) and performed a Fisher’s exact test to assess the (in)dependence between mTOR-related de novo variants and autism. Analyses were performed in R and the significance threshold was set at 0.05. Under- or over-enrichment of autism among carriers of variants in mTOR-related genes was tested based on the cumulative distribution function of the hypergeometric distribution using the webtool from the Graeber Lab.
As sensitivity analyses, we tested the enrichment of variants in genes from the RAS-MAPK sub-branch and the PIK3-AKT subbranch respectively in individuals with autism and in controls.

Next, we investigated if common genetic variants within the mTOR gene-set were linked to ASDs using the summary statistics resulting from the ASD genome-wide association study (GWAS) meta-analysis conducted by the Psychiatric Genomic Consortium (PGC) and the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH) that included 18,381 ASD cases and 27,969 controls (Grove et al., 2019; see Supplementary information). All the individuals included in the PGC and iPSYCH cohort samples provided written informed consent. We tested the full set of 96 autosomal mTOR-related genes, and post-hoc looked at the PI3K-AKT and the RAS-MAPK gene-sets. Gene-set analyses were performed using the Multi Maker Annotation for Genomic Annotation (MAGMA) software package (version 1.09) (de Leeuw et al., 2015) using the SNP-wise mean model that combines \( p \)-values of the gene-related SNPs, controlling for linkage disequilibrium (LD). In addition, we performed conditional analyses, where we conditioned on brain expressed genes considering that these genes are enriched for common variants associated with brain-related traits, such as ASDs and ICV (Brien et al., 2018). Gene-set results were considered significant if they reached the Bonferroni-corrected \( p \)-value threshold of \( p < 0.025 \) (i.e., corrected by the number of tests).

To investigate the genetic link between ASDs and ICV we analysed the summary statistics from the ASD-GWAS meta-analysis (18,381 ASD cases and 27,969 controls) (Grove et al., 2019) with the summary statistics resulting from the ICV-GWAS meta-analysis conducted by the Enhancing Neuro-Imaging Genetics Through Meta-Analysis (ENIGMA) and the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortia that included 25,974 individuals (Adams et al., 2016) and see Supplementary Information. All the
individuals included in the ENIGMA and CHARGE cohort samples provided written informed consent. We estimated the genetic correlation between the summary statistics of the PGC-ASD GWAS and the ENIGMA-ICV GWAS using LD bivariate score (LDSC) regression as implemented in the LDSCORE regression software (Bulik-Sullivan et al., 2015). Analyses used pre-computed LD scores based on the 1000 Genome Project reference panel that well-suits European-centred GWASs (available at https://github.com/bulik/ldsc). LDSC-based analysis consisted of two steps: 1) converting summary statistics data to the LDSC format (data filtering and merging to the reference panel as shown in https://github.com/bulik/ldsc/wiki/Heritability-and-Genetic-Correlation), 2) estimating genetic correlation. The regression model did not include any regression intercept given the absence of sample overlap between the investigated datasets. A block jack-knife procedure over SNP data was used to estimate standard errors and calculate corresponding $p$-values. Results were considered significant at $p<0.05$.

To estimate whether ASDs and ICV were genetically correlated through mTOR-related genes and if so in which direction, we used the GeNetic cOVariance Analyzer (GNOVA) method (https://github.com/xtonyjiang/GNOVA) (Lu et al., 2017). The GNOVA approach allows to stratify the genetic covariance between two traits of interest by pre-defined functional annotation, in our case the mTOR gene-set. In addition, it provides stratified correlation results by normalising the covariance estimates. We created annotations for the full mTOR gene-set using the make_annot.py script available at (https://github.com/bulik/ldsc/). Then, using the PGC-ASD and ENIGMA-ICV summary statistics as input we performed the stratified genetic correlation analysis. Subsequently, we used the annotation for the two mTOR sub-branches and respectively tested them in secondary GNOVA-based ASD-ICV correlation analyses.
**Results**

In the local ID cohort, we identified 39 *de novo* variants in mTOR-related genes. Macrocephaly was observed in 33% of the patients with mTOR-related variants, while it was reported in 6% of the patients with *de novo* variants in genes outside the mTOR-related gene-set (*Table 1*).

---*Table 1***

The Fisher’s exact test indicated that macrocephaly was significantly more frequent in carriers of *de novo* variants in mTOR-related genes than individuals carrying variants in genes outside the mTOR gene-set (OR= 8.9; CI: 3.9-19.61; *p* = 1.7e-7). The hypergeometric test also indicated an over-enrichment of macrocephaly in carriers of *de novo* mTOR-related variants (5.03-fold change, *p*-value = 5e-07). We did not observe differences in the frequency of *de novo* variants between the RAS-MAPK and PI3k-AKT gene sub-branches.

We observed microcephaly in 17% of the patients carrying *de novo* variants in mTOR-related genes. However, the frequency of microcephaly did not statistically differ between carriers of variants in mTOR-related genes and carriers of variants in genes not included in the mTOR gene-set (OR= 2.2; CI = 0.77-5.53; *p*=0.08, see *Table S2*).

Exploring the *de-novo-db* database, we counted a total of 34,389 *de novo* variants (7,623 in brain-expressed genes) reported across individuals with autism and controls. Of these variants, 135 variants (32 for brain-expressed genes) occurred in genes belonging to the mTOR gene-set and 97% (132/135) were reported in individuals with autism. Fisher’s exact test indicated a significantly higher incidence of *de novo* variants in the genes of the mTOR gene-set in individuals with autism than in controls (OR=2.06 CI: 1.90-2.24; *p*=2.2e-16; positive enrichment fold 1.06; hypergeometric *p*-value=0.01). Analyses restricted to the brain expressed genes showed a positive but non-significant enrichment of mTOR-related *de novo*
variants in autism as compared to variants in other brain expressed genes (enrichment fold 1.02; \( p > 0.5 \)) (Table S4). Sensitivity analyses on the two sub-branches indicated a significant enrichment of variants in genes of the RAS-MAPK sub-branch in autism (enrichment fold = 1.06, \( p = 0.01 \)), but not for variants in the PIK3-AKT sub-branch genes (enrichment fold = 1.02; \( p = 0.06 \)). There was no significant enrichment of variants in either of the two sub-branches in autism when restricting our analyses to genes expressed in the brain (Table S4).

For further details about the \textit{de novo} variants in mTOR-related genes from \textit{de-novo-db} see Table S3.

MAGMA-based gene-set analysis showed a significant association of the full set of 96 autosomal mTOR-related genes with both ASDs (\( \beta(\text{SD})=0.23(0.08); \ p=0.0033 \)) and ICV (\( \beta(\text{SD})=0.17(0.08); \ p=0.021 \)). Conditioning analyses on brain expressed genes confirmed these associations (ASDs: \( \beta=0.22; \ p=0.003 \), ICV: \( \beta=0.17 \); \( p=0.02 \)). Post-hoc analyses on the two mTOR-related sub-branches indicated a significant association of ASDs with the RAS-MAPK pathway (\( \beta=0.28; \ p=0.002 \); conditioned on brain-expressed \( p=0.001 \)) and of ICV with the PI3K-AKT pathway (\( \beta=0.29; \ p=0.00425 \); conditioned on brain-expressed \( p=0.00434 \))(Table 2).

\textit{--Table 2--}

The LDSC-based analysis indicates no significant genome-wide genetic correlation between ASDs and ICV (\( \rho_g=0.040, \ p=0.81 \)).

The GNOVA-based analysis of genetic correlation stratified by the mTOR-related gene-set revealed a significant positive genetic correlation between ASDs and ICV (\( \rho_g=0.001; \ p=0.027 \)). Subsequent post-hoc analyses revealed that the genetics of ASDs and ICV were significantly positively correlated when stratified by the RAS-MAPK pathway (\( \rho_g=0.001; \ p=0.002 \)), but not when stratified by the PI3K-AKT pathway (\( \rho_g=0.0004; \ p=0.10 \)) (Table 3).
Discussion

By relying on multiple data sources and a combination of analytical tools, our results support the hypothesis that genes involved in mTOR signalling constitute a genetic link between ASDs and brain volume. The observed mTOR-gene-set-based correlation between ASDs and brain volume occurs in the expected positive direction in line with the phenotypic association between large brain and head size and ASDs.

We extended and confirmed the over-representation of rare mTOR-related de novo variants in individuals with macrocephaly in line with previous work (Reijnders et al., 2017) using a subset of our local ID cohort and we found that de novo variants in mTOR-related genes are over-represented in individuals with autism. Additionally, we show that common variants in the mTOR gene-set are, next to being associated to ICV (Reijnders et al., 2017), also associated with ASDs. Lastly, we demonstrated that common variants in the mTOR gene-set mediate a positive genetic correlation between ASDs and ICV, while no ASDs-ICV correlation was detected at a genome-wide level (Figure 1).

Figure 1

There is evidence for a phenotypic correlation between ASDs and large brains (Sacco et al., 2015), yet previous studies did not find significant genetic correlations between ASDs and brain volume (Grove et al., 2019). Our data support a more specific, positive genetic correlation through mTOR-related genes between ASDs and ICV, indicating an overlap between alleles that increase the odds of having ASDs and having a large brain. This provides a potential mechanistic underpinning of the clinical association of large brains in individuals with ASDs. These results highlight that stratification of genetic correlations can reveal that complex traits are linked specifically through functionally defined portions of the
genome (Lu et al., 2017) and that this signal can be diluted when we look at the genetic correlation of the complete genome. Genetic stratification is particularly informative for those scenarios where genetic correlation between loci is confined to a specific genomic region, or when genetic factors show opposite effects at different loci that would cancel each other out in a standard, global-scale genetic correlation analyses (Werme et al., 2021). Dissecting the genetic relationship by function has the potential to delineate shared biological mechanisms between complex traits, such as brain volume and ASDs.

Our research focused on a specific set of mTOR-related genes, involving both RAS-MAPK and PI3K-AKT signalling, which have been studied separately in previous studies on ASDs and brain volume. The ENIGMA consortium, relying on the same ICV-GWAS data included in our study, showed that the top ICV-associated genetic variants were enriched for genes annotated to PI3K-AKT signalling (Adams et al., 2016). In line with those findings, we show here that the complete PI3K-AKT gene-set is associated with ICV. On the other hand, bioinformatic analysis of the ASDs risk genes from the SFARI database revealed an overrepresentation of genes within the RAS-MAPK mTOR sub-branch (Wen et al., 2016). Here we were able to expand on those studies by showing the positive genetic link between ASDs and ICV specifically for the RAS-MAPK pathway, and not for the PI3K-AKT pathway. These results were confirmed by our findings of a significant enrichment of de novo variants in the RAS-MAPK pathway in ASDs. Considering the functional role of these pathways, it is possible that the co-occurrence of ASDs and large ICV is explained by RAS-MAPK-mediated increased dendritic growth and spine formation, key to synaptic communication and regulated by the RAS-MAPK pathway (Kumar et al., 2005).

Our findings inform the need for further research on the mTOR signalling pathway as a potential target for treatment in ASDs. Previous studies already investigated mTOR inhibitors (e.g., rapamycin) in ASDs based on the evidence of up-regulation of the mTOR pathway in
monogenic syndromes that include ASDs as well as in animal models (Schneider et al., 2017). Specifically, mTOR-related inhibitors appear effective in the rescue of neuronal and behavioural alterations across a range of ASD-related syndromes. Recently, Ganesan et al. (2019), reviewed the possibility to translate the use of mTOR inhibitors to the treatment of idiopathic ASDs, and highlight the downstream regulators of the mTOR signalling as valuable molecular targets (Ganesan et al., 2019). Our data support research on these mTOR inhibitors also to compounds that modulate RAS-MAPK signalling, in rare as well as in common, non-syndromic forms of ASDs, especially in the presence of increased brain size.

Our study has some limitations. We did not have access to a cohort where we could assess the enrichment of mTOR-related \textit{de novo} variant jointly in macrocephaly and ASDs. Therefore, these analyses were performed in two separate, phenotypically distinct cohorts. We used the largest most powerful meta-analyses for ASD genetics, that include all genome-wide genetic datasets of ASDs. Therefore, there is no availability of an independent replication dataset. The consistency between different datasets, however all point to a link between mTOR-related genes and ASDs and/or brain size. We are limited by the current knowledge and understanding of mTOR signalling and its contributing genes for gene-set and gene-set stratified analyses. Lastly, our results cannot claim causality, as to translate genetic association findings to more mechanistic molecular understanding, it would require the use of \textit{in vitro} or \textit{in vivo} models. Such experiments would help to establish if variation in this mTOR gene-set indeed represent a causal factor for large brains in ASDs.

Our findings demonstrate that a set of mTOR-related genes associate with both ASDs and brain volume, and that variation in mTOR underlies a positive genetic correlation between them. Our findings suggest that the mTOR pathway is a biological mediator of the phenotypic association of ASDs and large brains.
Corresponding author: Janita Bralten, Department of Human Genetics, Radboud University Medical Center, Geert Grooteplein 10, route 855, 6500 HB, Nijmegen, The Netherlands; Tel. + 31 24 361 40 17; Fax + 31 24 366 87 52; Janita.Bralten@radboudumc.nl
References


Tables

Table 1. Intellectual disability cohort descriptives

Table 2. MAGMA-based gene-set analyses in autism spectrum disorders and intracranial volume

Table 3. Genetic correlation analyses between autism spectrum disorders and intracranial volume
Figure 1. A phenotypic correlation between head and brain size and autism spectrum disorders (ASDs) is known, however no genome-wide genetic correlation was found between ASDs and intracranial volume (ICV). In this study we show that both rare and common variants in the mTOR gene-set link to brain volume and ASDs, and that stratifying the genetic correlation to variants within the mTOR gene-set leads to a significant positive genetic correlation between ASDs and ICV.

Supplementary tables (excel spreadsheet)

Table S1. List of genes included in the full mTOR gene-set and its sub-branches

Table S2. De novo variants in genes of the mTOR gene-set in the ID cohort

Table S3. De novo variants in mTOR-related genes in individuals with autism from the de-novo-db

Table S4. Sensitivity analyses of de novo variants in individuals with autism for the two mTOR sub-branches and restricted to brain-expressed genes in the de-novo-db

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 847879 (PRIME, Prevention and Remediation of Insulin Multimorbidity in Europe). This publication is part of the project ‘No labels needed: understanding psychiatric disorders based on genetic traits’ (with project number 09150161910091) of the research program Veni, which is partly financed by the Dutch Research Council (NWO) Health Research and Development (ZonMW). The analyses were carried out on the Dutch national e-infrastructure, and it is part of the research programme Computing Time National Computing Facilities Processing Round pilots 2018 with project No. 17666, which is (partly) financed by the Dutch Research Council (NWO).
Brunner H.G. is in part supported by the Solve-RD project that has received funding from European Community’s Horizon 2020 research and innovation programme under agreement No. 779257. Mota R.N. is supported by European Community’s Horizon 2020 programme (H2020/2014-2020) under agreement No. 667302 (CoCA) and by funding from the Dutch National Science Agenda NeurolabNL project (grant 400-17-602).

**Ethical considerations**

For the local ID cohort, diagnostic WES was approved by the medical ethics committees of Radboudumc (Commissie Mensegebonden Onderzoek; registration number 2011-188). Written informed consent was obtained from all the families. Other datasets used in the current manuscript are based on publicly available datasets. In de-novo db all participants in each study provided written informed consents and agreed with the use of their data for research purposes. All the individuals included in the PGC, iPSYCH, ENIGMA and CHARGE cohort samples provided written informed consent.

---

**Key points**

- Autism spectrum disorders (ASDs) and brain volume show a phenotypic association, but no genetic correlation has been reported
- We show that rare and common variants in genes from the mTOR signalling pathway associate to both ASDs and brain volume
- Using a stratified genetic correlation approach, restricting to variants within mTOR-related genes, we show a positive genetic correlation between ASDs and brain volume
- By dissecting the genetic relationship by function we can delineate potential biological mechanisms between complex traits, like ASDs and brain volume.

---

**Table 1. Intellectual disability cohort descriptives**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>N</th>
<th>Age (mean)</th>
<th>Sex</th>
<th>Macrocephalic (OFC &gt; + 2SD)</th>
<th>Normocephalic (-2SD &lt; OFC &lt; + 2SD)</th>
<th>Microcephalic (OFC &lt; -2SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>2258</td>
<td>9</td>
<td>43.8% female</td>
<td>163</td>
<td>1,707</td>
<td>388</td>
</tr>
</tbody>
</table>
Patients with *de novo* variants

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Sample Size</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>mTOR-related</td>
<td>52%</td>
<td>13</td>
<td>19</td>
<td>7</td>
<td>39</td>
</tr>
<tr>
<td>de novo</td>
<td>48%</td>
<td>142</td>
<td>1,712</td>
<td>290</td>
<td>2144</td>
</tr>
</tbody>
</table>

OFC = occipitofrontal cortex, N = sample size, OFC >+ 2SD = occipitofrontal cortex bigger than 2 times the standard deviation from the mean, -2SD < OFC < +2SD = occipitofrontal cortex between 2 times the standard deviation bigger and smaller than the mean, OFC >-2SD = occipitofrontal cortex smaller than two times the standard deviation away from the mean.

Table 2. MAGMA-based gene-set analyses in autism spectrum disorders and intracranial volume

<table>
<thead>
<tr>
<th>Trait</th>
<th>Gene-set</th>
<th>Beta</th>
<th>SD</th>
<th>P</th>
<th>Cond-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASDs</td>
<td>Full mTOR</td>
<td>0.23</td>
<td>0.08</td>
<td>0.0029*</td>
<td>0.0033*</td>
</tr>
<tr>
<td></td>
<td>PI3K-AKT</td>
<td>0.17</td>
<td>0.11</td>
<td>0.059</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>RAS-MAPK</td>
<td>0.28</td>
<td>0.09</td>
<td>0.0014*</td>
<td>0.0016*</td>
</tr>
<tr>
<td>ICV</td>
<td>Full mTOR</td>
<td>0.17</td>
<td>0.08</td>
<td>0.021*</td>
<td>0.021*</td>
</tr>
<tr>
<td></td>
<td>PI3K-AKT</td>
<td>0.29</td>
<td>0.11</td>
<td>0.0042*</td>
<td>0.0044*</td>
</tr>
<tr>
<td></td>
<td>RAS-MAPK</td>
<td>0.14</td>
<td>0.09</td>
<td>0.07</td>
<td>0.07</td>
</tr>
</tbody>
</table>

ASDs = autism spectrum disorders, ICV = intracranial volume, SD = standard deviation, P = p-value, Cond-p = conditional p-value * Bonferroni-corrected significant p-values.

Table 3. Genetic correlations between autism spectrum disorders and intracranial volume

<table>
<thead>
<tr>
<th>Genetic annotation</th>
<th>rg</th>
<th>SD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full genome</td>
<td>0.04</td>
<td>0.05</td>
<td>0.81</td>
</tr>
<tr>
<td>mTOR gene-set</td>
<td>0.001</td>
<td>0.0004</td>
<td>0.027*</td>
</tr>
<tr>
<td>RAS-MAPK sub-branch</td>
<td>0.0012</td>
<td>0.0003</td>
<td>0.002*</td>
</tr>
<tr>
<td>PI3K-AKT sub-branch</td>
<td>0.00040</td>
<td>0.00024</td>
<td>0.105</td>
</tr>
</tbody>
</table>

Rg = genetic correlation, SD = standard deviation, p = p-value * Bonferroni-corrected significant p-values.
Supporting information

ASD IPSYCH - PGC GWAS sample

Summary statistics of the ASD GWAS came from the meta-analysis of the iPSYCH ASD sample and the PGC ASD samples. These samples consist of individuals of Caucasian, European ancestry. Summary statistics from the meta-analysis can be retrieved through the PGC website [https://www.med.unc.edu/pgc/about-us/](https://www.med.unc.edu/pgc/about-us/).

The iPSYCH ASD sample is a population-based case-control sample extracted from a baseline cohort consisting of all children born in Denmark between 1981 and 2005. Cases were identified from the Danish Psychiatric Central Research Register (DPCRR), which includes data on all individuals treated in Denmark at psychiatric hospitals. ASD diagnosis was based on ICD10 criteria, and include diagnosis for childhood autism, atypical autism, Asperger’s syndrome, other pervasive neurodevelopmental disorders and/or unspecified. Children without an ASD diagnosis were defined as controls. The PGC ASD samples includes 5 ASD case-control samples: The Geschwind Autism Center of Excellence (ACE; \( n = 391 \)), the Autism Genome Project (AGP; \( n = 2272 \)), the Autism Genetic Resource Exchange (AGRE; \( n = 974 \)), the Montreal/Boston Collection (MONBOS; \( n = 1396 \)), and the Simons Simplex Collection (SSC; \( n = 2231 \)). For further details, see the PGC website: [https://www.med.unc.edu/pgc/files/resultfiles/PGCASDEuro_Mar2015.readme.pdf](https://www.med.unc.edu/pgc/files/resultfiles/PGCASDEuro_Mar2015.readme.pdf). Here, only the CEU subset was selected based on the results of principal component analyses. All participants provided written informed consent.

ICV ENIGMA-CHARGE cohort

ICV GWAS data were based on the analysis of 25,974 individuals from
46 independent studies belonging to the Enhancing Neuroimaging Genetics Through Meta-analysis (ENIGMA) consortium (N= 13,171) and Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium (N = 12,803). The summary statistics of the meta-analysis of these two projects can be retrieved through the ENIGMA website http://enigma.ini.usc.edu/research/download-enigma-gwas-results/

The ENIGMA consortium includes samples with neuroimaging data in a range of neurological and neuropsychiatric disorders and in control populations. The CHARGE consortium includes mainly population-based studies investigating the underpinnings of health related traits and age-related disorders. Only individuals of Caucasian, European ancestry were taken along for analyses. All participants provided written informed consent.

**ID cohort diagnostic WES analyses**

Since September 2011 whole exome sequencing (WES) is part of the routine diagnostic work-up aimed at the identification of the genetic causes underlying intellectual disability (ID) diseases at the Radboud University Medical Center and Maastricht University Medical Center. For individuals with unexplained ID, a family-based WES approach is used which allows the identification of de novo variants as well as variants segregating according to other types of inheritance, including recessive variants and maternally inherited X-linked recessive variants in males. For this study, we selected all individuals with ID who had data from family-based WES in the time period of 2013-2018. This selection yielded a set of 2,418 individual probands, including 1040 females and 1378 males across 2387 different families. Of these, 161 were excluded because of lack of data, no ID or termination of pregnancy. The level of ID ranged between mild (IQ 50-70) to severe-profound (IQ<30). Families gave informed consent for both the diagnostic procedure as well as for forthcoming research that
could result in the identification of new genes underlying ID by meta-analysis. The exomes of 2,257 patient-parent trios were sequenced, using DNA isolated from blood, at the Beijing Genomics Institute (BGI) in Copenhagen. Exome capture was performed using Agilent SureSelect v4 and v5 and samples were sequenced on an Illumina HiSeq 4000 instrument with paired-end reads to a median target coverage of 112x. Sequence reads were aligned to the hg19 reference genome using BWA version v0.7.12 and duplicate marking by Picard v1.90. Variants were subsequently called by the GATK haplotypcaller (version v3.4-46). The diagnostic WES process as outlined above only reports (de novo) variants that can be linked to the individuals’ phenotype. In this study, we systematically collected all de novo variants in regions in or close to (200bp) a capture target. De novo variants were called as described previously. Briefly, variants called within parental samples were removed from the variants called in the child. For the remaining variants pileups were generated from the alignments of the child and both parents. Based on pileup results variants were then classified into the following categories: “maternal (when identified in the mother only)”, “paternal (when identified in the father only)”, “low coverage” (when there was insufficient read depth in either parent), “shared” (when identified in both parents)”, and “possibly de novo” (when absent in the parents). Variants classified as possibly de novo were included in this study. We applied various quality filters to ensure that only the most reliable calls were included in the study, based on stringent quality control. For this quality control we selected on GATK Quality score > 450, minimal number of variant reads of 10, minimal number of total reads of 20, minimal percentage of variant reads of 20%, frequency in dbSNP < 0.1%, coverage in parents of at least 10 reads, and we discarded 15 complex variants after manual inspection in the Integrative Genomics Viewer (https://igv.org/app/).
Rare and common variants in mTOR

mTOR

ASDs