A role for the male germline in the expansion of the mammalian brain

Stephen J. Bush[1*], Anne Goriely[2,3*]

[1] Interdisciplinary Research Centre for Information and Biomedical Sciences, School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China
[2] MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, OX3 9DS, UK
[3] NIHR Oxford Biomedical Research Centre, Oxford, OX3 7JX, UK

* Correspondence: stephen.bush@xjtu.edu.cn (S.J.B.), anne.goriely@imm.ox.ac.uk (A.G.)

ORCIDs:

Stephen J. Bush: 0000-0001-9341-2562
Anne Goriely: 0000-0001-9229-7216

Abstract

The brain and testis share a surprisingly high number of molecular and cellular similarities. We have previously hypothesised that, throughout evolution, many genetic variants contributing to brain size expansion first arose in spermatogonia where they conferred a selective advantage to the male germline stem cells via a process analogous to oncogenesis – known as 'selfish spermatogonial selection'. Once transmitted to the next generation, these selfish variants became constitutive, disproportionately accumulating in signalling pathways active in both spermatogenesis and neurogenesis and which regulate stem cell proliferation. Although the evidence supporting a close molecular relationship between the germline and brain is compelling, research in this area is stymied by the relative scarcity of spermatogonia and the inherent stochasticity of single-cell transcriptomic profiling. Accordingly, the molecular signatures of spermatogonia are incompletely understood, and their similarity with neural programs difficult to assess. To address this, we combine re-analyses of 34 adult human single-cell testis datasets with data from the Human Protein Atlas to assess the extent to which genes functionally associated with brain growth and development are expressed within testicular cell types. Consistent with our hypothesis, we find that among thousands of proteins with brain-associated functions, the majority are not only expressed in male germ cells, but show particular enrichment in spermatogonia. We contextualise these results with an extensive literature survey and conclude that further enquiry into the testis-brain connection may yield novel insight into the evolutionary processes that shaped the human condition.

Significance statement

The human brain and testis share unexpected molecular similarities, yet the evolutionary and biomedical implications of this overlap remain poorly understood. By integrating single-cell transcriptomic datasets with large-scale proteomic data, we show that genes implicated in neurodevelopment are widely expressed in the male germline and particularly enriched in spermatogonia. These findings support the idea that 'selfish' mutations arising in spermatogonial stem cells not only promote their own propagation in the testis but may also influence neural progenitor biology once inherited. Our results provide a systematic foundation for understanding how male germline-specific evolutionary forces could have contributed to the emergence of the large and complex human brain, while also offering insight into the origins of susceptibility to some congenital neurodevelopmental diseases.

Introduction

Prominent hallmarks of primate evolution include an increased brain volume and disproportionate expansion of its executive and integrative centre, the cortex (1,2). After correcting for allometric scaling with body mass, primates (and humans in particular) have larger brains than almost any other mammal (3), due to an increased number of neurons supporting more elaborate neural architectures and cognitive abilities (4). A recent review of human brain evolution offers a striking perspective on the extent to which neuron numbers have increased in our species: compared to other great apes, developing humans generate around three million more neurons per hour (5). How this came about is an active subject of debate and likely the result of a combination of different factors. In general, the mechanisms producing interspecies differences in neuron number and specialisation implicate the duration, timing, and relative rate of neural stem and progenitor cell (NSPC) divisions (6) during development and in what proportion and when mitoses are symmetrical (proliferative, or self-renewing) or asymmetrical (producing differentiation-inclined progeny) (7). Consequently, even subtle species-specific differences in NSPC activity can have pronounced effect upon the size and composition of both the brain as a whole, or its regions (8). Human-specific variants in, for instance, *ARHGAP11B* (9), *NOVA1* (10) and *TKTL1* (11) influence NSPC proliferation and by extension the overall size of the progenitor pool, and thereby brain. Hence, by modulating the rate of cell division, or by changing developmental decisions, genetic variants in a specific set of genes can alter the number of neurons produced over time and therefore play causal roles in neocortical expansion.

We have previously drawn attention to selfish spermatogonial selection – a mechanism by which *de novo* mutations affecting stem and progenitor cell regulation in the human testis gain a proliferative advantage and are preferentially introduced into the genome (12) – as we believe this may offer a fresh perspective on the molecular and developmental basis of brain expansion and its close connection to other phenotypes, in particular fertility. We have outlined this hypothesis briefly elsewhere (13) and illustrate its core principles in **Figure 1**. As our hypothesis approaches the topic of brain expansion from a different area of biology entirely, we refer readers to recent reviews of human brain evolution and development for subject-specific detail (see, for example, (5,14) and references therein). Here we focus instead on evaluating the evidence that bears directly on the role of selfish spermatogonial selection. Given this emphasis, a full treatment of other evolutionary forces shaping the brain-testis relationship is beyond the scope of this work, although selected aspects are considered in the **Supplementary Text**.

Selfish spermatogonial selection is a process taking place during adult spermatogenesis whereby a subset of spontaneously occurring oncogenic-like ('selfish') mutations confer a selective advantage to mutant spermatogonial stem cells (SSCs) which then clonally expand along the length of the seminiferous tubules as men age (see reviews (12,15,16)). As a consequence, selfish mutations disproportionately accrue in the sperm of older men and, compared to classical spontaneous mutations, have a higher likelihood of being inherited as a constitutive mutation in the next generation. To date, the best characterised examples of selfish mutations are pathogenic gain-of-function single nucleotide variants (SNVs), typically (but not exclusively (17)) occurring within components of the RTK/RAS/MAPK signalling pathway, a central regulator of testicular homeostasis (18,19) which maintains the most primitive SSC state (20). Documented examples include variants in *BRAF*, *CBL*, *FGFR2*, *FGFR3*, *HRAS*, *KRAS*, *MAP2K1*, *MAP2K2*, *NRAS*, *PTPN11*, *RAF1*, *RET*, *SMAD4* and *SOS1* (see (12,17) and references therein, plus (21) for additional candidates). These selfish mutations are consistently associated with severe conditions in the offspring including skeletal dysplasia and neurodevelopmental disorders, such as RASopathies (15). While their pathogenic phenotypes are complex and heterogenous (individuals carrying identical RTK/RAS/MAPK mutations can exhibit different clinical presentations (22)), cranial

overgrowth or macrocephaly, a disproportionately large head, are recurrent features (23). This is plausibly linked to the roles of many of these genes in regulating cell fate decisions during neurogenesis (discussed further in the **Supplementary Text**).

These observations provide the empirical grounds for our hypothesis – namely, that the clonal dynamics of SSCs have broader implications beyond fertility, and that changes in mammalian brain volume may, at least in part, have been shaped by the selfish properties of the testis.

Despite vast differences in structure and function, the testis and brain share many cellular and molecular similarities, including substantial overlap of both their proteomes and transcriptomes (24), larger repertoires of splice variants than any other organ (25), and – critically for our argument – coordinated cell fate programs, with the neural lineage and the germline co-regulated during early development (26). Indeed, not only is there a growing catalogue of genes with pleiotropic functions in both the germline and nervous system (27), but many characteristic spermatogonial markers also play functional roles in neurogenesis and brain expansion (13) (see also **Supplementary Text**). Moreover, both lineages exhibit similar age-dependent mutational signatures, in particular SBS5. Under a recent model, this signature arises when multiple sources of DNA damage converge into a common mutational outcome, producing similar signatures across diverse cell types. This has been interpreted to suggest the output of common, tightly coupled, endogenous processes; that is, in drawing on the same molecular machinery, the germline and neural lineages accrue the majority of their mutations the same way (28).

As an evolutionary force, selfish selection favours the introduction into our genome of functional variants that alter specific molecular pathways modulating SSC self-renewal (13). As such, when mutant SSCs differentiate, they generate disproportionate numbers of sperm carrying the selfish allele. Provided that the resulting heterozygous variant is not overly deleterious to the somatic health of the offspring, it can become heritable and persist across generations. Importantly, tight control of the self-renewal of stem cells at the onset of neurogenesis is essential for generating a large pool of progenitors (see review (29)), the size of which predicts both mature adult brain volume and the relative size of brain regions (30). Consistent with this, mathematical modelling suggests that an additional round of symmetric – rather than asymmetric (neurogenic) divisions – by neuroepithelial stem cells would double the number of neurons in the cortex (31). Further supporting a mechanistic link between male reproduction and neurogenesis, studies of human brain organoids show that androgens, key regulators of spermatogenesis, can increase the number of proliferative divisions in cortical progenitors (32).

Female reproductive tissues (which, like the testis, possess unusually broad proteomic overlap with the brain (33)) may also be relevant to brain expansion; they are enriched in genes associated with endocranial globularity, a uniquely human phenotype reflecting the timing and pattern of brain growth (34). Similarly, a recent study which performed *in silico* modelling of hominin brain expansion argued that increased brain size may have arisen not from direct selection on cognition but by developmental constraints on fertility (specifically, affecting the number of ovarian follicles) (35). Although the study lacked data for males, these constraints may have diverted evolutionary change in a direction that led to brain expansion, such that this phenotype evolved as a spandrel (36).

While these observations point to intriguing connections between the germline and brain development, our aim here is to evaluate this relationship more systematically. The outstanding challenge in characterising molecular similarities between the brain and either spermatogonia or SSCs (a subset of undifferentiated spermatogonia (37)) is relative paucity of data: not only are spermatogonia a rare cellular population (38) but compared to other human tissues, relatively few single-cell transcriptomic datasets are available for testis (39). Accordingly, the molecular identity of spermatogonia, and the heterogeneity of their subpopulations, remains incompletely understood (40). To address this limitation, we had previously created an integrated single-cell atlas of 34

publicly-available adult human testis samples, with the aim of increasing the resolution at which spermatogonial states and their developmental trajectory could be defined (41).

Here we combine transcriptomic data from multiple single-cell testis atlases (our own (41), that of a previous study (42) and the HumanTestisDB (43)) and protein-level expression data from the Human Protein Atlas (44–46) to assess the extent to which genes functionally associated with brain expansion and development are also expressed in the testis. We find that of thousands of proteins with established roles in the brain, the majority are expressed in male germ cells, with particular enrichment in spermatogonia. Furthermore, we also find the converse to be true: genes enriched in spermatogonia, or implicated in human-specific aspects of spermatogenesis, are widely expressed throughout the brain. We contextualise these results with an extensive discussion of the literature and argue that, taken together, they support further enquiry into the testis-brain relationship as a potential source of insight into the evolutionary processes that shaped the human condition. To support this work, we provide extensive datasets (**Supplementary Tables 1 to 7**) as a resource to the community.

Materials and Methods

Sources of raw data

Our analyses draw on several publicly accessible proteomic and transcriptomic datasets. Firstly, from the Human Protein Atlas (HPA) v23 (44–46), we obtained the files 'normal_tissue.tsv' and 'rna_tissue_hpa.tsv' which represent, respectively, protein-level expression for 56 tissues and 122 distinct cell types (of which 8 cell types were characterised for the testis) and transcript-level expression for 40 tissue types (https://v23.proteinatlas.org/download/normal_tissue.tsv.zip and https://v23.proteinatlas.org/download/rna_tissue_hpa.tsv.zip, downloaded 17th June 2023). Secondly, we obtained an integrated single-cell atlas of 60,427 adult human testicular cells (representing data from 29 individuals across 9 different studies (47–55) and an age range of 14-66 years), the bioinformatic methods for which we described in previous work (41) and briefly recapitulate here. To create this atlas, we generated cell x gene count matrices using the kallisto/bustools v0.26.3 'count' workflow (56), which were then processed using Seurat v4.0.3 (57) with SCTransform normalisation (58) and integrated using Seurat's rPCA 'anchor' technique. Using an unsupervised clustering approach (the Smart Local Moving algorithm (59)), we then identified 10 distinct cell clusters, which were annotated using both a global (all-against-all) differential gene expression analysis and established biomarkers for the major somatic and germ cell types (sourced from (47,49,54,60–63) and as previously described (41)). These comprised 3 somatic cell clusters ('myoid and Leydig cells', 'Sertoli cells', and 'endothelia and macrophages', collectively 10,788 cells) and 7 germ cell clusters ('undifferentiated spermatogonia' (undiff SPG), 'differentiating spermatogonia/early meiosis' (diff SPG), 'spermatocyte', 'early spermatid 1', 'early spermatid 2', 'late spermatid 1' and 'late spermatid 2', collectively 49,639 germ cells). As a validation that these clusters followed the expected developmental trajectory, the relative proportions of spermatogonia, spermatocytes, and spermatids approximated the ratio 1:2:4. The undiff SPG cluster contained 3298 cells, the spermatocyte cluster 6714 cells, and the four spermatid clusters collectively 36,897 cells (a ratio of 1 to 2.03 to 5.49); note that the germline clusters form a continuum, as shown in **Figure 1B**, and so do not necessarily represent discreet cell types. Differentially expressed genes in each cluster were identified using Seurat's FindAllMarkers function with parameters min.pct = 0.25, logfc.threshold = 0.25 and only.pos = TRUE which, respectively, require that a gene is expressed in >25% of the cells in a given cluster, that it has a $\log_2$ fold-change difference in expression relative to all other clusters of > 0.25 (that is, > 2-fold average expression) and that it is on average more highly expressed in that cluster compared to all others.

We also obtained comparable sets of Seurat differential expression analysis results from four previously published human single-cell whole-testis and male-germline atlases, one from a previous study (42) (which integrated data from (47,49,53,54)) and three from the HumanTestisDB (43) (which integrated data from (47,49,52–54,64,65)). Data sources, and details of both the expression atlas contents and analyses performed, are given in **Supplementary Table 4**.

Protein-coding genes functionally associated with the human brain

To assess the degree of overlap between genes expressed in the testis and functionally associated with the brain, we first compiled an inclusive list of 'brain genes' on the basis of either an explicit or suggestive phenotype (**Supplementary Table 1**). This comprised 7193 human brain-associated protein-coding genes (36% of the 19,871 in total in the hg38 primary assembly), collated from previous studies of genes associated with macrocephaly (66), megalencephaly (i.e. abnormal brain size) (67), postmortem brain weight (68,69), cognitive function (using intelligence quotient (70) and educational attainment (71) as provisional proxies), and human-accelerated regions (72) (otherwise conserved genomic loci with human-specific substitutions, the expression of which spatially correlates with patterns of cortical expansion (73)). We also collated genes from previous reviews of the molecular factors governing brain development (14), human cognition (74), neocortical evolution (75,76), autism spectrum disorder (77–81), schizophrenia (82–84), genetic epilepsies (85–88), and indirect neurogenesis, a process thought to underlie the evolution of the gyrencephalic cortex (89). The overall list is presented in **Supplementary Table 1**, alongside notes on the sources, summary statistics, and the classification of each gene by the nature of its association. As a resource to support further enquiry, genes are assigned, where possible, to one or more of the seven categories of macrocephaly/megalencephaly (n=379 genes), brain weight (n=879), autism (n=2562), schizophrenia (n=345), epilepsy (n=3372), intelligence/educational attainment (n=2303), or human-acceleration (n=1608). We consider the first five of these seven categories to be explicit rather than suggestive phenotypes and to collectively constitute a conservative set of 5212 brain-associated genes.

Spermatogenesis and male infertility-associated genes

We compiled a combined, non-redundant, set of 218 genes considered to be either a human spermatogonial marker (90–92), to have a human-specific role in spermatogenesis (93), or to be a core (conserved) component of the metazoan spermatogenic program (94), alongside a set of 154 genes moderately, strongly or definitively linked to a male infertility phenotype, sourced from two systematic reviews of monogenic gene-disease relationships (95,96). Genes, and their sources, are detailed more fully in **Supplementary Table 5**.

Results

Genes associated with brain growth and development are widely expressed in the male germline

Multiple genes have been functionally implicated with increased brain mass or identified as positively selected along the primate phylogeny or the lineage leading to humans (see reviews (9,14,75,76)). Given the extensive overlap between testicular and neural transcriptomes (24), we hypothesised that a substantial fraction of these brain-growth-associated genes would be active in the male germline. To assess this, we draw upon both the Human Protein Atlas v23 (44–46) and an integrated single-cell transcriptomic atlas of 60,427 human testicular cells (**Figure 2**), which we have previously described (41).

We first examined the expression levels of a list of 7193 brain-associated protein-coding genes using the single-cell atlas (**Supplementary Table 1**). This list is intentionally broad with no filters applied beyond those of the original authors (see notes to **Supplementary Table 1**); rather, our aim was simply to determine whether even an unselected list of 'brain genes' would show significant

overlap with expression in the testis. We summarise the testicular expression of each gene at the protein-level in **Supplementary Table 1** and at the transcript-level in **Supplementary Tables 2 and 3**.

Of the 7193 genes in **Supplementary Table 1**, 2392 (33%) lacked usable protein expression data (either no data or data marked by the HPA as 'uncertain') and were excluded from summary statistics. Of the remaining 4801 genes, experimental antibody staining was available; for 2467 (51%) genes staining had been performed for three somatic (peritubular, Leydig, and Sertoli) and five germ cell types, including spermatogonia (detailed further in the notes to **Supplementary Table 1**). For the other 2334 genes (49%), staining was only performed for two cell types: Leydig cells and the broader category of 'cells in seminiferous ducts' (that is, seminiferous tubules, comprising all germ cells and Sertoli cells). The HPA only contains the more specific level of antibody staining data (eight cell types, rather than two) for genes with testis-elevated mRNA expression (detailed at https://www.proteinatlas.org/humanproteome/tissue/testis, accessed 12th May 2024). We considered protein expression in seminiferous ducts to represent highly probable germline expression as Sertoli cells, the only somatic cells within them, represent only a few percent of their total (97). Although we cannot exclude the possibility that some signals in seminiferous ducts reflect Sertoli cell-specific proteins, this seems unlikely to affect our conclusions as of the 2467 genes stained for eight cell types, only 76 (3.1%) were exclusively detected in Sertoli cells (**Supplementary Table 1**).

Overall, we found that 73% of brain-associated genes had either experimentally-supported or highly probable germline expression (that is, 3512 of 4801 genes for which data was available, irrespective of whether antibody staining was performed only for seminiferous ducts or for specific germ cells; **Supplementary Table 1**). Furthermore, of the 2467 genes where antibody staining data was available for specific germ cells, 2049 brain-associated genes (83%) had detectable protein-level expression in at least one germ cell type, more of which (n = 1476, 59%) were expressed before meiosis (in spermatogonia) than either during (in preleptotene or pachytene spermatocytes; n = 1258, 51%) or after (in round or elongated spermatids; n = 1328, 54%) (**Supplementary Table 1**). The number and distribution of brain-associated genes expressed in the testis is summarised in **Figure 2**.

To further assess whether genes associated with brain growth and development have enriched expression in spermatogonia, we next considered transcript-level expression (**Supplementary Tables 2 and 3**). The testicular transcriptome and proteome are surprisingly dissimilar (98): while the testis is abundant in mRNAs, many proteins with high testicular mRNA expression remain undetected by antibody-based or mass spectrometry approaches, with no obvious technical explanation available. One compelling suggestion is that these 'missing' proteins are enriched for spermatogenesis-related processes and may have transient, specialised, functions, being rapidly degraded (98).

Using data from the testicular single-cell RNA-seq atlas (41), we found that 6293 of the 7193 brain-associated genes (87%) were detectably expressed in the germline (having one or more sequenced reads in >1% of the germ cells; **Supplementary Table 3**). The number of genes differentially expressed in a germ cell cluster declined across the spermatogenic trajectory, being at its highest for 'undiff SPG' (1420 genes, or 20% of the total, including *TMEM14B*, a primate-specific gene which promotes cortical expansion (99), and both *NOVA1* (10) and *TKTL1* (11), two genes in which human-specific substitutions have been implicated in neocortical expansion), then decreasing to 1312 ('diff SPG'), 831 ('spermatocyte'), 763 ('early spermatid 1'), 731 ('early spermatid 2'), 245 ('late spermatid 1') and 286 ('late spermatid 2') (**Figure 2** and **Supplementary Table 2**).

Overall, these findings indicate that many genes functionally associated with brain growth are expressed not only in the testis but show particularly enriched expression in undifferentiated spermatogonia, more so than either meiotic or post-meiotic adult germ cells (**Figure 2C**). We note, however, that some of the gene categories in **Figure 2A** represent suggestive (such as IQ (70) or

educational attainment (71), broad proxies for cognitive function) or indirect (such as 'human-accelerated sequence' (72)) associations with the brain, rather than discreet phenotypes. Nevertheless, repeating the analysis after excluding genes found only in these categories, we obtain quantitively similar results (**Supplementary Figure 1**). Moreover, brain-associated genes were significantly more highly expressed in undifferentiated spermatogonia than all other germ cell types, except spermatocytes (Mann-Whitney U $p < 0.05$; **Supplementary Figure 2**).

In terms of single cell expression profiles, the male germline resembles a continuum of 'states' (**Figure 2B**) with its division into discrete cell types imposed by a clustering algorithm. As such, our conclusions about the relative enrichment of brain-associated genes across germline populations (i.e. their differential expression; **Figure 2C**) could be sensitive to methodological choices in clustering and annotation. To ensure robustness, we repeated the analysis using comparable Seurat clustering results from four previously published human single-cell whole-testis and male-germline atlases (42,43), each of which annotated their cell types independently and varied in the cells used, age range, granularity, and methodology (described further in **Supplementary Table 4**). Across all datasets, we found that brain-associated genes were more frequently differentially expressed in germ than somatic cells (**Supplementary Figures 3 and 4**). Consistent with our own expression atlas (41), differentially expressed genes were also particularly abundant towards the start of the spermatogenic trajectory (**Supplementary Figures 3 and 4**). However, unlike our own atlas, which only included adult samples, these alternative datasets also comprised foetal, pre-, and peri-pubertal cells (**Supplementary Table 4**), and hence the highest proportion of differentially expressed genes often appeared in germ cells 'adjacent' to spermatogonia, namely leptotene spermatocytes and primordial germ cells (PGCs; **Supplementary Figures 3 and 4**). Notably, PGCs give rise to fetal spermatogonia (100) and are transcriptomically similar to 'state 0' adult spermatogonia, the most undifferentiated spermatogonial state (65).

To test whether our results were driven by specific brain phenotypes, we repeated our analysis for each of the seven phenotypic categories separately. We consistently found differentially expressed genes enriched in the spermatogonia of our own atlas (**Supplementary Figure 5**) and, with few exceptions, around the start of the spermatogenic trajectory in the others (**Supplementary Figures 6, 7, 8 and 9**).

Finally, to assess whether the observed testis-brain overlap simply reflected a broader enrichment for developmental or housekeeping genes (given their disproportionate association with cell proliferation), we examined, for each phenotypic category, the number of proteins detected in each of the 56 tissues comprising the HPA. As expected, the greatest proportion of each gene (protein) set was expressed in cerebral cortex, cerebellum, and testis (**Figure 3**). This suggests that the overlap is not a generic consequence of ubiquitous protein expression but instead reflects shared molecular programs among these tissues.

Further supporting a bidirectional link between genes functional in the brain and male germline, we also found evidence for the converse: genes considered human spermatogonial markers (90–92), having human-specific roles in spermatogenesis (93), associated with monogenic male infertility (95,96), or thought to be core components of the metazoan spermatogenic program (94), are all widely expressed throughout the brain (discussed further in the **Supplementary Text**; see also **Supplementary Tables 5 and 6**).

The breadth of the HPA antibody staining data offers an additional line of evidence for a functional connection between the male germline and brain: genes whose protein expression is restricted to the two tissues. Although we found only 16 genes out of 5379 (0.3%) whose proteins were only detectable in both a male germ and brain cell, and no other cell type in any other (healthy) tissue (**Supplementary Table 6**), these represent potential entry points for further exploring the molecular parallels between the two organs (discussed further in the **Supplementary Text**). Notably, several of these have been associated with brain-expansion phenotypes. Three have been associated with increased postmortem brain weight (68) – *ERMN* (which promotes oligodendroglial differentiation

(101)), *HTR2A* (ectopic expression of which increases the number of basal progenitor cells in embryonic mice (102)), and *OPALIN* (which promotes oligodendrocyte terminal differentiation (103)) – one with macrocephaly (66) (*HYLS1*, a conserved centriole protein (104)), and another (*OMG*, an oligodendrocyte glycoprotein which inhibits cell proliferation (105)) has a co-evolutionary relationship with primate brain mass (69) (**Supplementary Table 1**). We note, however, that some of these proteins may also be expressed at low levels or in tissues not assayed by the HPA, and cannot be assumed to be strictly brain-testis specific.

Having shown that many genes expressed in spermatogonia are functionally associated with the brain, including the brain-expansion phenotypes of macrocephaly and megalencephaly, we next considered whether the germline was also enriched for genes harbouring human-specific variants, as these are disproportionately associated with neurodevelopment (73).

Proteins with human-specific amino acid changes are simultaneously enriched for neurodevelopmental functions and widely expressed in male germ cells

Advances in genome sequencing have enabled detailed comparisons of modern humans with archaic hominins, and identified genetic variants that define human-specific traits (see review (106)). Although the precise catalogue of these variants depends on archaic sample availability, reference genome annotation and coverage, as well as genotype calling and filtering methods, there is nevertheless a broad consensus on the strongest candidates – protein-coding changes in cell cycle-related genes to which the brain is especially sensitive (107). A study comparing modern human, Denisovan, and Neanderthal genomes identified 571 genes with non-synonymous amino acid changes in humans (all present at high variant allele frequencies), and an additional 36 genes in which 42 amino acid changes were fixed in modern humans (107). Enrichment analyses suggested that these genes mostly affect cell division and early brain growth trajectories, with notable expression in the infant frontal cortex (107). However, this same gene set also shows a strong connection to the testis: it is significantly enriched for the GO term 'spermatoproteasome complex'. Furthermore, 3 of the 42 fixed amino acid changes occur in *SPAG5* (sperm-associated antigen 5), a gene crucial for meiosis and spermatid morphogenesis (108) (yet dispensable for fertility (109)) and which is also expressed in the fetal ventricular zone during cortical neurogenesis (110). *SPAG5* encodes a mitotic spindle protein, with mutant phenotypes consistent with a critical role in whether a cell divides symmetrically or asymmetrically (106). This is particularly relevant to NSPCs as the orientation of the mitotic cleavage plane can impact cell fate decisions, and thereby the composition of the progenitor pool (111).

Motivated by these observations, we hypothesised that genes containing human-specific amino acid changes would also be detectably expressed (and putatively functional) in the male germline, and spermatogonia in particular. To test this, we assessed protein-level expression in the male germline for 309 of the 607 genes in the human-specific substitution set (genes with no usable data were excluded; **Supplementary Table 7**). Of these, 70% had either experimentally-supported or highly probable (that is, in seminiferous ducts) germline expression. Furthermore, of the 152 genes where protein expression data was available for five germ cell types, 22 were only detectable in spermatogonia (compared to 4 in spermatocytes and 15 in spermatids). These include *ASCC1*, mutations in which downregulate genes associated with neurogenesis and neuronal migration (112), *LRRTM1*, an imprinted gene expressed throughout forebrain development and associated paternally with cerebral asymmetry (113), and *TCF3*, which represses neuronal differentiation and increases the self-renewal of NSPCs during neocortical development (114).

Discussion

We have previously hypothesised that the dysregulation of male germline homeostasis provides a novel perspective on the mechanisms underlying mammalian brain expansion (13). In the present work we develop this hypothesis by systematically examining the extent to which genes with functional roles in brain growth, development or evolution are also expressed in the male germline. In conjunction with numerous observations from the literature, we summarise our results, alongside other evidence supporting our hypothesis, in **Table 1**.

Our entryway into this topic was the detection of spontaneous oncogenic-like mutations in spermatogonial stem cells, in genes that are functionally critical to brain development. Selfish selection, the underlying mechanism, provides a causal link between the otherwise disparate fields of spermatogenesis, neurogenesis, and oncogenesis, and offers a novel perspective on a number of observations. For example, a genome-wide association study of the genetic variants associated with human head size (115) (a highly heritable trait and strong correlate of brain size) identified 67 candidate loci. These loci were disproportionately located near genes preferentially expressed in intermediate neural progenitors, whose increased proliferation has been linked to human brain expansion (116), as well as being enriched for pathways involved in both macrocephaly and cancer. Consistent with our hypothesis, many proteins encoded by these genes are abundant in spermatogonia, including ALAS1, BNIP3L, FOXO3, HMGA2, IGF2BP1, NCK1, NFIX, PPP6R3, and STAG1 (**Supplementary Table 1**).

Similarly, another genome-wide association study of human brain MRI data (117) identified 199 loci significantly associated with either cortical surface area or thickness, traits linked to the NSPC proliferation rate and their number of neurogenic divisions, respectively. Many of these loci lie in genes involved with the establishment and maintenance of the male germline, or in regions functionally linked to Wnt signalling, a key regulator of spermatogenic cell fate (118). Examples include variants in *PAX7* (a marker of spermatogonial stem cells in mice (119)), *MOV10* (knockdown of which affects spermatogonial cell fate decisions in mice (120)), *DAAM1* (which regulates the actin cytoskeleton of sperm (121)), and *HDAC9* (which, in chickens, promotes the differentiation of embryonic stem cells into male germ cells (122).

Our hypothesis of a testis-brain relationship also intersects with the biology of human-accelerated regions (HARs), many of which act as neurodevelopmental enhancers (123). The evolutionary changes underpinning human cognitive abilities have been conceptualised as “Achilles’ heels” as they may predispose to the high burden of psychiatric disorders in our species – that is, they may represent maladaptive by-products of adaptations in brain development (124). Supporting this view, many HARs with enhancer activity in NSPCs show evidence of compensatory evolution to maintain ancestral activity. That is, variants may have initially induced large shifts in regulatory activity in the brain which were then moderated by nearby substitutions (124). Although the underlying forces responsible for accelerated substitution rates in HARs remain poorly understood, this ‘back and forth’ pattern is compatible with a scenario in which variants were initially adaptive in another context (i.e. male germline) and only indirectly affected the brain, with phenotypic consequences that may be mildly deleterious but tolerated.

As discussed above, these testis-brain associations reflect the shared developmental and regulatory origins of the neural lineage and the germline. In humans, this coordination centres on the early specification of primordial germ cells (PGCs), mediated by the core transcriptional network of *OCT4* (*POU5F1*), *PAX5*, and *PRDM1* (*BLIMP1*). This regulatory ‘master switch’ simultaneously activates germline and represses somatic differentiation programs (the default for which in mice is neuronal (125)), thereby committing cells to a germline fate (reviewed in (90,126)). The establishment and maintenance of this fate requires extensive chromatin modelling (127); disruption of this process can lead to reprogramming, with PGCs aberrantly expressing somatic genes. Tellingly, when germ cells reprogram as soma in *C. elegans*, they often express pan-neuronal markers, extrude neurite-like projections, and differentiate into neurons (128).

Although our focus here has been on the conceptual links between SSCs and NSPCs, the genetic architecture of brain size is multifaceted and complex (129). It includes, among others, gene duplication (130,131), gene family size variation (132–134), and changes to coding (135,136), non-coding (137), and regulatory (138) sequences, these mechanisms collectively giving rise to relatively large brains through multiple independent evolutionary trajectories (139). Moreover, larger brains are not only defined by a greater number of neurons: they also differ in neuronal size, regional specialisation, and packing density, as well as their number and proportion of somatic support cells (glia) (140). Importantly, not all genes implicated in brain size expansion are functionally active in the testis as the overlap between testis and brain proteomes, although substantial, is not exhaustive. For example, human-specific *NOTCH2NL* genes, which arose from recent gene duplication events to play critical roles in cortical neurogenesis (141), have no known association with spermatogenesis.

Taken together, we have drawn upon the biology of the testis to offer a fresh perspective on an old question: how did the human brain become so large? We believe new insight into this problem can arise from selfish spermatogonial selection, a male germline-specific evolutionary force that enriches genomes for oncogenic-like mutations in pathways involved in the regulation of spermatogonial stem cells – and which are also functional in neural stem and progenitor cells. In this respect, we can conceive of the male germline as both a generator of mutations and a filter, disproportionately passing to the next generation those of a particular ('selfish') nature – ones benefiting reproductive success and robust sperm production whilst ageing. It follows that if selfish mutations disproportionately accumulate within pathways essential for germline proliferation and fertility, they may also influence the phenotype of other (somatic) cells in which the same genes are expressed. Consistent with this, protein-coding changes in genes involved in cell cycling rank among the strongest candidates for human-specific sequence variation (107).

Scaled up to the population level, the long-term effect of this additive, non-random, germline mutagenesis may have been to expose to selection variants that modulate baseline rates of cell division, potentially exerting second-order effects in related lineages, such as neural progenitors. This relates to current theories of human brain evolution through the concept of bradychrony ('slowed time'), which refers to the relative slowing of equivalent developmental processes between species (142). Notably, differences in developmental tempo correlate with differences in brain size. In humans, brain maturation is markedly protracted (bradychronic) compared to other primates, taking a longer proportion of the lifespan to reach adult brain weight (142). Early human neurodevelopment is particularly prolonged, with the period of tangential cortical expansion, driven by symmetrically-dividing radial glial cells, lasting approximately a month in humans but a fortnight in macaques (142). The mechanisms driving differences in neurodevelopmental tempo are an area of active enquiry with calcium regulation (a process fundamental to sperm function (24)) emerging as a key candidate (143). We suggest that variants selfishly selected in the testis could also contribute by subtly modulating stem cell turnover rates, thereby influencing the duration and extent of neurogenic output.

To conclude, we highlight a pronounced, yet largely unexplored, role for the male germline – and spermatogonia in particular – in shaping both the development and evolution of the brain. Beyond demonstrating a substantial overlap between genes expressed in spermatogonia and those functionally associated with brain development, we draw conceptual parallels between the neural and spermatogonial stem cell lineages noting that the timing and balance of self-renewing and differentiating cell divisions are not only critical to both but especially so for longer-lived species (such as humans), who reproduce over many decades. We propose that variants selfishly selected in the testis (12,15,16) may bias this balance towards increasing cellular output (i.e. more sperm and/or neurons), thereby influencing both fertility and cognitive potential. We hope our analyses will encourage further enquiry into this topic as we find it particularly compelling to consider how

mutations selectively advantageous to human spermatogonia may have been co-opted by other tissues, contributing to the genomic innovations underlying our unusually large brains.

Authors' contributions

Stephen J. Bush: conceptualisation, methodology, formal analysis, data curation, investigation, validation, visualisation, software, writing – original draft, writing – review & editing.
Anne Goriely: conceptualisation, validation, funding acquisition, writing – review & editing.

Competing interests statement

The authors declare that there are no conflicts of interest.

Funding

This work was supported by grants from the Wellcome Trust (219476/Z/19/Z) and the European Society of Human Reproduction and Embryology (ESHRE G20-0016) to Anne Goriely, and by the Sanqin Talent Introduction Plan of Shaanxi Province (2023SYJ01) and the Xi'an Jiaotong University Young Talent Program class A to Stephen J. Bush. The funders had no role in study design, data collection and analysis, preparation of the manuscript or decision to publish.

Data availability

The data supporting this work is available either in the article itself or in its electronic supplementary material.

Code availability

All scripts used to perform the analyses and create both figures and tables are available at www.github.com/sjbush/testis_brain.

References

1. Nomura T, Gotoh H, Ono K. Changes in the regulation of cortical neurogenesis contribute to encephalization during amniote brain evolution. Nat Commun. 2013;4:2206.

2. Kaas JH. The evolution of brains from early mammals to humans. Wiley Interdiscip Rev Cogn Sci. 2013 Jan;4(1):33–45.

3. Barton RA. Primate brain evolution: Integrating comparative, neurophysiological, and ethological data. Evolutionary Anthropology: Issues, News, and Reviews. 2006 Nov 1;15(6):224–36.

4. Geschwind DH, Rakic P. Cortical Evolution: Judge the Brain by Its Cover. Neuron. 2013 Oct 30;80(3):633–47.

5. Lancaster MA. Unraveling mechanisms of human brain evolution. Cell. 2024 Oct 17;187(21):5838–57.

6. Otani T, Marchetto MC, Gage FH, Simons BD, Livesey FJ. 2D and 3D Stem Cell Models of Primate Cortical Development Identify Species-Specific Differences in Progenitor Behavior Contributing to Brain Size. Cell Stem Cell. 2016 Apr 7;18(4):467–80.

7. Homem CCF, Repic M, Knoblich JA. Proliferation control in neural stem and progenitor cells. Nature Reviews Neuroscience. 2015 Nov 1;16(11):647–59.

8. Dyer MA. Stem Cells Expand Insights into Human Brain Evolution. Cell Stem Cell. 2016 Apr 7;18(4):425–6.

9. Huttner WB, Heide M, Mora-Bermúdez F, Namba T. Neocortical neurogenesis in development and evolution-Human-specific features. J Comp Neurol. 2024 Feb;532(2):e25576.

10. Trujillo CA, Rice ES, Schaefer NK, Chaim IA, Wheeler EC, Madrigal AA, et al. Reintroduction of the archaic variant of NOVA1 in cortical organoids alters neurodevelopment. Science. 2021 Feb 12;371(6530):eaax2537.

11. Pinson A, Xing L, Namba T, Kalebic N, Peters J, Oegema CE, et al. Human TKTL1 implies greater neurogenesis in frontal neocortex of modern humans than Neanderthals. Science. 377(6611):eabl6422.

12. Wood KA, Goriely A. The impact of paternal age on new mutations and disease in the next generation. Fertility and Sterility. 2022 Dec 1;118(6):1001–12.

13. Bush SJ, Goriely A. Can the male germline offer insight into mammalian brain size expansion? Andrology [Internet]. 2024 Sep 18 [cited 2024 Sep 18];n/a(n/a). Available from: https://doi.org/10.1111/andr.13766

14. Zhou Y, Song H, Ming G li. Genetics of human brain development. Nature Reviews Genetics. 2024 Jan 1;25(1):26–45.

15. Goriely A, Wilkie AOM. Paternal age effect mutations and selfish spermatogonial selection: causes and consequences for human disease. Am J Hum Genet. 2012 Feb 10;90(2):175–200.

16. Goriely A, McGrath JJ, Hultman CM, Wilkie AOM, Malaspina D. 'Selfish spermatogonial selection': a novel mechanism for the association between advanced paternal age and neurodevelopmental disorders. Am J Psychiatry. 2013 Jun;170(6):599–608.

17. Wood KA, Tong RS, Motta M, Cordeddu V, Scimone ER, Bush SJ, et al. SMAD4 mutations causing Myhre syndrome are under positive selection in the male germline. Am J Hum Genet. 2024 Sep 5;111(9):1953–69.

18. Li MWM, Mruk DD, Cheng CY. Mitogen-activated protein kinases in male reproductive function. Trends Mol Med. 2009 Apr;15(4):159–68.

19. Wei BH, Hao SL, Yang WX. Regulation of spermatogonial stem cell self-renewal and proliferation in mammals. Histol Histopathol. 2022 Sep;37(9):825–38.

20. Capponi C, Smith C, Medica A, Hsieh TC, Orwig KE, Tan K. Decoding the heterogeneity of human undifferentiated spermatogonia reveals RAS-dependent regulation of stem cell fate. Cell Reports. 2026 Feb 24;45(2):116868.

21. Seplyarskiy V, Moldovan MA, Koch E, Kar P, Neville MDC, Rahbari R, et al. Hotspots of human mutation point to clonal expansions in spermatogonia. Nature. 2025 Nov 1;647(8089):429–35.

22. Krab LC, Goorden SMI, Elgersma Y. Oncogenes on my mind: ERK and MTOR signaling in cognitive diseases. Trends Genet. 2008 Oct;24(10):498–510.

23. Williams CA, Dagli A, Battaglia A. Genetic disorders associated with macrocephaly. American Journal of Medical Genetics Part A. 2008 Aug 1;146A(15):2023–37.

24. Matos B, Publicover SJ, Castro LFC, Esteves PJ, Fardilha M. Brain and testis: more alike than previously thought? Open Biology. 2021 Jun 2;11(6):200322.

25. Naro C, Cesari E, Sette C. Splicing regulation in brain and testis: common themes for highly specialized organs. Cell Cycle. 2021 Mar;20(5–6):480–9.

26. Schulz M, Teissandier A, De La Mata Santaella E, Armand M, Iranzo J, El Marjou F, et al. DNA methylation restricts coordinated germline and neural fates in embryonic stem cell differentiation. Nature Structural & Molecular Biology. 2024 Jan 1;31(1):102–14.

27. Kulkarni A, Lopez DH, Extavour CG. Shared Cell Biological Functions May Underlie Pleiotropy of Molecular Interactions in the Germ Lines and Nervous Systems of Animals. Frontiers in Ecology and Evolution. 2020;8:215.

28. Spisak N, de Manuel M, Przeworski M. Collateral mutagenesis funnels multiple sources of DNA damage into a ubiquitous mutational signature. United States; 2025.

29. Casas Gimeno G, Paridaen JTML. The Symmetry of Neural Stem Cell and Progenitor Divisions in the Vertebrate Brain. Front Cell Dev Biol. 2022;10:885269.

30. Jean F, Stuart A, Tarailo-Graovac M. Dissecting the Genetic and Etiological Causes of Primary Microcephaly. Front Neurol. 2020;11:570830.

31. Montgomery SH, Capellini I, Venditti C, Barton RA, Mundy NI. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. Mol Biol Evol. 2011 Jan;28(1):625–38.

32. Kelava I, Chiaradia I, Pellegrini L, Kalinka AT, Lancaster MA. Androgens increase excitatory neurogenic potential in human brain organoids. Nature. 2022 Feb;602(7895):112–6.

33. Kosti I, Jain N, Aran D, Butte AJ, Sirota M. Cross-tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues. Sci Rep. 2016 May 4;6:24799.

34. Molz B, Eising E, Alagöz G, Schijven D, Francks C, Gunz P, et al. Imaging genomics reveals genetic architecture of the globular human braincase. bioRxiv. 2024 Jan 1;2024.03.20.585712.

35. González-Forero M. Evolutionary–developmental (evo-devo) dynamics of hominin brain size. Nature Human Behaviour. 2024 Jul 1;8(7):1321–33.

36. González-Forero M, Gómez-Robles A. Why did the human brain size evolve? A way forward. Philos Trans R Soc Lond B Biol Sci. 2025 Jun 26;380(1929):20240114.

37. Lord T, Oatley JM. Spermatogenic Stem Cells: Core Biology, Defining Features, and Utilities. Mol Reprod Dev. 2024 Oct;91(10):e23777.

38. Fayomi AP, Orwig KE. Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. Stem Cell Res. 2018 May;29:207–14.

39. Tirumalasetty MB, Bhattacharya I, Mohiuddin MS, Baki VB, Choubey M. Understanding testicular single cell transcriptional atlas: from developmental complications to male infertility. Front Endocrinol (Lausanne). 2024;15:1394812.

40. Di Persio S, Neuhaus N. Human spermatogonial stem cells and their niche in male (in)fertility: novel concepts from single-cell RNA-sequencing. Hum Reprod. 2023 Jan 5;38(1):1–13.

41. Bush SJ, Nikola R, Han S, Suzuki S, Yoshida S, Simons BD, et al. Adult Human, but Not Rodent, Spermatogonial Stem Cells Retain States with a Foetal-like Signature. Cells. 2024;13(9).

42. Salehi N, Totonchi M. The construction of a testis transcriptional cell atlas from embryo to adult reveals various somatic cells and their molecular roles. Journal of Translational Medicine. 2023 Nov 27;21(1):859.

43. Wang M(汪梦杰), Li L(李来花), Cheng Q(程青), Zhang H(张浩), Liu Z(刘昭德), Cui Y(崔益强), et al. HumanTestisDB: A Comprehensive Atlas of Testicular Transcriptomes and Cellular Interactions. Genomics Proteomics Bioinformatics. 2025 May 10;23(1).

44. Djureinovic D, Fagerberg L, Hallström B, Danielsson A, Lindskog C, Uhlén M, et al. The human testis-specific proteome defined by transcriptomics and antibody-based profiling. Mol Hum Reprod. 2014 Jun;20(6):476–88.

45. Pineau C, Hikmet F, Zhang C, Oksvold P, Chen S, Fagerberg L, et al. Cell Type-Specific Expression of Testis Elevated Genes Based on Transcriptomics and Antibody-Based Proteomics. J Proteome Res. 2019 Dec 6;18(12):4215–30.

46. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419.

47. Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, et al. The adult human testis transcriptional cell atlas. Cell Res. 2018 Dec;28(12):1141–57.

48. Di Persio S, Tekath T, Siebert-Kuss LM, Cremers JF, Wistuba J, Li X, et al. Single-cell RNA-seq unravels alterations of the human spermatogonial stem cell compartment in patients with impaired spermatogenesis. Cell Rep Med. 2021 Sep 21;2(9):100395.

49. Sohni A, Tan K, Song HW, Burow D, de Rooij DG, Laurent L, et al. The Neonatal and Adult Human Testis Defined at the Single-Cell Level. Cell Rep. 2019 Feb 5;26(6):1501-1517.e4.

50. Shami AN, Zheng X, Munyoki SK, Ma Q, Manske GL, Green CD, et al. Single-Cell RNA Sequencing of Human, Macaque, and Mouse Testes Uncovers Conserved and Divergent Features of Mammalian Spermatogenesis. Dev Cell. 2020 Aug 24;54(4):529-547.e12.

51. Chen S, An G, Wang H, Wu X, Ping P, Hu L, et al. Human obstructive (postvasectomy) and nonobstructive azoospermia – Insights from scRNA-Seq and transcriptome analysis. Genes & Diseases. 2022 May 1;9(3):766–76.

52. Zhao L, Yao C, Xing X, Jing T, Li P, Zhu Z, et al. Single-cell analysis of developing and azoospermia human testicles reveals central role of Sertoli cells. Nat Commun. 2020 Nov 10;11(1):5683.

53. Hermann BP, Cheng K, Singh A, Roa-De La Cruz L, Mutoji KN, Chen IC, et al. The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. Cell Rep. 2018 Nov 6;25(6):1650-1667.e8.

54. Guo J, Nie X, Giebler M, Mlcochova H, Wang Y, Grow EJ, et al. The Dynamic Transcriptional Cell Atlas of Testis Development during Human Puberty. Cell Stem Cell. 2020 Feb 6;26(2):262-276.e4.

55. Nie X, Munyoki SK, Sukhwani M, Schmid N, Missel A, Emery BR, et al. Single-cell analysis of human testis aging and correlation with elevated body mass index. Developmental Cell. 2022 May 9;57(9):1160-1176.e5.

56. Melsted P, Booeshaghi AS, Liu L, Gao F, Lu L, Min KH (Joseph), et al. Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nature Biotechnology. 2021 Jul 1;39(7):813–8.

57. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021 Jun 24;184(13):3573-3587.e29.

58. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019 Dec 23;20(1):296.

59. Waltman L, van Eck NJ. A smart local moving algorithm for large-scale modularity-based community detection. The European Physical Journal B. 2013 Nov 13;86(11):471.

60. Laurentino S, Heckmann L, Di Persio S, Li X, Meyer zu Hörste G, Wistuba J, et al. High-resolution analysis of germ cells from men with sex chromosomal aneuploidies reveals normal transcriptome but impaired imprinting. Clinical Epigenetics. 2019 Aug 28;11(1):127.

61. Li L, Dong J, Yan L, Yong J, Liu X, Hu Y, et al. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. Cell Stem Cell. 2017 Jun 1;20(6):858-873.e4.

62. Tan K, Song HW, Thompson M, Munyoki S, Sukhwani M, Hsieh TC, et al. Transcriptome profiling reveals signaling conditions dictating human spermatogonia fate in vitro. Proc Natl Acad Sci U S A. 2020 Jul 28;117(30):17832–41.

63. Wang M, Liu X, Chang G, Chen Y, An G, Yan L, et al. Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. Cell Stem Cell. 2018 Oct 4;23(4):599-614.e4.

64. Hwang YS, Suzuki S, Seita Y, Ito J, Sakata Y, Aso H, et al. Reconstitution of prospermatogonial specification in vitro from human induced pluripotent stem cells. Nature Communications. 2020 Nov 9;11(1):5656.

65. Guo J, Sosa E, Chitiashvili T, Nie X, Rojas EJ, Oliver E, et al. Single-cell analysis of the developing human testis reveals somatic niche cell specification and fetal germline stem cell establishment. Cell Stem Cell. 2021 Apr 1;28(4):764-778.e4.

66. Bastos GC, Tolezano GC, Krepischi ACV. Rare CNVs and Known Genes Linked to Macrocephaly: Review of Genomic Loci and Promising Candidate Genes. Genes (Basel). 2022 Dec 4;13(12).

67. DeCasien AR, Trujillo AE, Janiak MC, Harshaw EP, Caes ZN, Galindo GA, et al. Equivocal evidence for a link between megalencephaly-related genes and primate brain size evolution. Sci Rep. 2022 Jun 28;12(1):10902.

68. Seidlitz J, Mallard TT, Vogel JW, Lee YH, Warrier V, Ball G, et al. The molecular genetic landscape of human brain size variation. Cell Rep. 2023 Nov 28;42(11):113439.

69. Boddy AM, Harrison PW, Montgomery SH, Caravas JA, Raghanti MA, Phillips KA, et al. Evidence of a Conserved Molecular Response to Selection for Increased Brain Size in Primates. Genome Biol Evol. 2017 Mar 1;9(3):700–13.

70. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet. 2018 Jul;50(7):912–9.

71. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature Genetics. 2018 Aug 1;50(8):1112–21.

72. Wei Y, de Lange SC, Scholtens LH, Watanabe K, Ardesch DJ, Jansen PR, et al. Genetic mapping and evolutionary analysis of human-expanded cognitive networks. Nat Commun. 2019 Oct 24;10(1):4839.

73. Guardiola-Ripoll M, Fatjó-Vilas M. A Systematic Review of the Human Accelerated Regions in Schizophrenia and Related Disorders: Where the Evolutionary and Neurodevelopmental Hypotheses Converge. Int J Mol Sci. 2023 Feb 10;24(4).

74. Somel M, Liu X, Khaitovich P. Human brain evolution: transcripts, metabolites and their regulators. Nat Rev Neurosci. 2013 Feb;14(2):112–27.

75. Wilsch-Bräuninger M, Florio M, Huttner WB. Neocortex expansion in development and evolution - from cell biology to single genes. Curr Opin Neurobiol. 2016 Aug;39:122–32.

76. Lui JH, Hansen DV, Kriegstein AR. Development and Evolution of the Human Neocortex. Cell. 2011 Jul 8;146(1):18–36.

77. Banerjee-Basu S, Packer A. SFARI Gene: an evolving database for the autism research community. Dis Model Mech. 2010 Apr;3(3–4):133–5.

78. Qiu S, Qiu Y, Li Y, Cong X. Genetics of autism spectrum disorder: an umbrella review of systematic reviews and meta-analyses. Translational Psychiatry. 2022 Jun 15;12(1):249.

79. Rylaarsdam L, Guemez-Gamboa A. Genetic Causes and Modifiers of Autism Spectrum Disorder. Front Cell Neurosci. 2019;13:385.

80. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, et al. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron. 2015 Sep 23;87(6):1215–33.

81. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020 Feb 6;180(3):568-584.e23.

82. Owen MJ, Legge SE, Rees E, Walters JTR, O'Donovan MC. Genomic findings in schizophrenia and their implications. Mol Psychiatry. 2023 Sep;28(9):3638–47.

83. Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD, et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. Nature. 2022 Apr 1;604(7906):509–16.

84. Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014 Jul 1;511(7510):421–7.

85. Perucca P, Bahlo M, Berkovic SF. The Genetics of Epilepsy. Annu Rev Genomics Hum Genet. 2020 Aug 31;21:205–30.

86. Rastin C, Schenkel LC, Sadikovic B. Complexity in Genetic Epilepsies: A Comprehensive Review. International Journal of Molecular Sciences. 2023;24(19).

87. Thakran S, Guin D, Singh P, Singh P, Kukal S, Rawat C, et al. Genetic Landscape of Common Epilepsies: Advancing towards Precision in Treatment. International Journal of Molecular Sciences. 2020;21(20).

88. Dhindsa RS, Weido BA, Dhindsa JS, Shetty AJ, Sands CF, Petrovski S, et al. Genome-wide prediction of dominant and recessive neurodevelopmental disorder-associated genes. The American Journal of Human Genetics. 2025 Mar 6;112(3):693–708.

89. Thor S. Indirect neurogenesis in space and time. Nature Reviews Neuroscience. 2024 Aug 1;25(8):519–34.

90. Fang F, Iaquinta PJ, Xia N, Liu L, Diao L, Reijo Pera RA. Transcriptional control of human gametogenesis. Human Reproduction Update. 2022 Jun 1;28(3):313–45.

91. Cheng H, Shang D, Zhou R. Germline stem cells in human. Signal Transduct Target Ther. 2022 Oct 2;7(1):345.

92. von Kopylow K, Spiess AN. Human spermatogonial markers. Stem Cell Research. 2017 Dec 1;25:300–9.

93. Liao C, Walters BW, DiStasio M, Lesch BJ. Human-specific epigenomic states in spermatogenesis. Computational and Structural Biotechnology Journal. 2024 Dec 1;23:577–88.

94. Correia RB, Almeida JM, Wyrwoll MJ, Julca I, Sobral D, Misra CS, et al. The conserved genetic program of male germ cells uncovers ancient regulators of human spermatogenesis. 2024 Sep 3; Available from: http://dx.doi.org/10.7554/eLife.95774.2

95. Houston BJ, Riera-Escamilla A, Wyrwoll MJ, Salas-Huetos A, Xavier MJ, Nagirnaja L, et al. A systematic review of the validated monogenic causes of human male infertility: 2020 update and a discussion of emerging gene-disease relationships. Hum Reprod Update. 2021 Dec 21;28(1):15–29.

96. Zhao Q, Peng H, Chen J, Zhang H, Ma Y, Jiang H. A systematic review and evidence assessment of monogenic gene-disease relationships in human male infertility. Front Endocrinol (Lausanne). 2025;16:1643543.

97. Thumfart KM, Mansuy IM. What are Sertoli cells? Historical, methodological, and functional aspects. Andrology. 2023 Jul 1;11(5):849–59.

98. Wang D, Eraslan B, Wieland T, Hallström B, Hopf T, Zolg DP, et al. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol Syst Biol. 2019 Feb 18;15(2):e8503.

99. Liu J, Liu W, Yang L, Wu Q, Zhang H, Fang A, et al. The Primate-Specific Gene TMEM14B Marks Outer Radial Glia Cells and Promotes Cortical Expansion and Folding. Cell Stem Cell. 2017 Nov 2;21(5):635-649.e8.

100. Nikolic A, Volarevic V, Armstrong L, Lako M, Stojkovic M. Primordial Germ Cells: Current Knowledge and Perspectives. Stem Cells Int. 2016;2016:1741072.

101. Wang S, Wang T, Liu T, Xie RG, Zhao XH, Wang L, et al. Ermin is a p116(RIP) -interacting protein promoting oligodendroglial differentiation and myelin maintenance. Glia. 2020 Nov;68(11):2264–76.

102. Xing L, Kalebic N, Namba T, Vaid S, Wimberger P, Huttner WB. Serotonin Receptor 2A Activation Promotes Evolutionarily Relevant Basal Progenitor Proliferation in the Developing Neocortex. Neuron. 2020 Dec 23;108(6):1113-1129.e6.

103. de Faria O, Dhaunchak AS, Kamen Y, Roth AD, Kuhlmann T, Colman DR, et al. TMEM10 Promotes Oligodendrocyte Differentiation and is Expressed by Oligodendrocytes in Human Remyelinating Multiple Sclerosis Plaques. Scientific Reports. 2019 Mar 5;9(1):3606.

104. Hou Y, Wu Z, Zhang Y, Chen H, Hu J, Guo Y, et al. Functional Analysis of Hydrolethalus Syndrome Protein HYLS1 in Ciliogenesis and Spermatogenesis in Drosophila. Front Cell Dev Biol. 2020;8:301.

105. Vourc'h P, Andres C. Oligodendrocyte myelin glycoprotein (OMgp): evolution, structure and function. Brain Res Brain Res Rev. 2004 May;45(2):115–24.

106. Pääbo S. The human condition-a molecular approach. Cell. 2014 Mar 27;157(1):216–26.

107. Kuhlwilm M, Boeckx C. A catalog of single nucleotide changes distinguishing modern humans from archaic hominins. Scientific Reports. 2019 Jun 11;9(1):8463.

108. Salemi M, Longo GA, La Vignera S, Romano C, Condorelli RA, Romano C, et al. SPAG5 mRNA is over-expressed in peripheral blood leukocytes of patients with Down's syndrome and cryptorchidism. Neurological Sciences. 2013 Apr 1;34(4):549–51.

109. Xue J, Tarnasky HA, Rancourt DE, van Der Hoorn FA. Targeted disruption of the testicular SPAG5/deepest protein does not affect spermatogenesis or fertility. Mol Cell Biol. 2002 Apr;22(7):1993–7.

110. Chen M, Wang D, Xu Y, Yang C. Upregulation of sperm-associated antigen 5 expression in endometrial carcinoma was associated with poor prognosis and immune dysregulation, and promoted cell migration and invasion. Scientific Reports. 2024 Jun 11;14(1):13415.

111. Fietz SA, Huttner WB. Cortical progenitor expansion, self-renewal and neurogenesis-a polarized perspective. Curr Opin Neurobiol. 2011 Feb;21(1):23–35.

112. Knierim E, Hirata H, Wolf NI, Morales-Gonzalez S, Schottmann G, Tanaka Y, et al. Mutations in Subunits of the Activating Signal Cointegrator 1 Complex Are Associated with Prenatal Spinal Muscular Atrophy and Congenital Bone Fractures. The American Journal of Human Genetics. 2016 Mar 3;98(3):473–89.

113. Francks C, Maegawa S, Laurén J, Abrahams BS, Velayos-Baeza A, Medland SE, et al. LRRTM1 on chromosome 2p12 is a maternally suppressed gene that is associated paternally with handedness and schizophrenia. Molecular Psychiatry. 2007 Dec 1;12(12):1129–39.

114. Kuwahara A, Sakai H, Xu Y, Itoh Y, Hirabayashi Y, Gotoh Y. Tcf3 represses Wnt-β-catenin signaling and maintains neural stem cell population during neocortical development. PLoS One. 2014;9(5):e94408.

115. Knol MJ, Poot RA, Evans TE, Satizabal CL, Mishra A, Sargurupremraj M, et al. Genetic variants for head size share genes and pathways with cancer. Cell Reports Medicine. 2024 May 21;5(5):101529.

116. Hansen DV, Lui JH, Parker PRL, Kriegstein AR. Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature. 2010 Mar 1;464(7288):554–61.

117. Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP, et al. The genetic architecture of the human cerebral cortex. Science. 2020 Mar 20;367(6484).

118. Young JC, Kerr G, Micati D, Nielsen JE, Rajpert-De Meyts E, Abud HE, et al. WNT signalling in the normal human adult testis and in male germ cell neoplasms. Hum Reprod. 2020 Sep 1;35(9):1991–2003.

119. Aloisio GM, Nakada Y, Saatcioglu HD, Peña CG, Baker MD, Tarnawa ED, et al. PAX7 expression defines germline stem cells in the adult testis. J Clin Invest. 2014 Sep;124(9):3929–44.

120. Fu K, Tian S, Tan H, Wang C, Wang H, Wang M, et al. Biological and RNA regulatory function of MOV10 in mammalian germ cells. BMC Biology. 2019 May 14;17(1):39.

121. Venditti M, Fasano C, Minucci S, Serino I, Sinisi AA, Dale B, et al. DAAM1 and PREP are involved in human spermatogenesis. Reprod Fertil Dev. 2020 Mar;32(5):484–94.

122. Li X, Yu Y, Zhang Q, Luo X, Yu L, Zhao Z. Effect of HDAC9 on the differentiation of chicken embryonic stem cells into male germ cells. Anim Reprod. 2024;21(2):e20240011.

123. Girskis KM, Stergachis AB, DeGennaro EM, Doan RN, Qian X, Johnson MB, et al. Rewiring of human neurodevelopmental gene regulatory programs by human accelerated regions. Neuron. 2021 Oct 20;109(20):3239-3251.e7.

124. Whalen S, Inoue F, Ryu H, Fair T, Markenscoff-Papadimitriou E, Keough K, et al. Machine learning dissection of human accelerated regions in primate neurodevelopment. Neuron. 2023 Mar 15;111(6):857-873.e8.

125. Sasaki K, Yokobayashi S, Nakamura T, Okamoto I, Yabuta Y, Kurimoto K, et al. Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. Cell Stem Cell. 2015 Aug 6;17(2):178–94.

126. Yao C, Yao R, Luo H, Shuai L. Germline specification from pluripotent stem cells. Stem Cell Res Ther. 2022 Feb 21;13(1):74.

127. Strome S, Updike D. Specifying and protecting germ cell fate. Nature Reviews Molecular Cell Biology. 2015 Jul 1;16(7):406–16.

128. Updike DL, Knutson AK, Egelhofer TA, Campbell AC, Strome S. Germ-granule components prevent somatic development in the C. elegans germline. Curr Biol. 2014 May 5;24(9):970–5.

129. Barton RA, Harvey PH. Mosaic evolution of brain structure in mammals. Nature. 2000 Jun 29;405(6790):1055–8.

130. Marques-Bonet T, Kidd JM, Ventura M, Graves TA, Cheng Z, Hillier LW, et al. A burst of segmental duplications in the genome of the African great ape ancestor. Nature. 2009 Feb 1;457(7231):877–81.

131. Holland LZ, Short S. Gene duplication, co-option and recruitment during the origin of the vertebrate brain from the invertebrate chordate brain. Brain Behav Evol. 2008;72(2):91–105.

132. Castillo-Morales A, Monzón-Sandoval J, de Sousa AA, Urrutia AO, Gutierrez H. Neocortex expansion is linked to size variations in gene families with chemotaxis, cell-cell signalling and immune response functions in mammals. Open Biol. 2016 Oct;6(10).

133. Castillo-Morales A, Monzón-Sandoval J, Urrutia AO, Gutiérrez H. Increased brain size in mammals is associated with size variations in gene families with cell signalling, chemotaxis and immune-related functions. Proc Biol Sci. 2014 Jan 22;281(1775):20132428.

134. Kilili H, Padilla-Morales B, Castillo-Morales A, Monzón-Sandoval J, Díaz-Barba K, Cornejo-Paramo P, et al. Maximum lifespan and brain size in mammals are associated with gene family size expansion related to immune system functions. Scientific Reports. 2025 Apr 29;15(1):15087.

135. Uddin M, Opazo JC, Wildman DE, Sherwood CC, Hof PR, Goodman M, et al. Molecular evolution of the cytochrome c oxidase subunit 5A gene in primates. BMC Evol Biol. 2008 Jan 15;8:8.

136. Vallender EJ, Lahn BT. A primate-specific acceleration in the evolution of the caspase-dependent apoptosis pathway. Hum Mol Genet. 2006 Oct 15;15(20):3034–40.

137. Zhang R, Wang YQ, Su B. Molecular evolution of a primate-specific microRNA family. Mol Biol Evol. 2008 Jul;25(7):1493–502.

138. Haygood R, Fedrigo O, Hanson B, Yokoyama KD, Wray GA. Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. Nat Genet. 2007 Sep;39(9):1140–4.

139. Smaers JB, Rothman RS, Hudson DR, Balanoff AM, Beatty B, Dechmann DKN, et al. The evolution of mammalian brain size. Sci Adv. 2021 Apr;7(18).

140. Herculano-Houzel S. The glia/neuron ratio: how it varies uniformly across brain structures and species and what that means for brain physiology and evolution. Glia. 2014 Sep;62(9):1377–91.

141. Fiddes IT, Lodewijk GA, Mooring M, Bosworth CM, Ewing AD, Mantalas GL, et al. Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical Neurogenesis. Cell. 2018 May 31;173(6):1356-1369.e22.

142. Lindhout FW, Krienen FM, Pollard KS, Lancaster MA. A molecular and cellular perspective on human brain evolution and tempo. Nature. 2024 Jun;630(8017):596–608.

143. Lindhout FW, Szafranska HM, Imaz-Rosshandler I, Guglielmi L, Moarefian M, Voitiuk K, et al. Calcium dynamics tune developmental tempo to generate evolutionarily divergent axon tract lengths. bioRxiv. 2025 Jan 1;2024.12.28.630576.

144. Moreira BP, Oliveira PF, Alves MG. Molecular Mechanisms Controlled by mTOR in Male Reproductive System. Int J Mol Sci. 2019 Apr 2;20(7).

145. Arenella M, Mota NR, Teunissen MWA, Brunner HG, Bralten J. Autism spectrum disorder and brain volume link through a set of mTOR-related genes. J Child Psychol Psychiatry. 2023 Jul;64(7):1007–14.

146. Reijnders MRF, Kousi M, van Woerden GM, Klein M, Bralten J, Mancini GMS, et al. Variation in a range of mTOR-related genes associates with intracranial volume and intellectual disability. Nat Commun. 2017 Oct 20;8(1):1052.

147. Aitken RJ. Paternal age, de novo mutations, and offspring health? New directions for an ageing problem. Hum Reprod. 2024 Dec 1;39(12):2645–54.

148. Khoshkhoo S, Bae M, Wang Y, Tillett A, Ramirez RB, Finander B, et al. Activating Ras-MAPK pathway variants drive hippocampal clonal competition in human epilepsy. bioRxiv. 2026 Jan 1;2026.01.26.701822.

149. Yavuz BR, Arici MK, Demirel HC, Tsai CJ, Jang H, Nussinov R, et al. Neurodevelopmental disorders and cancer networks share pathways, but differ in mechanisms, signaling strength, and outcome. npj Genomic Medicine. 2023 Nov 4;8(1):37.

150. Neville MDC, Lawson ARJ, Sanghvi R, Abascal F, Pham MH, Cagan A, et al. Sperm sequencing reveals extensive positive selection in the male germline. Nature. 2025 Nov 1;647(8089):421–8.

151. Skene NG, Bryois J, Bakken TE, Breen G, Crowley JJ, Gaspar HA, et al. Genetic identification of brain cell types underlying schizophrenia. Nat Genet. 2018 Jun;50(6):825–33.

530

Figures

533

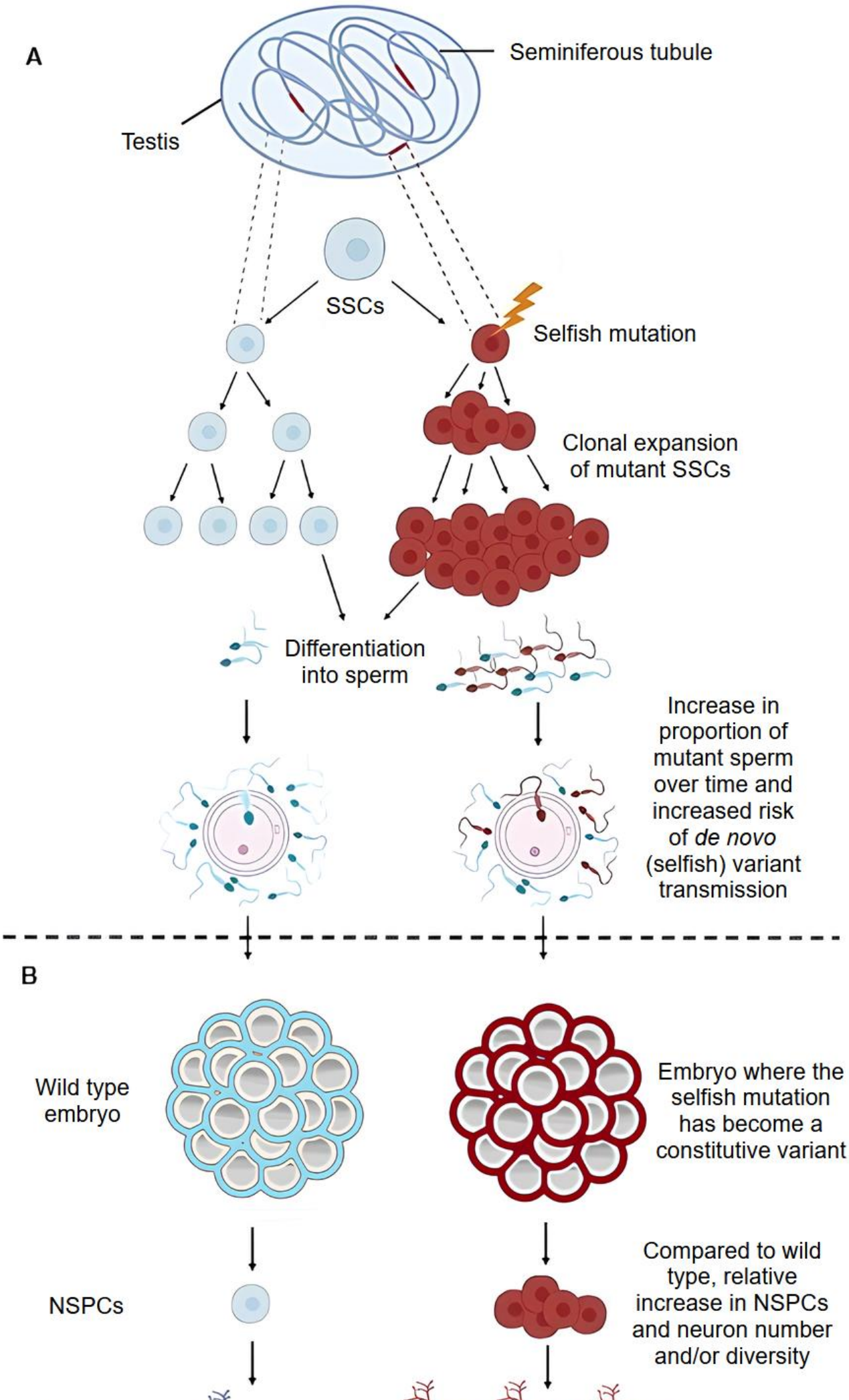

A
Seminiferous tubule
Testis
SSCs
Selfish mutation
Clonal expansion
of mutant SSCs
Differentiation
into sperm
Increase in
proportion of
mutant sperm
over time and
increased risk
of *de novo*
(selfish) variant
transmission
B
Wild type
embryo
Embryo where the
selfish mutation
has become a
constitutive variant
NSPCs
Compared to wild
type, relative
increase in NSPCs
and neuron number
and/or diversity

Figure 1. Variants underlying relative increases in brain size (through expansion of the neural progenitor pool) can be introduced into the genome by a male germline-specific evolutionary force, selfish spermatogonial selection.

(A) In selfish spermatogonial selection, spontaneous oncogenic-like mutations occur in genes functionally involved in spermatogonial stem cell (SSC) regulation. These selfish mutations confer a selective advantage to SSCs over their wild-type counterparts (blue), resulting in the clonal expansion of mutant SSCs (red) within the seminiferous tubules, and increasing the relative proportion of mutant sperm over time. Sperm carrying selfish mutations often give rise to offspring with neurodevelopmental disorders, although not all selfish mutations are inherently deleterious. This panel is adapted from that originally published in (12). (B) As human development begins from a single zygotic cell, this early bottleneck effectively fixes any selfish variants transmitted via sperm in the resulting embryo. As such, 'mild' selfish mutations (ones selectively beneficial to SSCs and fertility, yet not overly detrimental when present constitutively in the offspring) may become a source of heritable material and propagate across generations. As selfish selection preferentially enriches the genome for mutations affecting SSC regulation, and given that many of the same genes are also expressed in neural stem and progenitor cells (NSPCs), we propose that this process may have introduced functional variants that expanded the NSPC pool. In principle, such variants could contribute to increased brain size, regional diversification and/or developmental complexity.

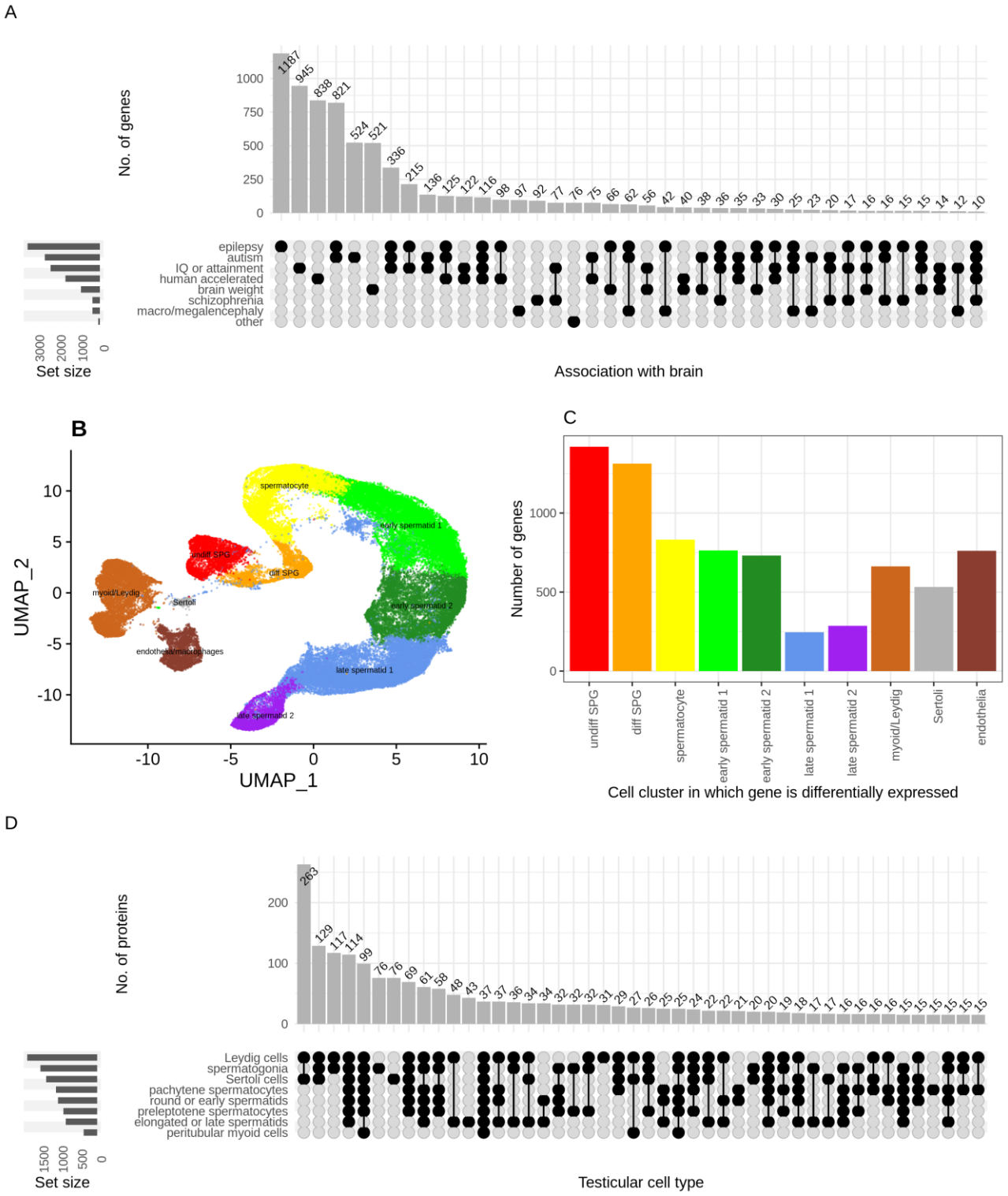


Figure 2. Brain-associated genes and their expression in the human testis.
(A) Distribution of 7193 genes among an inclusive set of brain-associated phenotypes, both explicit (such as macro- or megalencephaly) and suggestive (such as 'educational attainment', a proxy for cognitive function), detailed further in notes to **Supplementary Table 1**. Sets containing < 10 genes are not shown. **(B)** Single-cell expression atlas of the adult human testis (n = 60,427 cells), annotated into ten cell clusters (3 somatic, 7 germline). **(C)** Number of genes differentially expressed, at the transcript level, in at least one of these ten clusters (raw data in **Supplementary**

Tables 2 and 3). **(D)** Protein-level expression of 2049 brain-associated genes for which antibody staining data was available for each of 8 testicular cell types (3 somatic, 5 germline) and with a protein detected in at least one (raw data from https://v23.proteinatlas.org/about/download/normal_tissue.tsv, accessed 17th June 2023). Sets represent detectable protein-level expression (i.e. classification into either of the HPA categories of low, medium or high expression) in each cell type (raw data in **Supplementary Table 1**). Sets containing < 15 proteins are not shown. A version of this figure restricted to 5212 genes with five explicit brain-associated phenotypes (postmortem brain weight, macrocephaly/megalencephaly, epilepsy, autism, and schizophrenia) is presented as **Supplementary Figure 1** and shows quantitatively similar results.

Figure 3. Number of brain-associated genes whose proteins are detectably expressed in human tissues.
(A-G) Distribution of brain-associated genes for seven different phenotypic categories defined in **Supplementary Table 1**. **(H-I)** Subsets of genes associated with at least 2 (H) or 3 (I) different phenotypic categories, excluding the two 'indirect' categories of IQ/attainment and human-acceleration. Raw data for this figure, representing protein expression level per cell type per tissue,

was obtained from the Human Protein Atlas v23.0 (https://v23.proteinatlas.org/about/download/normal_tissue.tsv, downloaded 17th June 2023). A protein was considered expressed in a given tissue if, in any cell type of that tissue, 'level' was either 'low', 'medium' or 'high' and 'reliability' was not 'uncertain'. Cell types per tissue are detailed in **Supplementary Table 6**. For ease of comparison, brain tissues are shown in blue, testis in red, and all others in grey, with the black dashed line indicating the number of proteins detected in the testis.

Tables

Evidence	Source
Strong (pathogenic) selfish spermatogonial mutations cause neurodevelopmental disorders widely associated with macrocephaly, establishing an empirical link between cell proliferation in the testis and developing brain.	Reviewed in (12,15,16).
Macrocephaly frequently occurs in individuals with pathogenic variants in mTOR-related genes, key mediators of spermatogonial stem cell maintenance and differentiation (144).	Results from (145), using 101 mTOR-related genes defined by (146).
Selfish spermatogonial selection is a prevalent natural phenomenon in the ageing testis, affecting numerous genes involved in testicular homeostasis – many of which are also expressed in the brain.	14 genes are already known to harbour selfish spermatogonial mutations (12,17), with 27 new candidates proposed by large-scale trio sequencing (21). While historically the majority of selfish mutations were identified in components of the RTK/RAS/MAPK pathway, they are not restricted to it (17); in principle, selfish mutations can occur in any gene capable of altering SSC proliferation or survival (147).
Activating mutations in RTK/RAS/MAPK pathway genes (which maintain the most primitive SSC state (20)) are associated with the positive clonal selection of hippocampal neural progenitor cells; this in turn leads to localised brain overgrowth, resulting in focal epilepsies.	Results from (148). In a mechanistic parallel to selfish spermatogonial selection, this study found a higher frequency of somatic RAS/MAPK variants in older individuals, whose first known seizures were as adults.
Non-penetrant (or mildly deleterious) selfish mutations can be inherited and maintained across generations, allowing their effects on specific signalling pathways to accumulative additively across generations.	Discussed in this work; see also (149).
Many genes under positive selection in the human male germline can be associated with macrocephaly or broader aspects of brain development.	Of 40 genes positively selected in the human male germline (150), 29 are associated with brain growth, function, development, or evolution, listed in **Supplementary Table 1** (*ARHGAP35, BMPR2, BRAF, CBL, CTNNB1, CUL3, DDX3X, EP300, FGFR2, FGFR3, HRAS, KDM5B, KDM5C, KMT2D, KMT2E, KRAS, MIB1, NF1, NSD1, PTEN, PTPN11, RAF1, RASA2, RBM12, RIT1, ROBO1, SCAF4, SMAD4, TCF12*). Three others (*CSNK2B*, *FAM222B*, *PRRC2A*) are family

	members of those in **Supplementary Table 1** (<u>*CSNK2A1*</u>, *FAM222A*, *PRRC2C*). Underlined genes are those in which pathogenic variants are associated with either macrocephaly or megalencephaly.
In silico evolutionary modelling suggests that hominin brain expansion arose as a by-product of selection acting on fertility-related traits, rather than from direct selection on brain size.	Results from (35,36).
Of thousands of proteins functionally associated with the brain, more are expressed in testis than any other organ beside the brain, with many expressed in male germ cells, particularly spermatogonia. The converse is also true: genes enriched in human spermatogonia, implicated in human-specific roles in spermatogenesis, associated with monogenic male infertility, or part of the conserved metazoan spermatogenic program, are widely expressed throughout the brain.	Results shown in **Figures 2 and 3**, **Supplementary Figures 2 to 9**, **Supplementary Tables 1 and 5**, and discussed further in the **Supplementary Text**, having drawn on multiple gene lists from previous studies (90–96).
Many genes have been functionally characterised with dual roles in both the male germline and neural stem and progenitor cells, with several harbouring specific mutations associated with human brain expansion.	Evidence for a dual functional role in humans is given for ten genes in (13), with additional examples in the **Supplementary Text**. Genes with shared roles in both the germline and nervous systems of other metazoans are given in (27).

Table 1. Overview of the lines of evidence implicating a link between the male germline, and spermatogonia in particular, and human brain expansion.

<u>Supplementary Figures</u>

Supplementary Figure 1. A conservative set of brain-associated genes and their expression in the human testis.
(A) Distribution of 5212 genes among an explicit set of brain-associated phenotypes, detailed further in notes to **Supplementary Table 1**. Sets containing < 10 genes are not shown. **(B)** Single-cell expression atlas of the adult human testis (n = 60,427 cells), annotated into ten cell clusters (3 somatic, 7 germline). **(C)** Number of genes differentially expressed, at the transcript level, in at least one of these ten clusters (raw data in **Supplementary Tables 2 and 3**). **(D)** Protein-level expression of 1573 brain-associated genes for which antibody staining data was available for each of 8 testicular cell types (3 somatic, 5 germline) and with a protein detected in at least one (raw data from https://v23.proteinatlas.org/about/download/normal_tissue.tsv, accessed 17th June 2023). Sets represent detectable protein-level expression (i.e. classification into either of the HPA categories of low, medium or high expression) in each cell type (raw data in **Supplementary Table 1**). Sets containing < 15 proteins are not shown.

Supplementary Figure 2. Average expression across, respectively, an (A) inclusive and (B) conservative set of 7193 and 5212 brain-associated genes for each of ten cell clusters in an adult human testis expression atlas.

The brain-associated gene sets are detailed in **Supplementary Table 1** with expression level per testicular cell type in **Supplementary Table 2**. Statistical comparisons are Mann-Whitney U tests, implemented using the R package 'ggpubr.'

Supplementary Figure 3. Number of differentially expressed brain-associated genes per cell cluster, for each of four testis and male germline single-cell atlases, and using an inclusive set of brain-associated genes.
For each of the four single-cell atlases, barplots show the number of genes significantly differentially expressed within an inclusive set of brain-associated genes (n=7193). The cell atlases and their respective annotations are described in **Supplementary Table 4**, alongside the differential expression analysis results, which were filtered according to a common set of criteria. The brain-associated gene sets are detailed in **Supplementary Table 1**.

Supplementary Figure 4. Number of differentially expressed brain-associated genes per cell cluster, for each of four testis and male germline single-cell atlases, and using a conservative set of brain-associated genes.
For each of the four single-cell atlases, barplots show the number of genes significantly differentially expressed within an inclusive set of brain-associated genes (n=5212). The cell atlases and their respective annotations are described in **Supplementary Table 4**, alongside the differential expression analysis results, which were filtered according to a common set of criteria. The brain-associated gene sets are detailed in **Supplementary Table 1**.

Supplementary Figure 5. Number of brain-associated genes differentially expressed, at the transcript level, in one or more of ten testicular cell clusters.
The list of brain-associated genes is given in **Supplementary Table 1**, with raw data for this figure in **Supplementary Tables 2 and 3**. Panels A to G show the distribution of differentially expressed genes for seven different categories of phenotype, whereas panels H and I require that each gene is associated with at least 2 or 3 different categories, respectively, not including the two 'indirect' phenotypes of IQ/attainment and human-acceleration.

Supplementary Figure 6. Number of brain-associated genes differentially expressed, at the transcript level, in one or more of cell clusters from the 'Salehi 2023' single-cell expression atlas.
The 'Salehi 2023' expression atlas, its respective annotations, and the differential expression analysis results used for this figure, are detailed in **Supplementary Table 4**. The brain-associated gene sets are detailed in **Supplementary Table 1**. Panels A to G show the distribution of differentially expressed genes for seven different categories of phenotype, whereas panels H and I require that each gene is associated with at least 2 or 3 different categories, respectively, not including the two 'indirect' phenotypes of IQ/attainment and human-acceleration.

Supplementary Figure 7. Number of brain-associated genes differentially expressed, at the transcript level, in one or more of cell clusters from the 'Wang 2025 (all cells)' single-cell expression atlas.
The 'Wang 2025 (all cells)' expression atlas, its respective annotations, and the differential expression analysis results used for this figure, are detailed in **Supplementary Table 4**. The brain-associated gene sets are detailed in **Supplementary Table 1**. Panels A to G show the distribution of differentially expressed genes for seven different categories of phenotype, whereas panels H and I require that each gene is associated with at least 2 or 3 different categories, respectively, not including the two 'indirect' phenotypes of IQ/attainment and human-acceleration.

Supplementary Figure 8. Number of brain-associated genes differentially expressed, at the transcript level, in one or more of cell clusters from the 'Wang 2025 (germ cells)' single-cell expression atlas.
The 'Wang 2025 (germ cells)' expression atlas, its respective annotations, and the differential expression analysis results used for this figure, are detailed in **Supplementary Table 4**. The brain-associated gene sets are detailed in **Supplementary Table 1**. Panels A to G show the distribution of differentially expressed genes for seven different categories of phenotype, whereas panels H and I require that each gene is associated with at least 2 or 3 different categories, respectively, not including the two 'indirect' phenotypes of IQ/attainment and human-acceleration.

Supplementary Figure 9. Number of brain-associated genes differentially expressed, at the transcript level, in one or more of cell clusters from the 'Wang 2025 (germ cells – part 1)' single-cell expression atlas.
The 'Wang 2025 (germ cells – part 1)' expression atlas, its respective annotations, and the differential expression analysis results used for this figure, are detailed in **Supplementary Table 4**. The brain-associated gene sets are detailed in **Supplementary Table 1**. Panels A to G show the distribution of differentially expressed genes for seven different categories of phenotype, whereas panels H and I require that each gene is associated with at least 2 or 3 different categories, respectively, not including the two 'indirect' phenotypes of IQ/attainment and human-acceleration.

Supplementary Figure 10. Beeswarm (violin scatter) plot of the cellular specificity, in the mouse brain, of 14 genes known to harbour selfish spermatogonial mutations.
Cellular specificity values, from the Karolinska superset (151), represent the proportion of the total expression of a gene attributable to one of 24 brain cell types (a specificity of 0 means that the gene is not expressed in that cell type and a value of 1 that it is only expressed in that cell type). The set of 14 genes known to harbour one or more selfish spermatogonial mutations have been identified in humans (12,17) although for the purpose of visualisation here, specificity data from the one-to-one mouse orthologue is used.

Supplementary Figure 11. Violin plot of the cellular specificity of 218 spermatogonia and spermatogenesis-associated genes in the mouse brain.
Cellular specificity values, from the Karolinska superset (151), represent the proportion of the total expression of a gene attributable to one of 24 brain cell types (a specificity of 0 means that the gene is not expressed in that cell type and a value of 1 that it is only expressed in that cell type). Violins are sorted from left to right in ascending order of median specificity. The genes used for this plot are listed in **Supplementary Table 5** and drawn from previous studies (90–94). The genes in this table are human and for the purpose of visualisation here, the one-to-one mouse orthologue (if extant) is used.

Supplementary Figure 12. Violin plot of the cellular specificity of 154 male infertility-associated genes in the mouse brain.
Cellular specificity values, from the Karolinska superset (151), represent the proportion of the total expression of a gene attributable to one of 24 brain cell types (a specificity of 0 means that the gene is not expressed in that cell type and a value of 1 that it is only expressed in that cell type). Violins are sorted from left to right in ascending order of median specificity. The genes used for this plot are listed in **Supplementary Table 5** and drawn from two previous systematic reviews (95,96). The genes in this table are human and for the purpose of visualisation here, the one-to-one mouse orthologue (if extant) is used.

Supplementary Figure 13. Boxplot of the expression in the cerebellar cortex over time of 14 genes known to harbour selfish spermatogonial mutations, using the Allen BrainSpan Atlas. The set of 14 genes known to harbour one or more selfish spermatogonial mutations have been described in previous studies (12,17).

Supplementary Figure 14. Violin plot of the expression in the cerebellar cortex over time of 218 spermatogonia and spermatogenesis-associated genes, using the Allen BrainSpan Atlas. The genes used for this plot are listed in **Supplementary Table 5** and drawn from previous studies (90–94).

Supplementary Figure 15. Violin plot of the expression in the cerebellar cortex over time of 154 male infertility-associated genes, using the Allen BrainSpan Atlas. The genes used for this plot are listed in **Supplementary Table 5** and drawn from two previous systematic reviews (95,96).

Supplementary Tables

Supplementary Table 1. Summary of the testicular expression, at the protein-level, of 7193 genes associated with brain growth, function, development, or evolution.

Supplementary Table 2. Transcript-level expression of genes associated with brain growth, function, development, or evolution, using an adult human testis single-cell expression atlas.

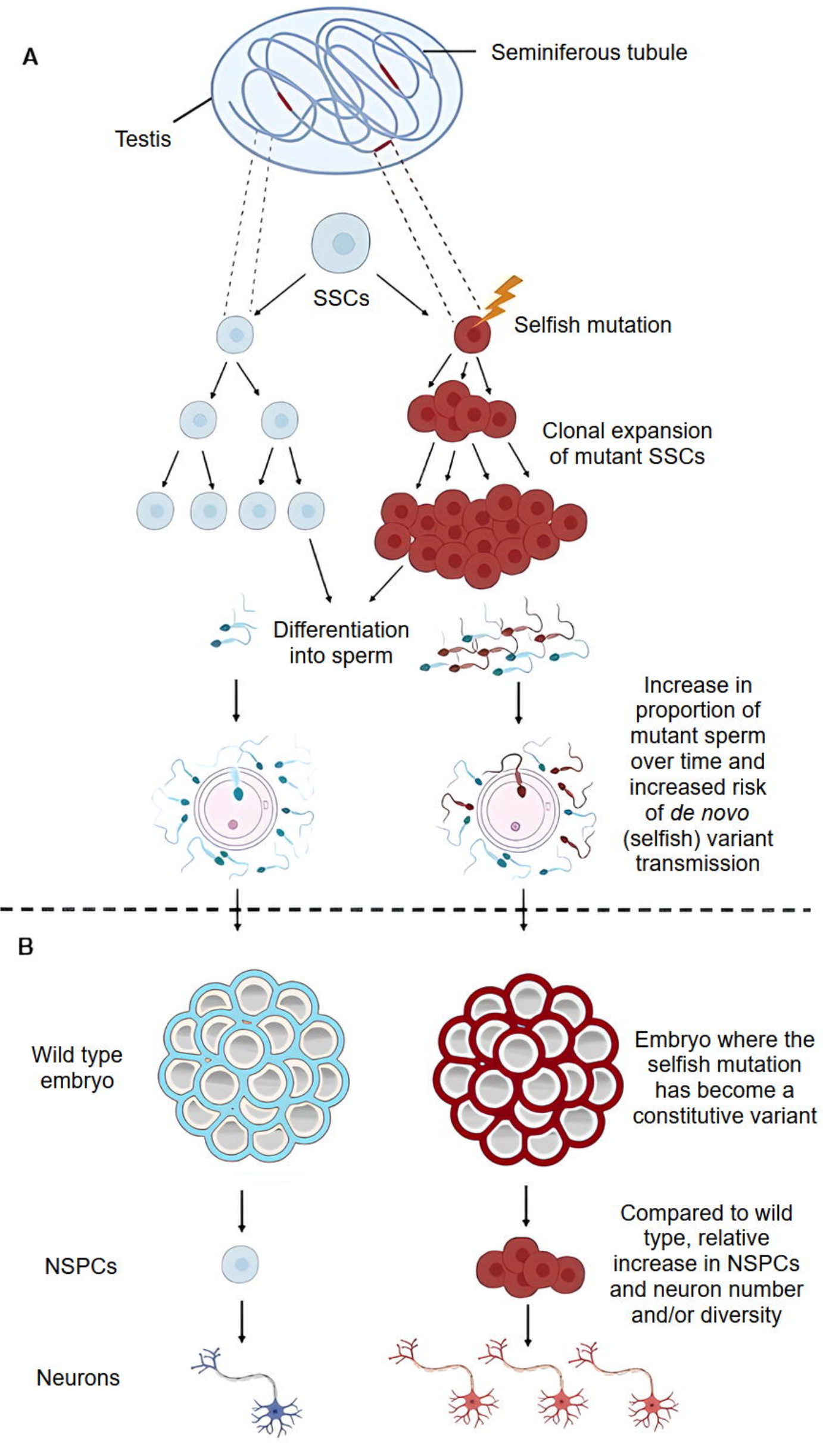
Supplementary Table 3. Summary of the testicular expression, at the transcript-level, of genes associated with brain growth, function, development, or evolution.

Supplementary Table 4. Sources of cluster (cell type) annotations, and Seurat differential gene expression analysis results, for four previously published single-cell human testis atlases.

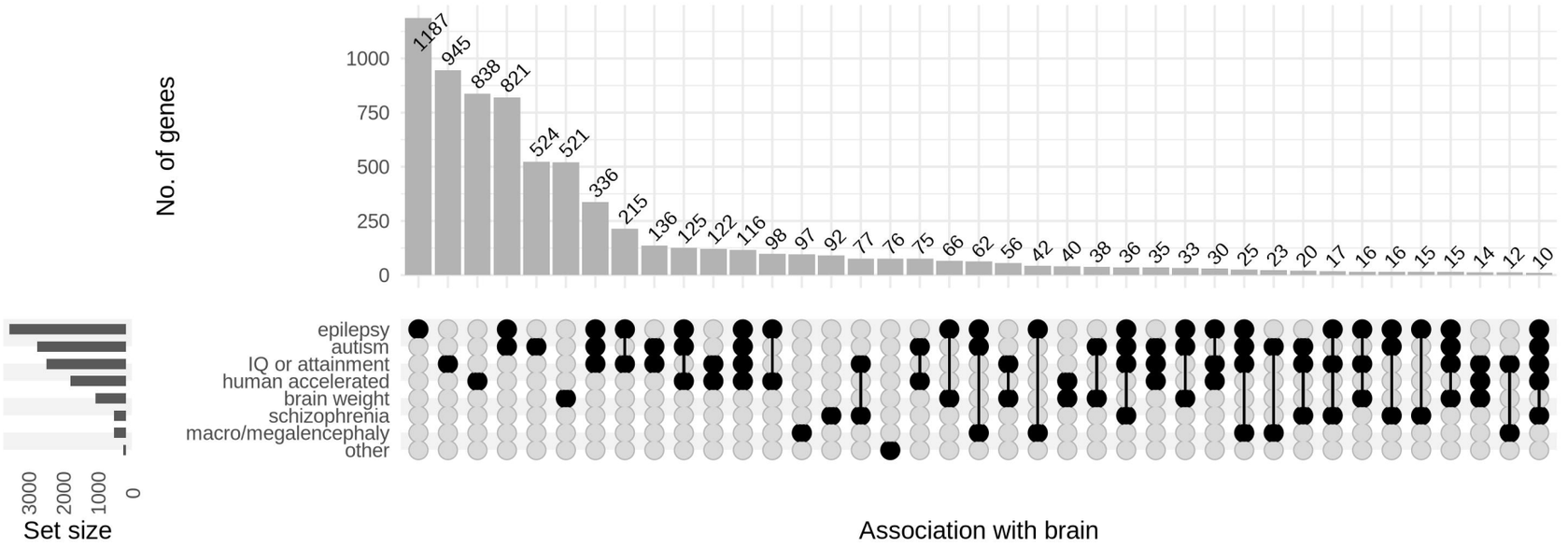
Supplementary Table 5. Transcript-level expression in the brain of spermatogonia-enriched genes, human-specific germline-associated genes, genes associated with human male infertility, and core components of the metazoan spermatogenic program.

Supplementary Table 6. Proportion of male germline-expressed proteins detectably expressed in other tissues.

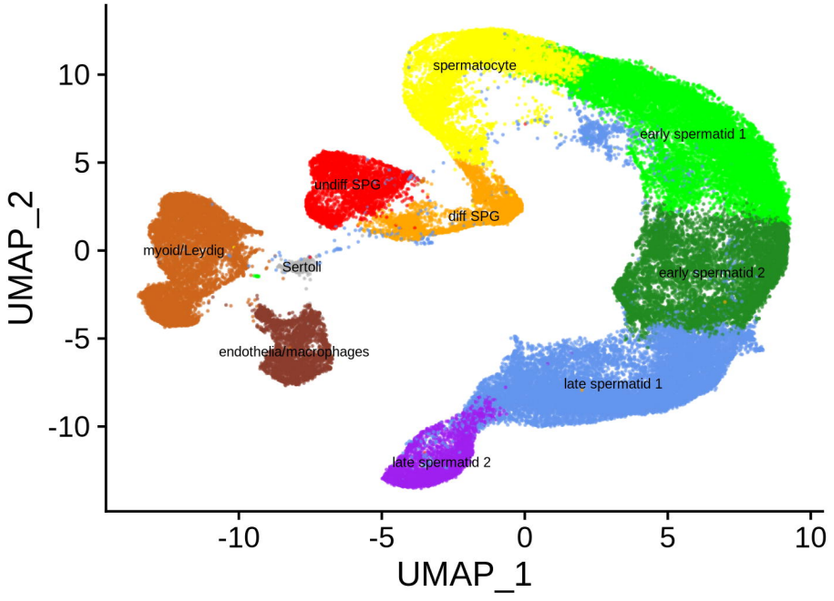
Supplementary Table 7. Summary of the testicular expression, at the protein-level, of protein-coding genes containing human-specific high-frequency missense changes.


A
Seminiferous tubule
Testis
SSCs
Selfish mutation
Clonal expansion of mutant SSCs
Differentiation into sperm
Increase in proportion of mutant sperm over time and increased risk of *de novo* (selfish) variant transmission
B
Wild type embryo
Embryo where the selfish mutation has become a constitutive variant
NSPCs
Compared to wild type, relative increase in NSPCs and neuron number and/or diversity
Neurons

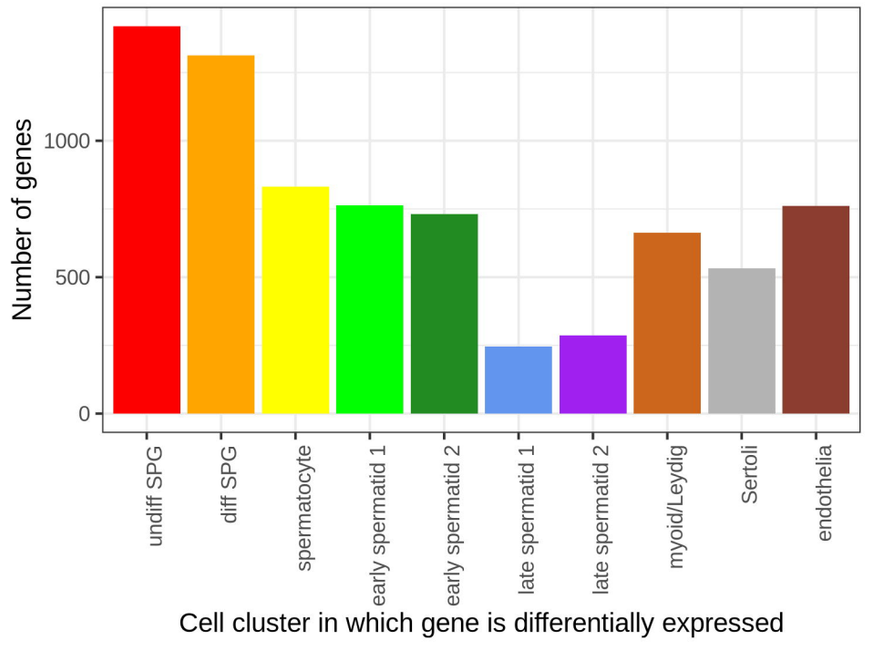
A


No. of genes
1000
750
500
250
0
1187 945 838 821 524 521 336 215 136 125 122 116 98 97 92 77 76 75 66 62 56 42 40 38 36 35 33 30 25 23 20 17 16 16 15 15 14 12 10
epilepsy
autism
IQ or attainment
human accelerated
brain weight
schizophrenia
macro/megalencephaly
other
3000
2000
1000
0
Set size
Association with brain

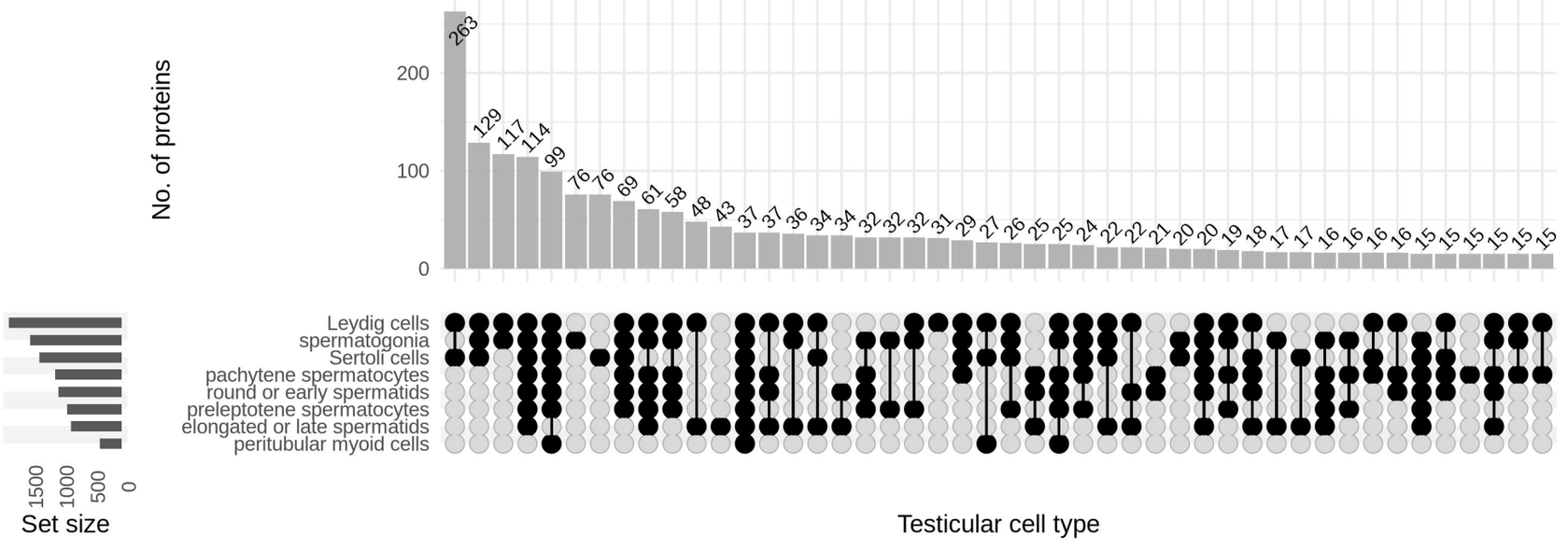
B

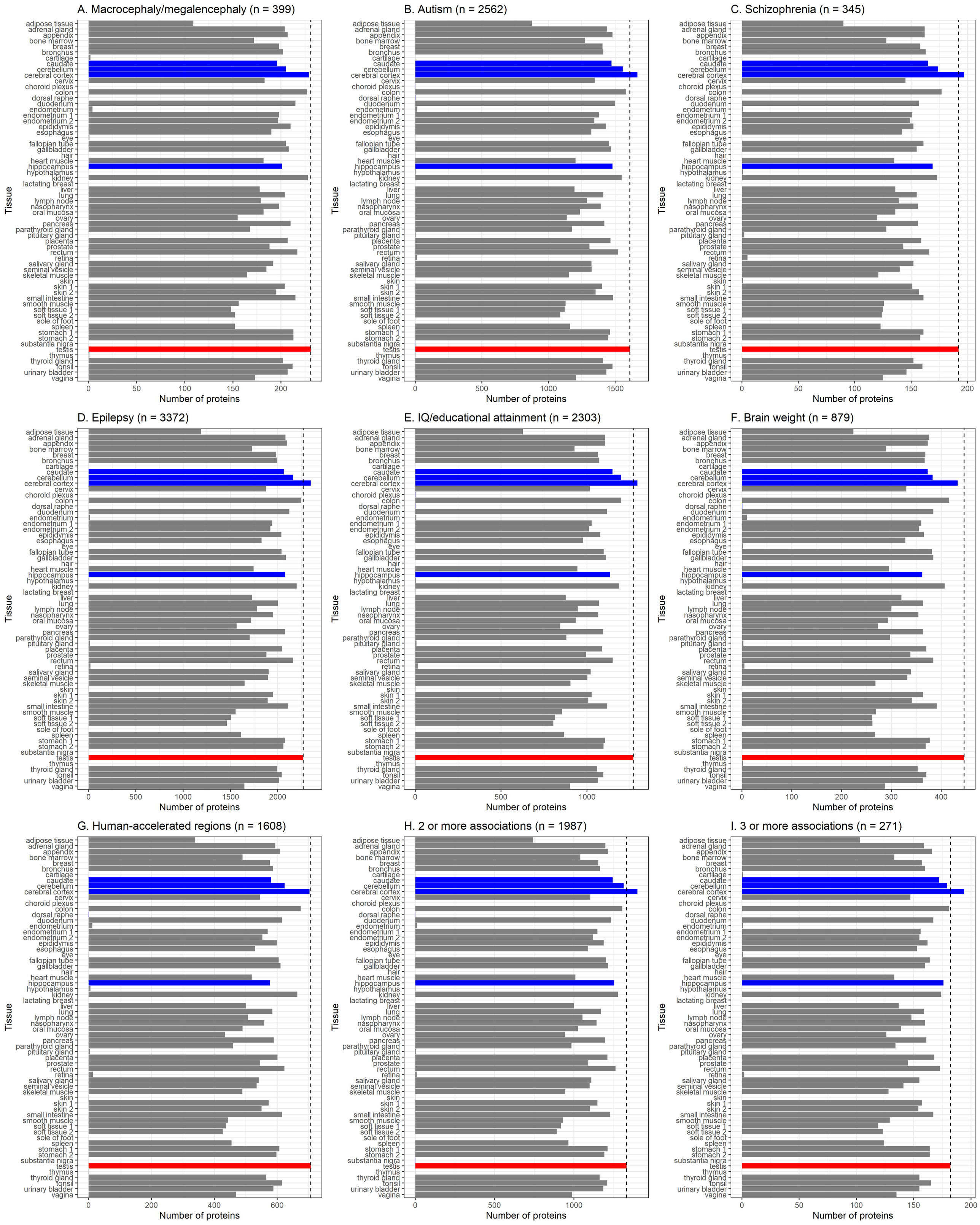

spermatocyte
early spermatid 1
undiff SPG
diff SPG
myoid/Leydig
Sertoli
early spermatid 2
endothelia/macrophages
late spermatid 1
late spermatid 2
UMAP_2
UMAP_1
10
5
0
-5
-10

C


Number of genes
1000
500
0
undiff SPG
diff SPG
spermatocyte
early spermatid 1
early spermatid 2
late spermatid 1
late spermatid 2
myoid/Leydig
Sertoli
endothelia
Cell cluster in which gene is differentially expressed

D


No. of proteins
200
100
0
263 129 117 114 99 76 76 69 61 58 48 43 37 37 36 34 34 32 32 32 31 29 27 26 25 25 24 22 22 21 20 20 19 18 17 17 16 16 16 16 15 15 15 15 15 15
Leydig cells
spermatogonia
Sertoli cells
pachytene spermatocytes
round or early spermatids
preleptotene spermatocytes
elongated or late spermatids
peritubular myoid cells
1500
1000
500
0
Set size
Testicular cell type


A. Macrocephaly/megalencephaly (n = 399)
B. Autism (n = 2562)
C. Schizophrenia (n = 345)
D. Epilepsy (n = 3372)
E. IQ/educational attainment (n = 2303)
F. Brain weight (n = 879)
G. Human-accelerated regions (n = 1608)
H. 2 or more associations (n = 1987)
I. 3 or more associations (n = 271)
Tissue
adipose tissue
adrenal gland
appendix
bone marrow
breast
bronchus
cartilage
caudate
cerebellum
cerebral cortex
cervix
choroid plexus
colon
dorsal raphe
duodenum
endometrium
endometrium 1
endometrium 2
epididymis
esophagus
eye
fallopian tube
gallbladder
hair
heart muscle
hippocampus
hypothalamus
kidney
lactating breast
liver
lung
lymph node
nasopharynx
oral mucosa
ovary
pancreas
parathyroid gland
pituitary gland
placenta
prostate
rectum
retina
salivary gland
seminal vesicle
skeletal muscle
skin
skin 1
skin 2
small intestine
smooth muscle
soft tissue 1
soft tissue 2
sole of foot
spleen
stomach 1
stomach 2
substantia nigra
testis
thymus
thyroid gland
tonsil
urinary bladder
vagina
Number of proteins