Review

CRISPR-Enabled Functional Genomics in hPSC-Derived Neural Models for Autism Spectrum Disorder

S.T. Gopukumar , Madhumita Saha , Tanveen Kaur Soni , Samer Shamshad , Sunil Tadakod , Alok Sharma , Uddalak Das *

Posted Date: 26 November 2025

doi: 10.20944/preprints202511.1997.v1

Keywords: CRISPR-Perturbomics; hPSC-derived neural models; Autism Spectrum Disorder (ASD); functional genomics; single-cell multiomics

 doi:10.20944/preprints202511.1997.v1

Review

CRISPR-Enabled Functional Genomics in hPSC-Derived Neural Models for Autism Spectrum Disorder

S.T. Gopukumar [1], Madhumita Saha [2], Tanveen Kaur Soni [3], Samer Shamshad [4], Sunil Tadakod [5], Alok Sharma [6,7] and Uddalak Das [8,9,*]

[1] Nanobioinformatics Unit, Department of General Surgery, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai, Tamil Nadu, India 602105
[2] Institute of Nano science and Technology, Mohali, Knowledge City, Sector 81, Mohali, Punjab, India 140306
[3] Department of Biosciences, JIS University, Kolkata, West Bengal, India 700109
[4] ICAR- National Institute of Veterinary Epidemiology and Disease Informatics, Bengaluru, India 560119
[5] Reva University, School of Applied Science, Department of Biotechnology, Rukmini Knowledge Park, Kattigenahalli Yelahanka, Bengaluru 560064
[6] NeuroGen Brain & Spine Institute Nerul, Navi Mumbai, India, 400706
[7] KLS Wellness Institute, Sanpada, Navi Mumbai, India, 400705
[8] Developmental Biology and Genetics (DBG), Division of Biological Sciences, Indian Institute of Science (IISc), Bengaluru, India, 560012
[9] School of Biotechnology (SBT), Jawaharlal Nehru University (JNU), New Delhi, India, 110067
* Correspondence: uddalakdas@iisc.ac.in; uddalak_2025@jnu.ac.in

Abstract

Autism Spectrum Disorder (ASD) is a genetically heterogeneous neurodevelopmental condition driven by rare de novo variants, copy number variations, and polygenic risk. SFARI-curated genes show high mutational constraint and enriched expression in cortical neurons and glia. This review highlights recent advances in CRISPR-based functional genomics using human pluripotent stem cells and induced pluripotent stem cells differentiated into neural progenitors, excitatory and inhibitory neurons, astrocytes, microglia, and brain organoids. CRISPR modalities including knockouts, CRISPRi and CRISPRa, base and prime editing, and Cas13 enable pooled and arrayed screens with high coverage at low multiplicity of infection. Integration of multimodal readouts such as Perturb-seq, single-cell and spatial transcriptomics, proximity labeling proteomics, and functional assays including microelectrode arrays and calcium imaging provides system-level insights into ASD gene function. Computational frameworks like MIMOSCA and SCEPTRE facilitate network reconstruction and pseudo-time inference. Case studies reveal Wnt and BAF complex dysregulation, microglial pruning deficits, and non-cell autonomous effects. Translational approaches target haplo-insufficient genes such as CHD8 and SCN2A using AAV or antisense oligonucleotides supported by isogenic iPSC models. Remaining challenges include model immaturity and scalability, while future directions focus on spatial perturb-omics, AI-driven causal inference, and standardized biobanks for precision ASD therapeutics.

Keywords: CRISPR-Perturbomics; hPSC-derived neural models; Autism Spectrum Disorder (ASD); functional genomics; single-cell multiomics

Introduction

Autism Spectrum Disorder (ASD) is a complex and heterogeneous neurodevelopmental condition characterized by persistent deficits in social communication and restrictive, repetitive behaviors [1]. Clinical presentation varies widely, ranging from individuals with mild impairments who achieve independence to those requiring lifelong support due to profound cognitive and

 doi:10.20944/preprints202511.1997.v1

adaptive deficits. ASD affects roughly 1–2% of the global population, with recent estimates indicating a prevalence of 1 in 31 children aged 8 years in the U.S. [2]. This increase reflects improved diagnostic criteria, greater awareness, and potentially environmental contributions, although diagnostic expansion accounts for a substantial proportion of the observed trend. In the U.S., prevalence is estimated at 1 in 31 among 8-year-olds (2022 cohort)[3];. ASD is diagnosed approximately four times more often in males than females, highlighting sex-specific vulnerability. The disorder is profoundly heterogeneous, encompassing a wide spectrum of cognitive abilities, language skills, and neurological comorbidities, including intellectual disability (~30–50% of cases), epilepsy (~20–30%), and anxiety disorders [4,5]. Heterogeneity extends to brain cell types and developmental stages, with excitatory neurons, interneurons, and glia affected in distinct ways. The term "the autisms" is increasingly used to capture this phenotypic diversity. Lifelong support is often required for daily living, and the cumulative personal and economic costs are substantial: lifetime costs for an individual with ASD can exceed $2.4 million in the U.S. [6], driven by healthcare, education, and lost productivity. Annual costs per child may exceed $17,000, with a disproportionate burden on schools and families. Worldwide, ASD prevalence is estimated at ~1 in 100 individuals [7].

ASD exhibits high heritability, with estimates ranging from 50% to 90% based on twin studies, showing 60–90% concordance in monozygotic twins versus 0–30% in dizygotic pairs [8]. Hundreds of ASD-associated genes and variants have been identified, yet traditional genetic approaches alone are insufficient to explain the full spectrum of disease. Many identified variants are classified as Variants of Unknown Significance (VUS), representing up to 80% of exome sequencing findings. The functional relevance of these variants cannot be inferred reliably without experimental assays [9]. The genetic landscape of ASD is further complicated by pleiotropy, wherein single genes, such as CHD2, CHD8, or SHANK3, influence multiple biological pathways and phenotypes. Pleiotropic effects extend beyond neurodevelopment to non-brain-limited mechanisms, including tubulin regulation, cell cycle progression, DNA damage, and sex-specific traits. This complexity interacts with phenomena such as the "female protective effect," where females may require a higher genetic load to manifest ASD [10]. Cell-type specificity adds another layer of complexity: ASD genes show preferential expression in specific neuronal and glial populations at particular developmental windows [11]. For instance, upper-layer cortical excitatory neurons are implicated in social deficits, whereas oligodendrocyte anomalies contribute to myelination defects. Transcriptomic and single-cell studies underscore that ASD pathology manifests in a cell-context-dependent manner, complicating the translation of genotype to phenotype. Constraint metrics, such as pLI >0.9 from population datasets (gnomAD v4), highlight loss-of-function intolerance in ~20% of ASD genes, yet these resources do not account for dynamic, context-dependent effects [12]. The genetic architecture comprises rare de novo variants, inherited risk alleles, copy number variants (CNVs), and polygenic contributions, with loci such as 16p11.2 showing variable penetrance; approximately 20–30% ASD diagnosis among carriers, encompassing a wide NDD spectrum [13], further emphasizing heterogeneity. Environmental modifiers, including maternal immune activation, may epigenetically influence gene expression, creating an interplay between genetics and environment [14].

Given these complexities—VUS, pleiotropy, genetic heterogeneity, and cell-type specificity—systematic functional genomics is essential to map causal relationships between genes, cell types, and phenotypes [15]. High-throughput experimental frameworks enable controlled perturbation of candidate genes in relevant cell systems, allowing scalable evaluation of gene function. Curated gene databases, such as the Simons Foundation Autism Research Initiative (SFARI) Gene resource, categorize approximately 1,200 ASD-associated genes across evidence tiers, from syndromic (Category S) to suggestive (Category 3) [16]. These curated lists provide a rational basis for prioritizing genes in functional screens. Functional assays must account for diverse mechanisms: some genes are haploinsufficient, others are influenced by regulatory elements or CNVs. Genetic association alone is insufficient; experimental validation through cellular models is required to interpret the pathogenicity of ASD variants.

 doi:10.20944/preprints202511.1997.v1

Recent technological advances in CRISPR gene-editing and human pluripotent stem cell (hPSC) platforms provide an unprecedented opportunity to trace the causal chain from variant to mechanism [17]. hPSCs, either embryonic (hESCs) or induced (iPSCs) from patients, can be differentiated into relevant neural cell types, including neurons, astrocytes, microglia, and oligodendrocytes. CRISPR tools—nucleases, base editors, prime editors, CRISPRi/a—allow precise perturbation of genes or individual mutations within their endogenous genomic context. By generating libraries of hPSC-derived cells with targeted knockouts or specific edits across multiple ASD genes, researchers can systematically assay cellular phenotypes, including transcriptomic, epigenomic, proteomic, and functional endpoints. For example, CRISPRi knockdown of high-confidence ASD genes in cortical neurons has revealed downregulation of synapse-related transcripts and reductions in excitatory synapses, implicating roles in synaptogenesis [18]. Organoid systems replicate aspects of early corticogenesis, including progenitor proliferation, neuronal migration, and laminar organization, offering human-specific models that overcome limitations of animal systems. Pooled perturbation screens, such as Perturb-seq, link genetic edits to single-cell transcriptomes, enabling network-level analysis. Anchored examples of pooled organoid screening, such as CHOOSE [19], highlight the ability to interrogate genetic perturbations at scale. Vascularized organoids address maturation and necrosis limitations, while AI-based trajectory inference predicts variant effects across development. These platforms allow interrogation of high-risk genes (SFARI categories 1–3) and have revealed convergent mechanisms, such as defects in neuronal protein quality control, despite diverse upstream genetic perturbations. This convergence supports a shift from “gene hunting” to pathway discovery, offering more tractable therapeutic targets than individual gene interventions.

ASD imposes substantial societal and economic burdens. In 2021, an estimated 61.8 million individuals were autistic worldwide, making ASD a leading contributor to non-fatal health burdens among youth [20]. The lifetime cost of care ranges from $1.4 to $2.4 million per individual, with special education representing the largest contributor to non-healthcare expenditures [21]. Families face chronic stress, higher rates of mental health challenges, and limited support infrastructure [22]. Effective early interventions, such as applied behavior analysis, speech therapy, and occupational therapy, are resource-intensive yet variably efficacious, emphasizing the need for mechanistic insights to guide personalized therapeutic strategies.

ASD is therefore a genetically and phenotypically heterogeneous condition with profound societal, economic, and personal consequences. Variants of unknown significance, pleiotropy, cell-type specificity, and environmental modifiers challenge the translation of genetic findings into mechanistic understanding. Systematic functional genomics using CRISPR and hPSC platforms bridges this gap, enabling causal mapping from gene to cell type to phenotype to mechanism [23]. These approaches facilitate the identification of convergent pathways, offering a rational framework for targeted interventions and personalized medicine, while also informing early diagnosis, support strategies, and resource allocation.

2. Autism Genetics Primer

2.1. Rare De Novo Variants, Inherited Risk, CNVs, and Common Polygenic Risk

Recurrent CNVs such as 16p11.2 deletions/duplications and 22q11.2 deletions are well-established ASD risk factors. For example, the 16p11.2 CNV, affecting 29 genes, is associated with head size anomalies and occurs in ~1–2% of cases [23]. Importantly, 16p11.2 demonstrates variable penetrance; approximately 20–30% of carriers receive an ASD diagnosis, encompassing a wide neurodevelopmental disorder (NDD) spectrum [23]. De novo variants often contribute disproportionately to severe ASD cases, particularly those with intellectual disability, due to strong purifying selection.

Inherited rare variants, including private missense or protein-truncating mutations, contribute particularly in multiplex families or consanguineous populations (~10% of cases) [24]. These variants generally exhibit lower penetrance compared to de novo mutations, yet their additive effects can influence risk in the context of other genetic factors.

 doi:10.20944/preprints202511.1997.v1

Common variants, predominantly single nucleotide polymorphisms (SNPs), individually have small effects but collectively explain a substantial portion of ASD heritability (~20%) [25]. Polygenic risk scores (PRS) indicate that individuals in the top decile of PRS have an ~3-fold increase in ASD risk. Notably, rare de novo and common polygenic variants are not mutually exclusive; the combination of a high-impact de novo mutation and elevated polygenic risk can modulate penetrance and phenotypic severity ("double-hit" model) [26].

ASD impacts approximately 1–2% worldwide, with a prevalence of 1 in 31 among 8-year-olds in the US (2022 cohort) [27]. Global prevalence is estimated at ~1 in 100 children [17]. As illustrated in Figure 1, the schematic overview depicts the functional genomics pipeline for ASD gene discovery, detailing the workflow from gene prioritization to mechanistic inference and potential therapeutic insights.

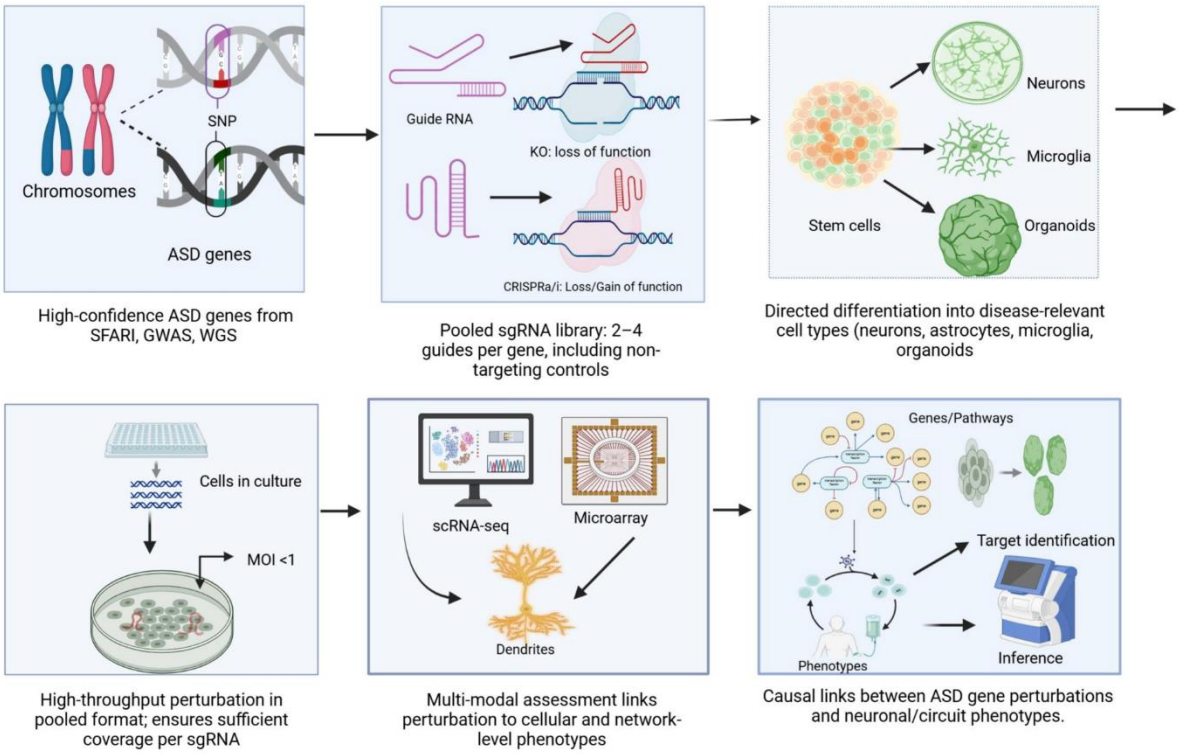

Figure 1. Functional genomics pipeline for ASD gene discovery. Schematic overview of a pooled CRISPR screening workflow in hPSC-derived neural models. Modules (from left to right) include gene prioritization (SFARI, WGS, GWAS), sgRNA library design (KO, CRISPRa/i), hPSC differentiation (neurons, glia, organoids), pooled perturbation screens, multimodal phenotyping (transcriptomics, electrophysiology, imaging), and mechanistic inference. Arrows indicate the flow from genetic perturbation to therapeutic insight.

2.2. High-Confidence vs Candidate ASD Genes

To systematically prioritize genes for functional studies, the Simons Foundation Autism Research Initiative (SFARI) maintains a gene scoring system based on the strength of evidence linking genes to ASD [28]. SFARI categorizes genes into four main groups as depicted in Table 1.

Table 1. SFARI gene categories and examples of ASD-associated genes.

SFARI Category	Description	Example Genes
Syndromic (S)	Monogenic syndromes featuring ASD (often with ID, seizures, etc.)	FMR1, TSC2, MECP2, PTEN

 doi:10.20944/preprints202511.1997.v1

1 (High Confidence)	Strong evidence (≥3 de novo LoF mutations)	CHD8, SCN2A, SYNGAP1, ADNP
2 (Strong Candidate)	Moderate evidence (e.g. ≥2 de novo LoF or replicated association)	GRIN2B, CTNND2, ANK2
3 (Suggestive)	Preliminary evidence (e.g. single de novo LoF or small studies)	CHRNA7, RAI1, DSCAM

Tiered functional genomics approaches use these categories: Tier 1 often includes Category S and 1 genes (with high constraint metrics, pLI >0.9), while Tiers 2–3 incorporate categories 2–3 genes with relevant developmental expression. A comprehensive list of SFARI autism risk genes is presented in Table 2, including their SFARI category, genetic constraint metrics (pLI score), developmental brain expression patterns, and associated phenotypes.

Table 2. Comprehensive Overview of SFARI Autism Risk Genes with Genetic Constraint, Developmental Expression, and Associated Phenotypes.

Gene	SFARI Category	pLI Score (gnomAD v4.0)	Constraint Metric (o/e)	Developmental Brain Expression Pattern	Associated Phenotypes
FMR1	Syndromic	0.99	0.10	Widely expressed in neurons; early development	Fragile X syndrome, intellectual disability, seizures
TSC2	Syndromic	0.99	0.08	Neuronal progenitors; cortical regions	Tuberous sclerosis complex, epilepsy, ID
MECP2	Syndromic	0.98	0.09	Methyl-CpG binding; glia and neurons	Rett syndrome, regression
PTEN	Syndromic	0.98	0.07	High in early cortical progenitors	Macrocephaly, ASD, tumor susceptibility
CHD8	1 (High Confidence)	1.00	0.05	High in prenatal cortex; excitatory neurons	Macrocephaly, intellectual disability
SHANK3	1	1.00	0.08	Synaptic regions; postnatal peaks	Phelan-McDermid syndrome, social deficits

 doi:10.20944/preprints202511.1997.v1

SCN2A	1	0.99	0.12	Neuronal ion channels; early development	Epilepsy, severe ASD
ARID1B	1	1.00	0.04	Chromatin remodeling; progenitors	Coffin-Siris syndrome, growth delays
SYNGAP1	1	1.00	0.06	Synaptic plasticity; excitatory neurons	Epilepsy, intellectual disability
ADNP	1	1.00	0.03	Chromatin; neural progenitors	Helsmoortel-Van der Aa syndrome
GRIN2B	2 (Strong Candidate)	0.99	0.11	Glutamate receptors; synapses	West syndrome, seizures
CTNND2	2	0.97	0.17	Neuronal adhesion; cortical layers	Intellectual disability, ASD
ANK2	2	0.98	0.13	Cytoskeleton; excitatory neurons	ASD, cardiac arrhythmias
FOXP1	2	0.98	0.15	Transcription; cortical layers	Language impairments, motor delays
TCF4	2	0.97	0.18	Neuronal differentiation; interneurons	Pitt-Hopkins syndrome, hypotonia
CHRNA7	3 (Suggestive)	0.96	0.20	Cholinergic neurons; hippocampus	Seizures, cognitive impairment
RAI1	3	0.95	0.22	Chromatin regulator; multiple brain regions	Smith-Magenis syndrome, behavioral issues
DSCAM	3	0.94	0.25	Axon guidance; cortical layers	Down syndrome-related phenotypes

 doi:10.20944/preprints202511.1997.v1

2.3. Functional Annotation and Population Resources

Prioritization of autism spectrum disorder (ASD) risk genes can be strengthened through the integration of population genetics data, gene constraint metrics, and developmental expression profiles. Population genetics resources such as gnomAD (v4.0, 2024) provide allele frequency information and constraint parameters including the probability of loss-of-function intolerance (pLI), loss-of-function observed/expected upper bound fraction (LOEUF), and observed/expected (o/e) ratios. Genes exhibiting high pLI scores or low o/e values (for example, CHD8 and SCN2A, with SCN2A displaying an o/e of 0.12) are generally considered haploinsufficient, indicating that disruptive mutations in these loci are likely to be deleterious [29]. Clinical variant repositories such as ClinVar further contribute pathogenicity annotations, which assist in filtering out common benign variants; however, coverage of ASD-associated genes within ClinVar remains limited, necessitating integration with functional evidence to ensure robust gene prioritization [30]. In parallel, developmental expression atlases such as BrainSpan and the Allen Brain Atlas enable the mapping of spatiotemporal gene expression patterns. Notably, ASD risk genes are enriched in the mid-gestational cortex, particularly within neural progenitors, deep-layer excitatory neurons, and specific interneuron populations. Single-cell transcriptomic studies corroborate this cell-type specificity, further implicating synaptic and chromatin regulatory networks in ASD pathogenesis [31].

2.4. Challenges in ASD Genetics

Autism spectrum disorder (ASD) research is challenged by several inherent complexities that complicate gene discovery and mechanistic understanding. A major factor is genetic heterogeneity, as hundreds to thousands of genes contribute to ASD risk, with each accounting for only a small proportion of cases. Consequently, traditional recurrence-based genetic approaches have limited statistical power when mutations are largely unique across individuals [32]. Variable penetrance further complicates interpretation, as not all carriers of pathogenic variants manifest ASD phenotypes; for instance, 22q11.2 deletions demonstrate approximately 50% penetrance. Such outcomes are strongly influenced by polygenic background and environmental modifiers, underscoring the context-dependent nature of phenotypic expression [33]. Pleiotropy also represents a key challenge, as single genes frequently contribute to multiple neurodevelopmental disorders. For example, mutations in SCN2A have been associated with ASD, epilepsy, and intellectual disability, while SHANK3 variants are implicated in both ASD and schizophrenia [34]. Moreover, cell-type specificity plays an important role, as ASD risk genes may exert distinct effects across neuronal and glial subpopulations, necessitating high-resolution models to delineate pathogenic mechanisms. Epistasis and broader genetic interactions further complicate analyses, as network-level effects can modulate the impact of individual variants, requiring integrative frameworks that combine genetic, transcriptomic, and functional datasets [35]. Finally, inference from polygenic additive models has been shown to amplify the contribution of rare variants and provide explanatory power for phenotypic heterogeneity. In this context, functional genomics combined with artificial intelligence (AI)-driven integration of polygenic risk scores (PRS) offers promise for early risk prediction and the development of tailored therapeutic interventions.

3. Human Stem Cell Platforms for ASD Functional Genomics

3.1. Human Pluripotent Stem Cell Types: hESCs vs iPSCs

Human pluripotent stem cells (hPSCs) comprise both human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs), each with unique characteristics that make them suitable for modeling neurodevelopmental disorders such as autism spectrum disorder (ASD) [36]. Table 3 compares the main features of hESCs and iPSCs. hESCs are derived from early embryos and exhibit “naïve” pluripotency, are genetically stable, and can be expanded robustly over long-term culture. They provide a uniform baseline and are ideal as reference lines, but their use is limited by ethical considerations, regulatory constraints, and restricted availability of lines with specific patient

 doi:10.20944/preprints202511.1997.v1

genotypes [36]. Furthermore, hESCs may not perfectly capture genetic diversity and are associated with potential immune mismatch if used for therapeutic applications.

Table 3. Comparative Features of hESCs and iPSCs.

Aspect	hESCs	iPSCs	Notes
Source	Surplus embryos	Patient fibroblasts / somatic cells	Defines origin
Genetic Background	Limited diversity; allogenic	Patient-specific; captures variants	Impacts disease modeling
Reprogramming Artifacts	None	Epigenetic memory; somatic mutations	Only in iPSCs
Ethical Concerns	Embryo destruction	Minimal; informed consent	Critical for approval
Expansion Potential	High; stable karyotype	High; clonal instability possible	Affects scale-up
Differentiation Efficiency	Consistent across lines	Variable; donor-dependent	Important for reproducibility
Disease Modeling	Requires genome editing (e.g., CRISPR)	Direct from patients; idiopathic conditions (e.g., ASD)	Relevance to patient-specific studies
Immune Compatibility	Allogenic mismatch	Autologous potential	Consider for therapy
Cost / Accessibility	Restricted access	Widely available; biobanks	Affects practical use

In contrast, iPSCs are reprogrammed from adult somatic cells, such as fibroblasts or blood, and retain the donor's genetic background, including disease-associated variants. This makes iPSCs particularly valuable for generating patient-specific models and isogenic controls, allowing direct investigation of the causal effects of genetic variants. iPSCs circumvent the ethical concerns associated with embryonic material and enable the study of personalized disease mechanisms. However, the reprogramming process can introduce genetic and epigenetic variability, including clonal differences, aneuploidy, and residual epigenetic memory, which may influence differentiation outcomes and introduce experimental noise. Despite these limitations, both hESCs and iPSCs are amenable to genome engineering, such as CRISPR–Cas9-mediated gene editing, and the choice between them depends largely on whether uniform reference lines or patient-specific insights are desired.

3.2. Differentiation Endpoints: Neural Lineages

A critical strength of hPSCs is their capacity to differentiate into diverse central nervous system lineages, enabling the study of ASD-relevant biology in the most pertinent cell types [37]. Neural progenitor cells (NPCs) represent multipotent early precursors expressing markers such as SOX2 and PAX6. NPCs can be expanded extensively, providing a scalable platform for high-throughput screening, but they represent an early developmental stage and are primarily suited for probing proliferation or cell fate decisions.

 doi:10.20944/preprints202511.1997.v1

Excitatory glutamatergic neurons, typically generated through dual-SMAD inhibition combined with WNT pathway modulation, express markers such as TBR1, VGLUT, and CTIP2, form synapses *in vitro*, and display spontaneous activity over several weeks. These neurons are highly relevant for ASD studies because many disease-associated genes are enriched in excitatory cortical neurons, although full maturation may take months and cultures often retain fetal-like characteristics.

Inhibitory interneurons, patterned with SHH agonists to mimic medial ganglionic eminence development, express GAD67, DLX2, and subtype-specific markers such as somatostatin and parvalbumin. These interneurons are essential for investigating excitation–inhibition balance, a key feature in ASD pathology, but achieving homogeneous and fully mature populations is technically challenging.

Astrocytes arise from extended NPC differentiation or directed protocols using BMP and FGF signaling, express GFAP and S100β, and contribute to synapse formation, neurotransmitter recycling, and non-cell-autonomous effects on neuronal function.

Microglia, derived from yolk sac-like myeloid progenitors using cytokines such as IL-34 and GM-CSF, express IBA1 and P2RY12 and play crucial roles in synaptic pruning and neuroinflammation, processes increasingly recognized in ASD [38]. Because microglia originate from a different embryonic lineage, they are generated separately and co-cultured with neurons or organoids to study neuro–immune interactions.

Oligodendrocytes, the myelinating glia expressing OLIG2 and MBP, require prolonged differentiation protocols and are implicated in neural wiring and connectivity relevant to ASD. Although their inclusion in CRISPR screens is limited due to slow differentiation, organoids can give rise to oligodendrocytes during extended culture periods. Collectively, these differentiated cell types enable multifaceted modeling of ASD, reflecting both cell-autonomous and non-cell-autonomous contributions to neural development.

3.3. Three-Dimensional Models

Three-dimensional brain models extend the utility of hPSC-derived cultures by recapitulating cellular architecture, multicellularity, and early neural circuit formation. Brain organoids are self-organizing structures that form layered arrangements of progenitors, neurons, astrocytes, and sometimes oligodendrocytes. They allow investigation of cortical development, disease modeling, and patient-specific phenotypes [39].

Regional patterning using morphogens permits the generation of organoids representing specific brain regions, such as the cortex, hippocampus, or midbrain, producing more homogeneous cell populations. Recent advances have led to the creation of assembloids, formed by fusing region-specific organoids, which facilitate modeling of inter-regional neuronal migration and long-range circuit connectivity, such as dorsal cortical interneuron migration from ventral forebrain regions. Multi-region brain organoids (MRBOs; defined here as fused assemblies of three or more patterned organoids) further integrate multiple brain regions and may include vascular or microglial components, enabling the study of non-cell-autonomous effects, circuit-level pathology, and interactions between distinct neural subtypes [40].

Vascularized organoids, achieved through co-culture with endothelial cells or *in vivo* engraftment, improve nutrient and oxygen delivery, enhance maturation, and can recapitulate rudimentary blood–brain barrier function. Despite these innovations, limitations persist: organoids are typically restricted to fetal-like development, lack full cortical folding, and exhibit batch-to-batch variability [41]. To mitigate these, best practice now includes documenting seed dimensions, media lot numbers, extracellular matrix (ECM) lot identifiers, and rotation velocity [19].

Recent engineering approaches, including spinning bioreactors and microfabricated scaffolds, have reduced variability, while single-cell genomics has mapped the diversity of progenitors, excitatory and inhibitory neurons, astrocytes, and other cell types. Patient-derived organoids with mutations in genes such as CHD8 or 16p11.2 demonstrate phenotypes including over-proliferation and synaptic deficits, illustrating their utility for studying ASD pathology [42].

 doi:10.20944/preprints202511.1997.v1

3.4. Co-Culture Systems and Microphysiological Platforms

To accurately model the brain's complex cellular environment, hPSC-derived neurons are frequently co-cultured with astrocytes, oligodendrocytes, or microglia. These systems recapitulate critical cell–cell interactions, including neuro–immune communication, synaptic pruning, and glial support for neuronal maturation [43].

Microfluidic platforms and microphysiological systems enhance these co-culture setups by providing compartmentalization, controlled connectivity via microchannels, and fluid perfusion that mimics shear stress and dynamic nutrient delivery. Blood–brain barrier chips incorporate endothelial cells, pericytes, and astrocytes in tubular flow chambers, modeling vascular interactions, while multi-organ chips and neurovascular organoids allow the study of neuroinflammation in a more physiologically relevant context [44].

Emerging approaches also highlight dual-barcode co-culture Perturb-seq methods, which distinguish between cell-autonomous and non-cell-autonomous influences, such as microglial pruning versus neuronal response [45]. Collectively, these advanced platforms provide a bridge between traditional 2D cultures and *in vivo* systems, facilitating mechanistic studies of multi-cellular interactions relevant to ASD [46].

3.5. Practical Considerations

Working with hPSC-derived neural models requires meticulous attention to batch effects, maturation stage, reproducibility, and scalability. Variability between iPSC clones, passages, or donors can influence differentiation efficiency and gene expression, and extended culture may introduce drift or altered phenotypes [47]. Neural maturation is critical, as hPSC-derived neurons remain developmentally immature relative to adult tissue, with synaptic density, ion channel expression, and network activity evolving over weeks to months.

Strategies to mitigate variability include using multiple clones, standardized protocols, cryopreservation of NPCs at defined stages, and the application of electrical stimulation to accelerate maturation [48]. Scaling experiments, particularly pooled CRISPR screens, demands large numbers of cells to maintain library representation; pooled designs require ≥500× coverage, and Perturb-seq requires ~100–200 cells/guide for robust power [49]. Transduction should occur at the NPC stage with MOI < 1 to ensure representation.

Additional considerations include randomizing guides across variations, adjusting with Harmony or MNN, and incorporating biological replicates [49]. Confirming karyotype and TP53 status is essential due to p53-dependent toxicity induced by DSBs; ribonucleoprotein (RNP) delivery with brief exposure, or the use of base/prime editors, reduces genomic instability [50].

Employing robust experimental designs with isogenic validation strategies (e.g., knockout plus cDNA rescue or CRISPRa) ensures causality [49]. Automated platforms such as liquid handlers and bioreactors facilitate high-throughput production.

Hybrid approaches are now being explored: for example, spatially resolved CRISPR screens in organoids to map genotype–phenotype interactions with respect to tissue structure [51]. Likewise, assay triage strategies can streamline workflows, beginning with pooled (viability/FACS), followed by multi-electrode arrays (MEA) for network outputs, and culminating in organoid-based or spatial platforms when structure is central.

Collectively, iPSC-derived MRBOs represent a powerful platform for capturing ASD heterogeneity, revealing non-cell-autonomous phenomena such as microglial pruning deficits, while limitations in cellular maturity suggest that hybrid approaches involving *in vivo* engraftment may be necessary for full validation [52]. A summary of the different human stem cell models, culture systems, and brain organoid platforms, along with their key features, cellular complexity, functional readouts, applications, and limitations, is provided in Table 4.

 doi:10.20944/preprints202511.1997.v1

Table 4. Overview of Human Stem Cell Models, Culture Systems, and Brain Organoid Platforms.

Model / System	Key Features / Cell Types	Cellular Complexity	Structural Fidelity	Functional Readouts	Applications / Pros	Limitations / Cons
hESCs (Embryonic Stem Cells)	Naïve pluripotent stem cells from embryos	Single-cell pluripotent	Minimal	Differentiation potential; lineage tracing	Highly pluripotent, uniform; stable karyotype; high expansion potential	Limited lines; ethical concerns (embryo destruction); immune mismatch
iPSCs (Induced Pluripotent Stem Cells)	Reprogrammed adult somatic cells; patient genotype	Single-cell pluripotent	Minimal	Differentiation potential; disease modeling	Patient-specific; scalable; autologous potential; widely available from biobanks	Variability from reprogramming; epigenetic memory; donor-dependent differentiation efficiency
Neural Progenitor Cells (NPCs)	Expandable neural precursors (SOX2+, PAX6+)	Early neural lineage	Minimal	Proliferation, differentiation assays	Easily amplified; suitable for high-throughput screening	Lacks mature neuronal/glial functions
Cortical Excitatory Neurons	Glutamatergic neurons (TBR1+, VGLUT+)	Post-mitotic neurons	Minimal to 2D networks	Electrophysiology, synapse assays	Model cortical synapses; relevant to ASD	Require weeks to mature; fetal-like features
GABAergic Inhibitory Neurons	Interneurons (GAD67+,	Post-mitotic neurons	Minimal to 2D networks	Electrophysiology, migration assays	Study E/I balance and migration	Complex induction (SHH patterning)

 doi:10.20944/preprints202511.1997.v1

	PVALB/SST)					; long maturation
Astrocytes	Glial cells (GFAP+, S100β+)	Glia	Minimal	Support synaptogenesis, calcium signaling	Support neurons and synapse formation; study cell-cell interactions	Require prolonged culture
Microglia	Brain macrophages (IBA1+, TMEM119+)	Myeloid-lineage glia	Minimal	Cytokine release, migration, synaptic pruning	Model neuroinflammation; study microglia-neuron interactions	Separate differentiation needed; no endogenous myeloid cells unless co-cultured
Oligodendrocytes	Myelinating glia (MBP+, OLIG2+)	Glia	Minimal	Myelination assays	Study myelination dynamics	Differentiation is very slow (months)
2D Monolayer Culture	Neurons, astrocytes, glia	Low	Minimal	Single-cell imaging, electrophysiology, bulk/scRNA-seq	Cost-effective, scalable; suitable for high-throughput screens	Lacks tissue architecture; limited to cell-autonomous effects
3D Brain Organoids	Self-organizing 3D tissues with neurons and glia	Multiple cell types	Recapitulates early development and cytoarchitecture	Gene expression, morphology, developmental trajectory	Models *in vivo* complexity; bridges 2D culture and *in vivo* brain	Diffusion limits; lack vascularization; limited long-term maturation
Assembloids / MRBO	Fused region-specific organoids	Multiple cell types, multipl	Models inter-regional	Electrophysiology, connectivity,	Model non-cell-autonomous effects and	High technical difficulty; costly;

 doi:10.20944/preprints202511.1997.v1

	(e.g., cortex + subpallium)	e regions	connectivity	gene expression	circuit pathology	limited long-term stability
Vascularized Organoids	Organoids plus endothelial cells or implanted vasculature	Multiple cell types	Enhanced structural fidelity with vasculature	Long-term viability, nutrient diffusion	Improved nutrient supply; prolonged longevity	Difficult to reproduce uniform vasculature
Microfluidic Co-culture	2–3 cell types (e.g., neurons + microglia)	Low to moderate	N/A	Microglial migration, inflammatory signaling, cytokine release	Enables controlled study of neuro-immune crosstalk	Does not model 3D architecture; limited to specific interactions

4. The CRISPR Toolbox: Modalities and Deliverables

4.1. CRISPR-Cas Modalities and Their Application in ASD

The CRISPR-Cas system has evolved into a versatile toolkit capable of interrogating the genetic and epigenetic complexity of disorders like autism spectrum disorder (ASD). Each modality offers distinct strengths for modeling loss-of-function, hypomorphic, gain-of-function, or precise variant effects. Table 5 provides an overview of CRISPR-Cas modalities employed for modeling genetic and epigenetic perturbations in Autism Spectrum Disorder (ASD), detailing their mechanisms, types of perturbations, library design strategies, relevant ASD use cases, and the respective advantages and limitations of each approach. Figure 2 provides a comparative overview of CRISPR-based approaches—including CRISPR-KO, CRISPRi, CRISPRa, and base/prime editing—highlighting their mechanisms, gene-level effects, and example applications in ASD research.

Table 5. CRISPR-Cas modalities for modeling genetic and epigenetic perturbations in Autism Spectrum Disorder (ASD).

Modality	Mechanism	Perturbation Type	Library Design	ASD Use Cases	Advantages	Disadvantages
CRISPR-KO (Nuclease-based editing)	Cas9 nuclease induces DSBs repaired by NHEJ,	Complete loss-of-function	sgRNA libraries targeting coding exons	Modeling high-penetrance LGD variants (e.g., *CHD8*)	Robust LoF phenotypes, simple	Potential lethality for essential genes; p53/DSB

 doi:10.20944/preprints202511.1997.v1

	introducing indels				design, scalable	toxicity; mosaicism
CRISPRi (Interference)	dCas9-KRAB represses transcription	Tunable hypomorph (partial knockdown)	Promoter-proximal sgRNA libraries	Modeling dosage-sensitive or toxic LoF (e.g., *SCN2A* variants)	Non-lethal, reversible, dose-dependent effects	Variable repression efficiency; off-target chromatin effects
CRISPRa (Activation)	dCas9 fused to activators (VP64-p65-Rta) upregulates transcription	Gain-of-function / compensation	Promoter-proximal sgRNA libraries	Rescue of haploinsufficient genes; modeling GoF variants	Endogenous regulation, rescue strategies feasible	Limited to genes with accessible promoters; variable activation strength
Base Editors (CBE/ABE)	Cas9 nickase + deaminase mediates targeted base conversion (C→T or A→G)	Single-nucleotide substitution	sgRNA libraries with editable PAM-proximal sites	Missense or nonsense variants (e.g., *SHANK3*)	Precise, efficient point mutations; avoids DSBs	Window and PAM constraints; bystander edits
Prime Editors	Cas9 nickase + RT + pegRNA enables programmable substitutions, small indels	Any base substitution, small indel	pegRNA libraries with optimized design (nick-to-edit distance, PBS length)	Patient-specific VUS modeling; fine-tuned variant editing	Broad editing scope; reduced mosaicism	Lower efficiency; complex design; delivery challenges

 doi:10.20944/preprints202511.1997.v1

Epigenetic Editors	dCas9 fused to epigenetic modifiers (e.g., DNMT3A, LSD1, p300)	Reversible transcriptional silencing/activation	sgRNA libraries targeting promoters, enhancers, non-coding elements	Modeling imprinting defects, non-coding variants	Reversible, non-genome-disruptive; context-specific modulation	Epigenetic changes may be unstable; off-target chromatin remodeling
RNA Editors (Cas13-based)	Cas13 targets RNA; with ADAR fusion mediates A→I editing	Transcript knockdown or RNA editing	sgRNA libraries targeting mRNA/transcripts	Transient knockdown or reversible editing in ASD genes	No DNA modification; reversible; RNA-level precision	Transient effects; off-target RNA cleavage; delivery limits

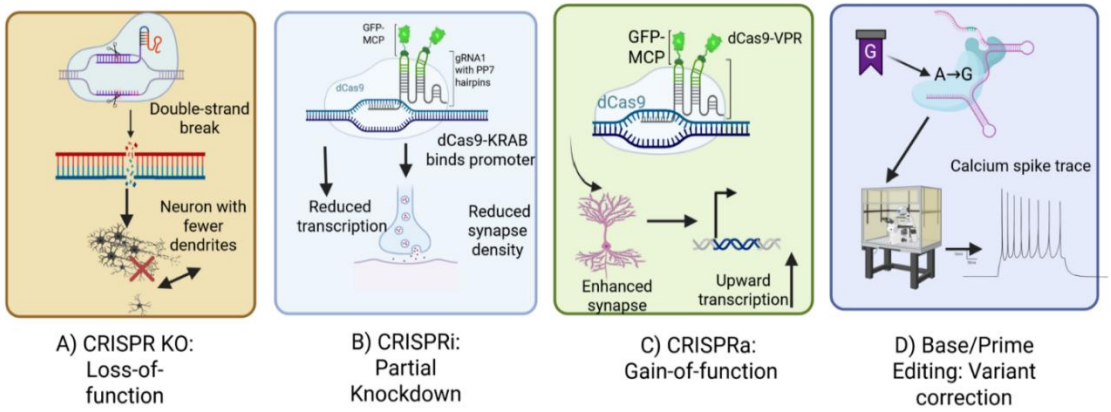

Figure 2. Comparative overview of CRISPR modalities for ASD functional genomics. Comparison of CRISPR-KO, CRISPRi, CRISPRa, and base/prime editing platforms in neuronal models. Each panel depicts the mechanism of action, gene-level perturbation, and example ASD-relevant applications. The central table summarizes key attributes, including perturbation type, library format, typical use cases, and modality-specific advantages and limitations.

4.1.1. Nuclease-Based Editing (CRISPR-KO)

The foundational approach, CRISPR knockout (CRISPR-KO), employs a Cas9 nuclease guided by a single-guide RNA (sgRNA) to induce double-strand breaks (DSBs) [53]. The cell's error-prone non-homologous end joining (NHEJ) repair pathway typically introduces small insertions or deletions (indels), resulting in frameshift mutations and complete loss-of-function alleles. CRISPR-KO is particularly suited for modeling high-penetrance, likely-gene-disrupting (LGD) variants identified in ASD patients (e.g., CHD8). Complete knockout of essential genes, however, can be lethal [54]. To minimize p53/DSB-related toxicity, ribonucleoprotein (RNP) delivery and brief Cas9 exposure are recommended, alongside confirmation of karyotype integrity and TP53 status post-editing [54].

 doi:10.20944/preprints202511.1997.v1

4.1.2. CRISPR Interference and Activation (CRISPRi/a)

To enable tunable modulation of gene expression without introducing genomic double-strand breaks, catalytically inactive Cas9 (dCas9) can be fused to transcriptional effector domains. In CRISPR interference (CRISPRi), dCas9 is coupled with the KRAB repressor domain to suppress transcription, thereby generating a hypomorphic state. This approach is particularly valuable for modeling partial dosage effects, which is critical for genes where complete knockout would be deleterious, such as SCN2A gain-of-function variants [55]. Conversely, CRISPR activation (CRISPRa) employs dCas9 fused to transcriptional activator domains (e.g., VP64-p65-Rta) to enhance expression from endogenous promoters, thereby facilitating gain-of-function analyses or rescuing haploinsufficient gene states. These complementary strategies are essential for modeling subtle regulatory perturbations and dosage-sensitive mechanisms, both of which are frequently implicated in ASD. For rigorous isogenic validation, edit-and-rescue strategies are recommended, such as combining loss-of-function mutations with cDNA-based rescue or employing CRISPRa-mediated compensation [56].

4.1.3. Base and Prime Editors

To enable precise genome modifications without the induction of double-strand breaks (DSBs), advanced editing platforms such as base editors and prime editors have been developed. Base editors, including cytosine base editors (CBE) and adenine base editors (ABE), function by combining a Cas9 nickase with a cytidine or adenosine deaminase, thereby facilitating the direct conversion of C→T (CBE) or A→G (ABE). These systems allow for the modeling of pathogenic missense variants or premature stop codons in genes such as *SHANK3* [57]. Importantly, editing efficiency is influenced by the sequence context and the preferred activity window of the base editor, factors that must be carefully considered during experimental design [57]. Prime editors expand the scope of genome modification by coupling a Cas9 nickase with a reverse transcriptase and employing a prime editing guide RNA (pegRNA). This system enables the introduction of all 12 possible base substitutions, as well as small insertions and deletions, thereby allowing precise modeling of patient-specific variants of uncertain significance (VUS) with reduced mosaicism, particularly in stem cell systems [57]. The efficiency of prime editing is critically dependent on pegRNA design parameters, such as the nick-to-edit distance, which must be optimized for successful editing outcomes [57].

4.1.4. Epigenetic and RNA Editors

Epigenetic editors, generated by fusing dCas9 with chromatin-modifying enzymes such as DNMT3A, LSD1, or p300, enable reversible modulation of gene activity through targeted DNA methylation or histone modifications. This approach is particularly advantageous for investigating non-coding variants or correcting imprinting defects [58]. In addition, RNA-targeting strategies utilizing Cas13 nucleases allow transient knockdown of transcripts or, when coupled with an adenosine deaminase, facilitate RNA base editing (A→I). Such methods provide reversible perturbation at the RNA level without introducing permanent alterations to the genome [59].

4.2. Guide RNA Design

Effective CRISPR-based experiments require the design and utilization of high-quality single-guide RNAs (sgRNAs) to ensure maximal on-target activity while minimizing off-target effects. On-target efficiency is typically predicted using computational algorithms, such as the Doench scoring system, which facilitates the selection of sgRNAs with strong cleavage potential. Minimization of off-target effects involves careful evaluation of potential unintended genomic binding sites, with the exclusion of sgRNAs predicted to exhibit a high probability of off-target interactions. In addition, protospacer adjacent motif (PAM) specificity must be considered, as the choice of Cas variant (e.g., SpCas9 requiring an NGG PAM sequence) determines the spectrum of compatible target sites. Furthermore, it is critical to avoid designing sgRNAs that overlap single nucleotide polymorphisms (SNPs), since such polymorphic variations may reduce binding efficiency, particularly in human

 doi:10.20944/preprints202511.1997.v1

pluripotent stem cell (hPSC) and induced pluripotent stem cell (iPSC) lines derived from genetically diverse backgrounds [60].

4.2.1. Tiling and Saturation Mutagenesis

Tiling and saturation mutagenesis represent powerful strategies to enhance the resolution and interpretability of functional genomic screens in autism spectrum disorder (ASD). In tiling approaches, single-guide RNAs (sgRNAs) are designed to systematically span coding or regulatory regions, thereby enabling the identification and mapping of critical functional domains within target genes. Complementarily, saturation mutagenesis employs base or prime editing technologies to generate all possible point mutations within a given locus, facilitating comprehensive interpretation of variant effects at single-nucleotide resolution. Together, these approaches maximize the sensitivity of screening platforms and enable functional annotation of both coding and non-coding variants that are implicated in ASD.

4.3. CRISPR Library Design

4.3.1. Genome-Wide vs. Focused Libraries

Genome-wide CRISPR libraries encompass nearly all genes, typically incorporating more than 100,000 single guide RNAs (sgRNAs), and thereby enable the unbiased discovery of novel genes influencing phenotypic outcomes. In contrast, focused libraries are designed to target pre-selected gene sets, such as ASD Tier 1 and Tier 2 genes, synaptic regulators (e.g., SYNGAP1), or chromatin remodelers. These focused approaches allow for deeper coverage per gene, reduce experimental complexity, and enhance statistical power in downstream analyses.

4.3.2. Library Formats

CRISPR screening approaches can be broadly categorized into pooled and arrayed library formats. In pooled library screens, all single guide RNAs (sgRNAs) are combined and delivered collectively into a population of cells, with phenotypic outcomes assessed through sequencing-based readouts, either at bulk or single-cell resolution. This format is highly scalable, cost-effective, and well-suited for relatively simple phenotypes such as cell viability, fluorescence-activated cell sorting (FACS), or single-cell RNA sequencing (scRNA-seq). To ensure sufficient statistical power and reproducibility, pooled screens typically require a coverage of at least 500-fold representation, with a minimum of 100–200 cells per sgRNA recommended for Perturb-seq experiments [60]. In contrast, arrayed library screens involve delivering individual perturbations into separate wells of multiwell plates (e.g., 96- or 384-well formats), enabling detailed phenotypic assessments such as high-content imaging or electrophysiological recordings. Although arrayed screens allow for more complex phenotyping, they are inherently lower in throughput compared to pooled designs.

4.4. Barcode Strategies and Readout Integration

Reliable assignment of phenotypes to specific perturbations is essential in pooled screening approaches. In this context, several strategies have been developed to link perturbations with transcriptional outcomes at single-cell resolution. For instance, CROP-seq embeds the sgRNA sequence within a polyadenylated transcript, enabling its capture alongside endogenous mRNAs during single-cell RNA sequencing (scRNA-seq). Alternatively, direct gRNA capture employs custom reverse transcription primers complementary to the sgRNA scaffold, thereby facilitating precise sequencing of guide identities. Another widely used approach involves barcode–sgRNA vector systems, which link unique molecular barcodes to sgRNAs to allow deconvolution of perturbations. Collectively, these strategies provide high-fidelity mapping of perturbations to cellular phenotypes, a requirement for high-content functional genomics studies in ASD. To ensure transparency and reproducibility, deposition of raw and processed data in repositories such as GEO or SRA, together with the submission of analysis notebooks (e.g., using Scanpy or Seurat) and guide–cell assignment maps, is strongly recommended [60].

 doi:10.20944/preprints202511.1997.v1

4.5. Advantages for ASD Research

The CRISPR-Cas toolbox offers a versatile set of approaches to model diverse genetic perturbations relevant to autism spectrum disorder (ASD). CRISPR-mediated knockout (CRISPR-KO) enables the generation of complete loss-of-function (LoF) models, mimicking de novo truncating mutations. CRISPR interference (CRISPRi) and base editing facilitate the modeling of hypomorphic effects or partial LoF, which are particularly informative for haploinsufficient genes. Conversely, CRISPR activation (CRISPRa) provides a means to rescue insufficient gene expression. Prime editing allows for precise modeling of patient-specific variants while reducing mosaicism in induced pluripotent stem cell (iPSC) systems. In addition, epigenetic editors and RNA-targeting Cas13 systems enable interrogation of regulatory variants and transient perturbations. Collectively, these methodologies provide an integrated platform for systematic interrogation of ASD genetics across genomic, epigenomic, and transcriptomic levels [61].

5. Building CRISPR-Engineered Stem Cell Libraries (Practical Guide)

5.1. Gene/Variant Selection Strategy

The first critical step in designing a functional genomics screen is defining a gene and variant selection strategy. A curated framework such as the SFARI Gene database provides tiered categories that prioritize genes for inclusion [62]. A multi-tier library can be constructed: Tier 1 includes Category S and 1 (high-confidence) genes (~100–150) along with control guides, while Tier 2 adds Category 2 genes and strong candidates, totaling several hundred more. This allows phased screening and focuses on convergent pathways in ASD.

Selection is further refined using computational and population-based metrics. Constraint scores, such as pLI > 0.9 or low LOEUF (as of gnomAD v4), identify genes intolerant to loss-of-function variants. Genes frequently mutated in ASD cohorts or encoding synaptic proteins or chromatin remodelers are prioritized [63]. Clinically reported Variants of Unknown Significance (VUS) can also be included for patient-specific studies, enabling functional reclassification. For variant-specific studies, recurrent missense mutations (e.g., in SCN2A or CHD8) are targeted using base or prime editors to recapitulate pathogenic alleles [64].

In summary, gene/variant selection integrates curated knowledge (SFARI, literature), population-genetic constraint, and experimental focus (e.g., synaptic vs transcription factor genes), enabling targeted and interpretable screening.

5.2. Library Architecture

CRISPR library architecture is strategically designed to maximize on-target activity, redundancy, and statistical confidence in screening outcomes. A standard approach involves incorporating 4–10 gRNAs per gene to ensure robust knockout or knockdown efficiency while simultaneously mitigating guide-specific off-target effects. The inclusion of redundant guides enhances confidence in gene-level hits, as true hits are expected to demonstrate consistent dropout or enrichment across multiple guides. Appropriate controls are critical for the validity of these experiments. Negative controls, typically comprising non-targeting or scrambled guides (~5–10% of the library), establish baseline phenotypes, whereas positive controls, such as guides targeting essential genes (e.g., POLR2A) or reporter genes, enable the monitoring of screen performance. More advanced library designs include dual-gRNA constructs, which facilitate exon or regulatory region deletions and enable saturation mutagenesis through systematic tiling of guides across specific loci. These constructs can be engineered with unique molecular handles for PCR amplification or cloning, thereby streamlining quality control and downstream analyses [65]. In the context of autism spectrum disorder (ASD), focused libraries are generally composed of 200–500 genes, providing a balance between scale and depth, and are particularly useful in uncovering convergent molecular networks such as Wnt signaling in neural progenitors.

5.3. Cloning, Synthesis, and Quality Control

 doi:10.20944/preprints202511.1997.v1

CRISPR libraries are generated as oligonucleotide pools, either synthesized commercially via microarray or prepared in arrayed format followed by pooling after PCR amplification. These oligonucleotides are subsequently cloned into lentiviral vectors employing techniques such as Gibson assembly or Golden Gate cloning. Transformation into bacterial hosts necessitates high coverage, typically ≥1000 colonies per single-guide RNA (sgRNA), to prevent bottleneck effects. Quality control (QC) of the library is performed using Next-Generation Sequencing (NGS) to ensure uniform guide representation, with read counts maintained within approximately 2–3 fold of the mean, minimal dropout of guides, and retention of representation following viral packaging [66]. Such rigorous QC measures are essential to guarantee that the outcomes of CRISPR screens accurately reflect biological effects rather than artifacts introduced by library preparation.

5.4. Delivery to hPSCs and Derivatives

The selection of an appropriate delivery method for CRISPR-based experiments is dictated by the specific experimental objectives. Lentiviral vectors are the standard choice for pooled screens, as they stably integrate single-guide RNAs (sgRNAs) into both dividing and non-dividing cells. Maintaining a low multiplicity of infection (MOI 0.3–1) ensures that each cell receives a single guide; however, lentiviral delivery carries risks of insertional mutagenesis and requires biosafety level 2 (BSL-2) containment [67]. Adeno-associated virus (AAV) offers a non-integrating alternative with low immunogenicity, making it suitable for post-mitotic neurons, though its small packaging capacity (~4.7 kb) and limited transduction efficiency in dividing human pluripotent stem cells (hPSCs) constrain its utility [68]. Ribonucleoprotein (RNP) electroporation allows direct delivery of Cas9 protein pre-loaded with sgRNA, providing transient but highly efficient genome editing (~80–90%) with minimal off-target effects, making it ideal for individual edits or small pooled screens [69]. The PiggyBac transposon system enables non-viral integration of large genetic cargos, facilitating stable expression of Cas9 or CRISPR interference (CRISPRi) machinery; however, random genomic integration may perturb endogenous genes. Additionally, the use of stable Cas9 or dCas9-KRAB hPSC lines simplifies library delivery by requiring only the sgRNA pool, ensuring uniform editing across cells. The typical workflow for pooled CRISPR knockout (CRISPR-KO) or CRISPRi screens in hPSC-derived neurons involves generating a stable Cas9/dCas9-KRAB line, producing a lentiviral sgRNA library, infecting cells at low MOI, differentiating the cells, and subsequently assaying relevant phenotypes.

5.5. Generation of Isogenic Clones for Validation

Follow-up validation of candidate hits is typically conducted using isogenic human pluripotent stem cell (hPSC) lines harboring precise genetic modifications. Traditional CRISPR-mediated homology-directed repair (HDR) allows replacement of endogenous sequences but is limited by efficiency and the potential for indel formation. In contrast, base editing and prime editing enable high-efficiency, double-strand-break-free introduction of pathogenic single-nucleotide variants or small deletions, such as modeling or correcting 16p11.2 duplications [70]. Edited clones are rigorously genotyped to confirm the presence of exact intended modifications without off-targ*et al*terations. Subsequent parallel differentiation of wild-type and mutant clones facilitates high-resolution phenotypic analyses, including electrophysiological recordings and imaging. This strategy establishes causality and permits rescue experiments to verify that observed phenotypes arise specifically from the targeted variant or gene. Notably, focused gene libraries encompassing 200–500 targets achieve an optimal balance between experimental depth and scale. The inclusion of multi-guide redundancy, appropriate controls, and stringent quality control measures ensures reliable data interpretation. Isogenic validation is most robust when implemented in edit-and-rescue designs—combining loss-of-function modifications with cDNA rescue or CRISPR activation (CRISPRa)—thereby confirming causal effects and supporting downstream mechanistic studies in hPSC-derived models.

 doi:10.20944/preprints202511.1997.v1

6. Screening Formats and Experimental Design

6.1. Pooled Negative/Positive Selection Screens

Pooled CRISPR screens are a scalable and efficient method for interrogating gene function in hPSC-derived cells, such as neural progenitors or early neurons [71]. In this approach, cells are transduced with a pooled gRNA library and cultured over time. Guides targeting essential genes, including "housekeeping" genes, result in cell death or growth arrest, causing depletion of those sgRNAs ("dropout"), whereas guides conferring a growth advantage become enriched. This negative/positive selection strategy is analogous to screens performed in cancer cells to identify essential genes [72].

In ASD models, such screens can identify genes required for neural progenitor survival or proliferation. For example, differentiating iPSC progenitors into neurons allows assessment of gene knockouts that lead to neuron loss (sgRNA depletion) or excessive proliferation (sgRNA enrichment). While many ASD-associated genes are not strictly essential for survival, this approach remains a useful initial screen to verify library performance using positive controls and detect strong viability phenotypes. Notably, in hPSCs, double-strand break (DSB)-triggered p53 activation can bias selection tendencies; mitigating strategies include use of RNP delivery, base or prime editing to reduce DSB load, and rigorous clone screening [50].

6.2. Phenotypic Pooled Screens with Bulk Readouts

Phenotypic pooled screens extend beyond simple viability measurements by linking genetic perturbations to quantifiable cellular phenotypes. In this approach, a reporter or marker is engineered to reflect the phenotype of interest, and cells are subsequently separated using fluorescence-activated cell sorting (FACS) or alternative selection methods, with the associated single-guide RNAs (sgRNAs) in each fraction identified via next-generation sequencing (NGS). Representative examples of this strategy include the use of synaptic marker expression, where GFP is placed under the control of a synaptic protein promoter, such as PSD95, enabling the sorting of neurons exhibiting low versus high synaptic protein levels; cell surface markers, which can be assessed through FACS with antibody labeling; neurite outgrowth, measured via staining and imaging to capture structural phenotypes; transcriptional reporters, achieved by knocking GFP into genes of interest; and functional reporters, such as calcium indicators used to monitor neural activity [73]. This screening format facilitates the identification of regulators of cellular differentiation, including neuronal versus glial fate decisions, signaling pathways such as those revealed by phospho-flow, and cellular stress responses. Notably, limitations of this approach include its reliance on robust phenotypic markers and the potential loss of subtle phenotypic variations. FACS-based pooled screens have been successfully applied to analyze neuronal differentiation markers and synaptic proteins.

6.3. High-Content Pooled Single-Cell Readouts ("Perturbomics")

High-content single-cell methodologies, including Perturb-seq and CROP-seq, integrate pooled CRISPR-based genetic screens with single-cell RNA sequencing (scRNA-seq) to directly link genetic perturbations with high-dimensional transcriptional phenotypes [74]. In these approaches, each cell's transcriptome is captured along with its single-guide RNA (sgRNA) identity, allowing for precise profiling of gene expression alterations and cell-state transitions induced by individual perturbations. As shown in Figure 3, various CRISPR screening strategies have been applied in stem cell-derived neural systems, ranging from single-cell transcriptomic perturbations to spatial and functional assays in organoids and microglia. In human pluripotent stem cell (hPSC)-derived neural models, Perturb-seq facilitates the dissection of complex phenotypes, such as changes in neuronal subtype composition following transcription factor knockout [75,76]. Key technical considerations include the lower transcript counts observed in post-mitotic neurons, which necessitate either increased sequencing depth or larger cell numbers, as well as the heterogeneous populations generated during hPSC differentiation, which single-cell approaches can deconvolute to reveal cell-type-specific effects. Additionally, adequate statistical power typically requires profiling approximately 100–200 cells per

 doi:10.20944/preprints202511.1997.v1

guide [49], while controlling for multiplet rates, dropout events, and batch effects is essential, achievable through the use of unique molecular identifiers (UMIs) and experimental replicates. Perturb-seq has been successfully applied to excitatory and inhibitory neurons, as well as cerebral organoids, enabling the elucidation of gene-specific transcriptional networks pertinent to ASD. For instance, perturbation of FOXP1 preferentially impacts interneurons, highlighting cell-type-specific biases in ASD-relevant models [75,76].

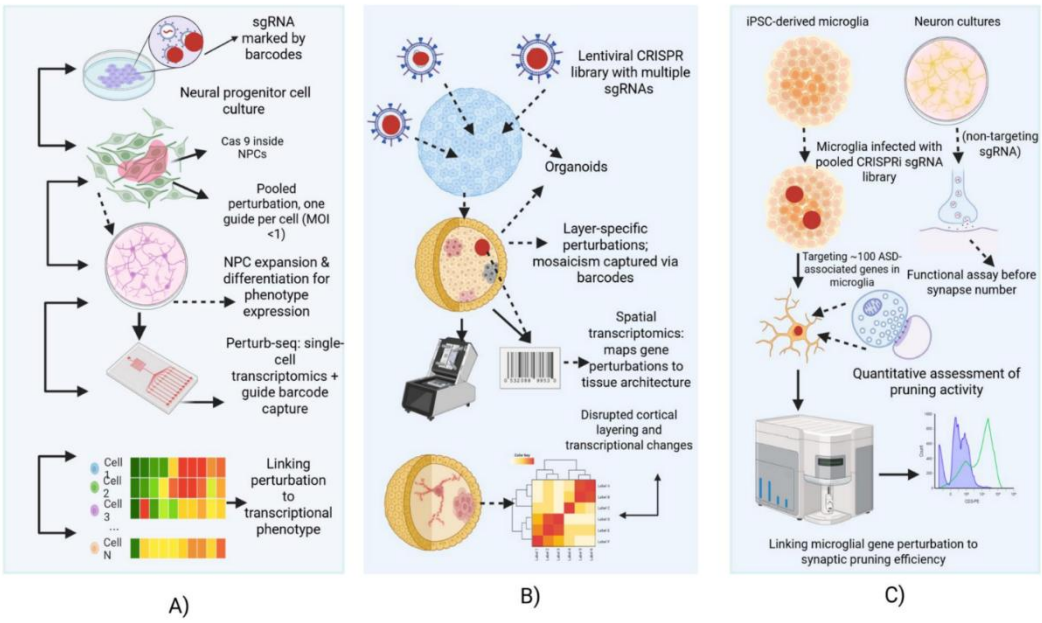

Figure 3. CRISPR screening strategies in stem cell-derived neural systems. (A) Pooled Perturb-seq in NPCs links single-guide perturbations to transcriptomic changes using scRNA-seq, (B) Spatial CRISPR screens in organoids combine pooled perturbation with spatial transcriptomics to reveal tissue-level phenotypes, and (C) CRISPRi screen in microglia assesses gene-specific effects on synaptic pruning via a fluorescent phagocytosis assay.

6.4. Spatially Resolved Pooled Screens and Integration with Spatial Transcriptomics

Emerging approaches integrate CRISPR perturbations with spatial transcriptomics to retain tissue context while profiling gene function [77]. Cells in organoids or neural cultures can be infected with a pooled guide library and later analyzed using spatial transcriptomic platforms, such as Slide-seq, MERFISH, or Visium. Spatial capture of sgRNA barcodes alongside gene expression allows assessment of non-cell-autonomous effects, tissue architecture, and cell–cell interactions [78].

Proof-of-concept studies demonstrate that spatial CRISPR screens can reveal phenotypes inaccessible via dissociated cell assays, including cortical layer formation and neuron migration in brain organoids [79]. Recent examples, such as CHOOSE [19], highlight pooled organoid screens as anchors for this approach. Next-generation platforms like AVITI24 aim to detect sgRNAs, protein markers, and morphology in a single automated workflow, streamlining spatial functional genomics.

6.5. Arrayed Screening for Electrophysiology and Calcium Imaging Readouts

Certain cellular phenotypes necessitate the use of arrayed formats, in which each perturbation is applied to an individual well or microculture, thereby enabling high-content assays for each specific perturbation [80]. This approach is particularly critical for morphological analyses, where automated imaging can quantify parameters such as neurite length, branching patterns, synapse number, and cell viability in 96- or 384-well plate formats [81]. In the context of electrophysiological studies, multi-electrode arrays (MEAs) facilitate the measurement of network-level activity, including firing rates and synchrony, whereas automated patch-clamp systems provide single-cell resolution [82]. Similarly, calcium imaging using GCaMP indicators combined with high-throughput microscopy or plate readers allows assessment of neuronal excitability and network connectivity [83].

 doi:10.20944/preprints202511.1997.v1

Although arrayed formats generally exhibit lower throughput compared to pooled screens, they are advantageous for longitudinal studies and for capturing subtle structural and functional phenotypes. Consequently, they are frequently employed to validate hits identified from pooled CRISPR screens.

7. Readouts and Multimodal Profiling

7.1. Single-cell RNA-seq

Single-cell RNA sequencing (scRNA-seq) has become a foundational readout for pooled CRISPR screens, providing an unbiased, high-resolution view of transcriptional consequences of genetic perturbations across thousands of individual cells [84]. In perturbation screens, each cell contains an associated guide RNA (gRNA), which can be captured via barcodes or CROP-seq cassettes. Unique molecular identifiers (UMIs) tag transcript molecules, allowing accurate quantification of transcript copies per cell [85].

Common platforms include microfluidic droplet systems (e.g., 10x Genomics Chromium) and plate-based methods (e.g., SMART-Seq on sorted cells). Droplet-based systems capture thousands of cells per run but at a shallow sequencing depth (~10,000–50,000 reads per cell), while plate-based methods achieve deeper coverage (~100,000+ reads per cell) at the cost of lower throughput. The choice between depth and throughput requires balancing sufficient cells per perturbation for statistical power versus adequate detection of genes of interest, especially in neural cells with large transcriptomes.

Quality control metrics include total UMIs per cell, number of genes detected, mitochondrial RNA fraction (a proxy for cell stress), doublet detection (two cells captured together), and the fraction of cells with exactly one guide RNA to ensure low multiplicity of infection (MOI) [86]. The scRNA-seq ultimately produces a digital gene expression matrix and a guide-cell matrix, enabling differential expression analysis per perturbation.

7.2. Single-cell ATAC-seq and Multiome for Regulatory Effects

Gene expression is modulated not only by DNA sequence but also by chromatin state. Single-cell ATAC sequencing (scATAC-seq) uses a Tn5 transposase to tag open chromatin regions, revealing regulatory elements such as enhancers and promoters [87]. Data are typically represented as “peak-by-cell” matrices, which can be aggregated by perturbation or cell type to identify differential chromatin accessibility.

The advent of single-cell multiome assays (e.g., 10x Multiome kit) enables simultaneous profiling of transcriptomes and epigenomes in the same cell [88]. This integration allows direct mapping of how perturbations alter regulatory element accessibility and gene expression, providing causal links between non-coding variants and transcriptional outcomes. Such analyses are particularly relevant in autism spectrum disorder (ASD), where a substantial portion of genetic risk is believed to reside in non-coding and regulatory regions [77].

Integrative analyses can link differential peak accessibility to target gene expression, constructing a comprehensive regulatory network affected by each perturbation. This is critical for understanding how mutations in chromatin remodelers (e.g., ARID1B, CHD8) impact neuronal gene networks.

7.3. Spatial Transcriptomics

While single-cell and multi-omic assays provide detailed molecular profiles, they lack spatial context, which is essential for studying tissue organization and non-cell-autonomous effects. Spatial transcriptomics technologies, such as 10x Visium, Slide-seq, MERFISH, and seqFISH, capture gene expression along with spatial coordinates [89].

In CRISPR perturbation screens, spatial transcriptomics can be applied to fixed sections of organoids or tissues. sgRNA identities can be co-detected spatially using barcoded probes, allowing researchers to map which perturbations occur in specific tissue regions [90]. For example, in cortical organoids, cells with a particular knockout may cluster in certain layers or exhibit disrupted migration. These

 doi:10.20944/preprints202511.1997.v1

approaches are vital for analyzing non-cell-autonomous phenotypes, a hallmark of ASD pathology, where a mutation in one cell type can affect neighboring cells or overall tissue architecture. Recently, spatially resolved CRISPR screens have been introduced, linking genetic modifications directly to tissue structure and organization [91].

7.4. Proteomics and Proximity Labeling

The transcriptome is an imperfect proxy for functional state, as mRNA levels do not always correlate with protein abundance, post-translational modifications, or activity. Proteomics, including mass spectrometry-based phosphoproteomics, provides a snapshot of global protein expression, signaling states, and modifications [92]. For perturbation screens, cells sorted by guide RNA can be analyzed to detect changes in protein abundance or phosphorylation, revealing post-transcriptional effects such as kinase activity or signaling cascade alterations.

Proximity labeling proteomics, using engineered enzymes like TurboID, BioID, or APEX, maps transient and weak protein-protein interactions [93]. This method can be combined with CRISPR to tag endogenous loci of genes, enabling interaction network mapping of gene products. Such approaches are particularly valuable for studying missense mutations, which may alter protein interactions or subcellular localization without changing expression levels. Integrating proteomics with transcriptomics uncovers mechanisms not apparent from RNA data alone.

7.5. Functional Phenotyping

High-throughput molecular datasets must be complemented with functional assays to establish causal relationships between genetic perturbations and cellular phenotypes [94]. Key functional phenotyping strategies encompass multiple complementary approaches. Electrophysiological techniques, including multi-electrode arrays (MEA) and patch-clamp recordings, allow measurement of neuronal activity and network properties, wherein genetic perturbations can alter firing patterns, burst activity, or synchrony. Synaptic assays employ imaging-based quantification of synapse number, morphology, and connectivity using markers such as PSD95 and Synapsin, with automated image analysis providing high-resolution insights into synaptic organization. Calcium imaging, utilizing fluorescent indicators such as GCaMP, enables monitoring of dynamic neural activity and serves as a proxy for neuronal excitability [95]. Synaptic pruning assays, involving co-culture of human iPSC-derived microglia with neurons, facilitate evaluation of microglial phagocytosis of synaptic material, and CRISPRi screens in microglia have identified key regulators of pruning and microglial states, including ADNP [96]. Collectively, these functional assays interrogate molecular, cellular, and network-level phenotypes, providing orthogonal evidence to elucidate the impact of genetic perturbations on neural circuits.

7.6. Multimodal Integration and Inferences

Combining scRNA-seq, scATAC-seq/multiome, spatial transcriptomics, proteomics, and functional phenotyping enables a holistic view of genetic perturbations [97]. Integrative analysis, often aided by AI or computational models, can infer causal chains, e.g., a CHD8 perturbation altering Wnt signaling, leading to progenitor cell imbalance [98]. This multimodal framework connects molecular, cellular, and tissue-level changes, providing mechanistic insights into ASD and other neurodevelopmental disorders.

8. Computational Pipelines and Statistical Analysis

The scale and complexity of data generated by pooled single-cell CRISPR screens necessitate sophisticated computational pipelines and statistical methods. This presents a significant bottleneck for researchers lacking bioinformatics expertise [99]. Rigorous computational workflows are essential to ensure high-quality, interpretable results from these high-dimensional perturbation experiments.

8.1. Preprocessing and Quality Control

 doi:10.20944/preprints202511.1997.v1

The first critical step in any pooled screening analysis is preprocessing the raw sequencing data. This includes aligning reads to the reference genome using tools such as STAR or CellRanger, demultiplexing cell barcodes, and accurately assigning sgRNA barcodes to each single-cell transcriptome to generate a cell-by-guide matrix [100]. Unique Molecular Identifiers (UMIs) are collapsed to count unique transcripts per gene per cell, producing a cell-by-gene count matrix. Quality control (QC) is paramount at this stage to filter out low-quality cells, doublets, and other artifacts that could confound downstream analysis. Standard QC metrics include the number of UMIs and genes per cell, the fraction of mitochondrial reads (high values indicate poor-quality cells), detection of doublets via anomalous transcript counts or software like DoubletFinder or Scrublet [101], and the distribution of cells per guide to ensure roughly uniform initial representation. Cells failing these QC thresholds are excluded, as the quality of all subsequent analyses depends on the rigor of this initial preprocessing step.

8.2. Differential Expression Analysis

Once the data is preprocessed and cleaned, differential expression (DE) analysis identifies genes whose expression is altered by specific perturbations [101]. In standard approaches, counts from cells sharing the same perturbation can be aggregated into pseudobulk data and analyzed using conventional RNA-seq tools such as edgeR, DESeq2, or limma-voom. However, specialized methods have been developed to handle the unique structure of single-cell perturbation data. MIMOSCA (Multi-Module Single-Cell CRISPR Analysis) uses a linear modeling framework to regress gene expression on perturbations while accounting for technical covariates and single-cell variability [102]. Similarly, SCEPTRE jointly models guide assignments and expression to control false positives [103]. Signature genes are identified transcriptome-wide for each perturbation, with statistical power dependent on the number of cells and effect sizes. Multiple testing correction, such as false discovery rate (FDR), is applied to avoid spurious findings. For Perturb-ATAC and related modalities, analogous analyses examine differential accessibility using logistic regression or other appropriate models [104]. Targeting 500–1000 cells per perturbation is typically required to achieve ~80% detection power [104].

8.3. Pseudotime and Trajectory Analyses

Pseudotime analysis is a powerful computational tool for mapping developmental trajectories, particularly in differentiating cell populations or organoid models relevant to neurodevelopmental disorders such as ASD [105]. Algorithms like Monocle and Slingshot order single cells along continuous developmental paths, enabling researchers to determine whether perturbations accelerate, delay, or divert cell fate decisions. By overlaying perturbation identities, one can detect developmental delays or arrests caused by gene knockouts. For example, ARID1B knockout in neural progenitors leads to a stall in progenitor states, reducing the number of cells reaching mature neuron stages. Perturbations can also be clustered by their effects along pseudotime, revealing functional groupings; in the CHOOSE brain organoid screen [19], members of the same chromatin remodeling complex exhibited similar shifts, highlighting shared regulatory mechanisms among ASD risk genes [106].

8.4. Network Reconstruction

Perturbation screens also facilitate the reconstruction of gene networks. Co-expression networks can be inferred across single cells using methods like WGCNA or graph-based clustering to identify modules of co-regulated genes. Perturbations of transcription factors or chromatin regulators allow the construction of causal networks, where consistent downregulation of a gene following a specific knockout indicates a directed influence (e.g., TF X → Gene Y) [107]. Tools such as SCENIC and ARACNe can infer regulatory networks, while Bayesian or dynamic causal modeling further refines the network by incorporating perturbation edges. Integrating co-expression with perturbation data

 doi:10.20944/preprints202511.1997.v1

prioritizes hub genes or pathways affected across multiple perturbations, providing a more tractable set of actionable targets for therapeutic or functional studies.

8.5. Integration of Multi-Omic Datasets

A growing frontier in functional genomics is the integration of multiple data modalities, including scRNA-seq, scATAC-seq, spatial transcriptomics, and proteomics. Tools like Seurat, Harmony, LIGER, and MOFA enable alignment of datasets by identifying shared latent factors or linking chromatin accessibility peaks to gene expression (e.g., Cicero or ArchR workflows). Spatial transcriptomic data can be annotated by transferring cell-type labels from scRNA-seq datasets (e.g., Seurat anchors or Tangram), while proteomics measurements are integrated via canonical correlation or multilayer graph approaches. Multi-omic integration allows researchers to observe coordinated effects across regulatory layers, such as a gene knockout simultaneously altering RNA transcripts, enhancer accessibility, and protein phosphorylation, providing a holistic view of cellular responses.

8.6. Handling Batch Effects and Statistical Rigor

The high-throughput nature of pooled single-cell screens, coupled with variability in human stem cell culture, presents statistical challenges. Batch effects—non-biological variations arising from sequencing runs, library preparation, or differentiation batches—must be corrected using methods such as ComBat, mutual nearest neighbors, or Harmony. Incorporating multiple biological replicates, randomized guide distributions, and pooling sufficient cells per perturbation (typically ~100–200 cells per guide) enhances statistical power. Power calculations should consider effect sizes, number of guides, desired FDR, and cells per guide[106,108]. Failure to address batch effects or inadequate replication can inflate false positives and compromise reproducibility.

Key Inferences: Rigorous preprocessing and QC are foundational for reliable single-cell CRISPR analyses. Specialized DE methods such as MIMOSCA enhance sensitivity to subtle perturbation effects, while pseudotime analysis provides temporal context for developmental perturbations. Network reconstruction moves beyond individual genes to prioritize key regulatory hubs and pathways. Multi-omic integration enables a comprehensive understanding of cellular states, and careful handling of batch effects and adequate replication ensures statistical robustness. Collectively, these strategies provide a framework for high-fidelity interpretation of pooled single-cell perturbation experiments.

9. Validation Strategies and Orthogonal Assays

9.1. Single-Clone Validation

Following a high-throughput screen, the critical next step is the validation of identified hits. Single-clone validation is considered the gold standard, wherein a specific gene is knocked out, mutated, or corrected in a clonal human pluripotent stem cell (hPSC) line to generate isogenic controls. This approach ensures that all cells within the clone carry the perturbation, eliminating heterogeneity seen in bulk populations, and confirms that the observed phenotype is due solely to the engineered mutation rather than confounding genetic background or off-target effects. Detailed assays are then performed after differentiating these clones into the relevant cell type. Key experiments include rescue experiments, where re-introduction of the wild-type gene or correction of the mutation in the mutant clone restores the normal phenotype. For example, if a knockout (KO) line shows reduced dendritic arborization, expressing a cDNA of the target gene to rescue expression should recover dendritic complexity, thereby providing a direct causal link between gene and phenotype. Single-clone isogenic validation combined with rescue experiments via CRISPRa further strengthens mechanistic confidence [109].

9.2. Orthogonal Perturbation to Confirm Directionality

To independently confirm gene–phenotype relationships, orthogonal perturbation strategies are employed, which use alternative methods to perturb the same gene.For instance, if a CRISPR-KO

 doi:10.20944/preprints202511.1997.v1

screen indicates that loss-of-function of gene X causes phenotype Y, RNA interference (shRNA or siRNA) can be used to knock down X, and concordant phenotypic outcomes are assessed. Conversely, CRISPRa-mediated overexpression can test whether gain-of-function produces the opposite phenotype, such as increased synapse number if a KO reduces it. Observing consistent or opposing effects across multiple technologies not only confirms directionality but also rules out method-specific artifacts. Convergent results from orthogonal perturbations, such as RhoA inhibition rescuing 16p11.2 migration defects, provide strong evidence for causality and reinforce the biological relevance of the gene–phenotype relationship.

9.3. Functional Assays: Electrophysiology, Synaptic, and Morphological Analysis

Beyond molecular or genetic validation, high-fidelity functional assays are essential to confirm that gene perturbations translate into meaningful cellular phenotypes. For neuronal studies, electrophysiological recordings are the gold standard. Patch-clamp techniques allow measurement of intrinsic properties such as action potential firing, synaptic currents, and ion channel function in individual neurons, whereas multi-electrode arrays (MEA) assess network-level excitability across populations. For example, a screen identifying gene X as reducing excitability would be validated by observing altered firing patterns on MEA or reduced currents in patch-clamp recordings.

Structural and synaptic validation can be achieved through high-resolution imaging, enabling dendritic tree reconstruction, spine counting, and quantification of synaptic density using tools such as NeuronStudio. Morphological assays are particularly informative when perturbations induce scaffold phenotypes, like reduced neurite outgrowth. Functional synaptic assays, including staining for vesicle release with FM dyes or recording miniature synaptic currents, provide additional validation of synapse integrity and functionality. Calcium imaging can also monitor neuronal activity in response to perturbations, linking molecular changes to functional output.

9.4. Cross-Model Validation (Patient iPSCs and Animal Models)

No single model perfectly recapitulates human brain biology; therefore, cross-model validation is essential. Findings from human iPSC-derived neurons or organoids can be corroborated using patient-derived iPSC lines carrying the same variant, assessing whether phenotypes are consistently reproduced. Complementary validation in animal models, such as mice or zebrafish engineered with the same genetic perturbations, allows examination of complex behaviors, neuroanatomical features, and circuit-level effects. Consistency across human cells, patient-derived lines, and animal models strengthens confidence in the relevance of the findings, whereas discrepancies can reveal species-specific mechanisms. This multi-model approach ensures that identified gene–phenotype relationships are robust, mechanistically informative, and broadly applicable for translational studies.

Inferences: Single-clone isogenic validation combined with rescue experiments confirms causal links between gene perturbation and phenotype. Orthogonal perturbations via RNAi or CRISPRa provide independent validation and confirm directionality. Functional assays including MEA, patch-clamp, synaptic imaging, and morphological reconstruction link molecular changes to neuronal behavior. Cross-model validation using patient iPSCs and animal models ensures robustness and translational relevance.

10. Case Studies: What's Been Done

10.1. Pooled CRISPR Screens in Neural Cells: Historic Milestones

Genome-wide CRISPR screens were first established in cancer cells over a decade ago, but their application to neural cells has only recently gained traction. Early studies primarily employed mouse neural stem cells or transformed lines to identify genes controlling neurite outgrowth, survival, or proliferation. A landmark 2016 study in primary mouse neurons used CRISPR to uncover regulators of axon regeneration, marking a shift toward more physiologically relevant models. Subsequently, human iPSC-derived neural progenitors and neurons were employed to interrogate complex

 doi:10.20944/preprints202511.1997.v1

neuronal phenotypes. For instance, combined CaMPARI2, a calcium integrator, with CRISPRi in iPSC neurons to screen hundreds of genes affecting neuronal excitability, identifying novel regulators that modulate firing rates. Similarly, pooled CRISPR-KO in 2024 highlighted synaptic regulators such as SYNGAP1, which alter neuronal excitability, with off-target effects mitigated using dual gRNAs [110].

Translationally, pooled CRISPR screens have targeted genes associated with neurodevelopmental and psychiatric disorders. In ASD, a small-scale screen of ~100 synaptic genes in iPSC neurons revealed that knockout of specific cell-adhesion genes impaired synaptic vesicle cycling. These studies collectively uncovered key regulators of neuronal survival, neurite dynamics, and electrophysiology, establishing a foundation for ASD-focused screens and highlighting the potential for discovering therapeutic targets.

10.2. CRISPRi Screens in Microglia and Circuit-Relevant Phenotypes (Synaptic Pruning)

Microglial CRISPRi screens have emerged as a critical tool to probe immune-neuron interactions in ASD. A 2023 *Nature* study conducted a pooled CRISPRi screen in human iPSC-derived microglia, targeting ~100 ASD-associated genes and measuring phagocytosis of fluorescently labeled synaptosomes. Knockdown of ADNP, an ASD-linked transcription factor, reduced synaptosome engulfment, implicating ADNP as a regulator of microglial synaptic pruning. Additional hits included autophagy and endocytosis genes, underscoring the mechanistic complexity of microglial function.

A complementary approach in 2025 revealed pruning deficits in SHANK3, linking excessive synapses to ASD-relevant phenotypes. These studies collectively demonstrate that microglia play a pivotal role in regulating circuit development and synaptic homeostasis. The screens highlighted successes such as profiling immune-neuron interactions and identifying regulators of phagocytosis, while also noting limitations including the immature state of *in vitro* microglia models and challenges in lentiviral transduction.

10.3. Pooled CRISPR Screens in Organoids — Successes, Bottlenecks, and Insights

CRISPR-based pooled screens have now been extended to 3D human organoids, enabling cell-type-specific and developmental analyses. The CHOOSE (CRISPR-Human Organoids-scRNA-seq) system [19] used a pooled lentiviral library targeting 36 high-confidence ASD transcriptional regulators, with two independent sgRNAs per gene, in human embryonic stem cell-derived cerebral organoids. Single-cell multiome profiling (RNA + ATAC) allowed mapping of gene perturbations to effects on cell fate and gene expression. Vulnerable cell types included dorsal intermediate progenitors, ventral forebrain progenitors, and upper-layer excitatory neurons, while gene regulatory network analyses revealed that BAF complex member knockouts (e.g., ARID1B) expanded ventral progenitors at the expense of excitatory neuron precursors. Patient iPSC-derived organoid validation confirmed phenotypes such as increased oligodendrocyte precursors.

Organoid screens have also revealed specific disruptions relevant to ASD. For example, FOXP1 knockout disrupted cortical lamination and highlighted convergence on Wnt signaling pathways (PMC 2024) [111]. Bottlenecks include the technical complexity of organoid culture, limitations in library delivery, and absence of vascularization, restricting long-term maturation. Nonetheless, these studies demonstrate the power of organoid models to reveal developmental and cell-type-specific phenotypes of ASD risk genes.

10.4. Spatially Integrated CRISPR Screens: Mapping Non-Cell-Autonomous Effects

Recent innovations integrate pooled CRISPR perturbations with spatial transcriptomics to map gene function in tissue context. A 2025 *Cell* study combined organoid infection with a pooled guide library and Slide-seq spatial transcriptomics, capturing sgRNA barcodes alongside mRNA to map perturbation effects to precise tissue locations. This approach revealed both cell-intrinsic and non-cell-autonomous effects; for instance, knockout of a cell-adhesion gene altered the local clustering of

 doi:10.20944/preprints202511.1997.v1

neurons, highlighting tissue-level architectural changes. Spatially integrated screens provide insights into cortical layering, glial-neuron interactions, and emergent phenotypes invisible to dissociated assays.

The field is also moving toward platforms such as gsMap and AVITI24, which integrate spatial transcriptomics, protein detection, and gRNA profiling to enable high-resolution mapping of gene perturbations across complex tissues [91]. These methods are poised to uncover critical ASD-related phenotypes that depend on tissue architecture and intercellular interactions, advancing our understanding of circuit-level dysfunction in neurodevelopmental disorders [91].

10.5. Inferences from Case Studies

Collectively, the reviewed studies highlight several key points: Microglia contribute to synaptic pruning deficits in ASD, suggesting potential for immune-targeted interventions. Organoid models capture cell-type-specific vulnerabilities and reveal convergent developmental pathways such as Wnt signaling and chromatin regulation [111]. Spatial CRISPR approaches enable mapping of non-cell-autonomous effects and tissue organization, providing a more holistic understanding of ASD gene function.

These case studies illustrate the complementary power of neural, glial, organoid, and spatially integrated CRISPR screens in elucidating the molecular and circuit-level underpinnings of ASD.

11. Reproducibility, Standardization, and Best Practices

11.1. Experimental Reporting Standards

Transparent and standardized reporting is crucial for reproducibility in pooled CRISPR and functional genomics screens. Key parameters to report include the complexity of the sgRNA library (number of guides and uniformity of distribution), the multiplicity of infection (MOI) used to transduce cells (ensuring mostly single-guide cells), and the coverage achieved (cells per guide). For example, a statement may read: “Library contained 5000 guides; cells were infected at MOI=0.3 to achieve ~20% transduction, followed by selection; screening maintained 2000× coverage (~2000 cells per guide).”90 Reporting these values allows assessment of screen quality and reproducibility. Recommended standards often include MOI <1 and coverage >500× to ensure robust statistical power. Additional details on Cas9 or dCas9 lines (expression levels, activity assays), duration of screens, and selective pressures (if any) should be documented. Utilizing standardized reporting checklists, analogous to ARRIVE guidelines, helps the community replicate and build upon published work.

11.2. Minimum Metadata to Report (Cell Line Source, Passage, Differentiation Protocol Details, Batch IDs)

Comprehensive metadata is essential for reproducibility and interpretability. Reports should specify the origin of the human pluripotent stem cell (hPSC) lines, including genetic background when relevant (e.g., “iPSC line C2 from an ASD patient with CHD8 mutation”) [109], along with passage numbers at the time of use. Culture medium composition, differentiation protocols (steps, timing, small molecules, growth factors), and batch identifiers for both cells and library preparation must also be included. For example, if differentiation occurred in two separate batches, this should be noted and ideally randomized with respect to perturbations. Slight variations in differentiation protocols can markedly influence outcomes, so transparent reporting is critical. Public repositories should include all metadata to enable replication under comparable conditions.

11.3. Data Deposition and Sharing (Raw FASTQ, Count Matrices, gRNA-Cell Maps, Metadata)

Open science principles require that raw and processed data be deposited in appropriate public repositories. Raw sequencing files (FASTQ) should be submitted to the Sequence Read Archive (SRA) or ArrayExpress, while single-cell count matrices can be shared via GEO or specialized portals like the Single Cell Portal (SCP) and Human Cell Atlas (HCA)[112]. The mapping of cell barcodes to sgRNA barcodes (“guide assignment” file) must be included, along with metadata describing sample conditions (e.g., library used, time point, cell type). When possible, analysis scripts or notebooks

 doi:10.20944/preprints202511.1997.v1

should accompany the datasets to improve transparency. For arrayed imaging or electrophysiology screens, raw image or voltage data, along with analysis code, should be archived. Recommended file formats include FASTQ, CSV/MTX matrices, CellML, NWB (electrophysiology), and OME-TIFF (images). Following FAIR principles—making data Findable, Accessible, Interoperable, and Reusable—is strongly encouraged.

11.4. Recommended QC Checklists and Power/Coverage Calculators

Prior to publication, screens should be evaluated using quality control checklists. Metrics include: ensuring ≥80% of guides are represented in plasmid counts, verifying that non-targeting controls behave neutrally, and confirming low Gini index values to indicate uniform guide distribution. For single-cell data, check that control cells form tight clusters and replicate guide results are consistent. Power calculations should be performed during experimental design, considering effect sizes, number of guides per gene, cells per guide, and replication. Tools such as simCAS can simulate single-cell CRISPR screens to estimate required scale. Coverage of at least 100 cells per guide per replicate is a common benchmark for detecting medium effect sizes. QC checklists may also include viability, editing efficiency, and Cas9 activity checks to ensure robust perturbation rates.

11.5. Community Standards Initiatives and Consortia (Recommended Standardized Nomenclature for Perturbation IDs)

The field of functional genomics benefits substantially from community-driven standardization efforts, analogous to large-scale genomic initiatives. Consortia such as the Human Cell Atlas (HCA), the International Genomics of Variation in Function (IGVF), and other collaborative initiatives aim to standardize single-cell data annotation, induced pluripotent stem cell (iPSC) biobanks, and differentiation protocols, thereby reducing variability across laboratories. Standardized nomenclature for genetic perturbations is strongly recommended, including the use of HGNC gene symbols and Addgene identifiers for constructs, alongside the deposition of detailed library descriptions in publicly accessible repositories. Establishment of a community registry for perturbation libraries would further facilitate cross-study comparisons [113]. Collaborative efforts among consortia such as SFARI, CRISPR-focused initiatives, and single-cell research groups can accelerate the adoption of unified protocols and data reporting standards, enabling interoperable datasets and large-scale meta-analyses. Key standards include library design with multiplicity of infection (MOI) less than 1 and coverage greater than 500×, detailed metadata encompassing cell source, passage number, differentiation protocols, and batch identifiers, and data deposition in repositories such as GEO or SRA for FASTQ files and matrices, as well as SCP or HCA for single-cell datasets. Quality control measures should assess library complexity, editing efficiency, control performance, and coverage. Community standards also encompass harmonized perturbation identifiers and standardized iPSC/CRISPR libraries. Initiatives such as IGVF are expected to further standardize iPSC biobanks, differentiation protocols, and perturbation nomenclature, thereby reducing experimental variability and enhancing reproducibility.

12. Ethical, Legal, and Social Implications (ELSI)

12.1. Human Stem Cell Research Governance and ISSCR Recommendations

Research involving human pluripotent stem cells (hPSCs), particularly embryonic stem cell lines and organoid models, raises substantial ethical considerations. The International Society for Stem Cell Research (ISSCR) provides internationally recognized guidelines to address these issues. The 2025 update focuses solely on embryo-model research; the 2021 framework remains unchanged, and uterine transfer is not allowed (International Society for Stem Cell Research, 2025). Brain organoids, as they gain structural and functional complexity, prompt debate regarding moral status; the current consensus holds that organoids lack sensory experience, yet regulatory frameworks are evolving. Researchers must obtain informed consent from donors, ensure compliance with local laws and institutional ethical review boards (IRB/IACUC), and maintain transparency, such as by publishing

 doi:10.20944/preprints202511.1997.v1

experimental protocols and approvals [114]. Governance following ISSCR 2025 ensures that ethical and scientific integrity is maintained while enabling responsible innovation.

12.2. Germline Editing vs. Somatic/Cell Models

It is critical to differentiate somatic or cell-based research from germline editing. The functional genomics studies in autism spectrum disorder (ASD) predominantly utilize patient-derived somatic cell models, including hPSCs and iPSCs, where genetic manipulations are confined to *in vitro* systems [115]. CRISPR-based screens in these models, including pooled organoid platforms such as CHOOSE, aim to elucidate mechanistic pathways and do not advocate heritable genome modification [19]. In contrast, germline editing involves permanent changes transmissible to future generations and is widely considered ethically impermissible. Ethical clarity in publications requires emphasizing the somatic, non-heritable nature of experimental manipulations, and avoiding speculative discussions about editing human embryos. Oversight ensures that research on severe ASD can proceed within an established ethical framework for somatic cell studies, while germline interventions remain prohibited due to heritability risks

12.3. Privacy and Data Sharing for Patient-Derived iPSCs

Patient-derived iPSCs carry the full genomic information of the donor, including ASD risk alleles and potentially unrelated health variants, raising privacy concerns. Informed consent must explicitly cover stem cell derivation, genomic sequencing, and controlled data sharing. Access to genomic datasets, such as whole-genome sequences, should be restricted to controlled-access repositories (e.g., dbGaP, EGA) to prevent inadvertent re-identification. Compliance with regulatory frameworks like HIPAA and GDPR, coupled with ongoing donor engagement regarding data use, is essential. Anonymization challenges are heightened by unique genetic mutations that could theoretically identify individuals. Ethical data management ensures that privacy risks are mitigated while supporting reproducible and collaborative research.

12.4. Societal Implications of Mechanistic Findings

Mechanistic insights into ASD genetics carry significant societal consequences. On one hand, such knowledge can improve diagnostic precision and inform therapeutic development. On the other, it raises ethical and social concerns, including potential stigma, genetic determinism, and implications for prenatal testing. For example, validation of a gene as causative may influence prenatal screening panels, creating complex decisions for families regarding preparation or pregnancy termination. Emphasis on genetic risk must be balanced with recognition of neurodiversity and the contributions of environmental and personal factors. Communication strategies should be developed in collaboration with ethicists, patient communities, and clinicians to ensure responsible translation of findings. Moreover, understanding biases in ASD diagnosis, such as the female-to-male ratio affected by underdiagnosis, can help contextualize research without reinforcing deterministic narratives.

12.5. Responsible Translational Pathways and Stakeholder Engagement

Translating scientific discoveries into clinical applications requires a coordinated, multi-stakeholder approach. Early involvement of patient advocacy organizations, such as the Simons Foundation, Autism Speaks, and the Autistic Self Advocacy Network, is critical to ensure that research priorities are aligned with patient values. Potential therapeutic strategies, including pharmacological and gene-based interventions, must undergo thorough preclinical evaluation followed by rigorous regulatory review by agencies such as the FDA and EMA. Continuous oversight guided by ethics, legal, and social implications (ELSI) frameworks, in conjunction with interdisciplinary collaboration among social scientists, clinicians, and patients, is essential to maximize benefits while minimizing potential harms. The overarching objective is to conduct

 doi:10.20944/preprints202511.1997.v1

inclusive, ethically guided research wherein mechanistic insights inform responsible clinical translation.

According to the ISSCR 2025 guidelines, clear governance is provided for embryo and organoid research, emphasizing ethical oversight and explicitly prohibiting uterine transfer (International Society for Stem Cell Research, 2025). Somatic cell models, including pooled organoid screens such as CHOOSE, are deemed ethically permissible for ASD research, whereas germline interventions remain prohibited due to heritability risks. Robust privacy safeguards, controlled data sharing, and informed consent are mandatory for the use of patient-derived iPSCs and genomic data. Mechanistic findings must be communicated responsibly to mitigate stigma, promote neurodiversity, and avoid deterministic interpretations of genetic information. Engagement of relevant stakeholders combined with rigorous ELSI oversight ensures that translational pathways remain safe, ethical, and inclusive.

13. Translational Potential and Therapeutic Discovery

13.1. Target Discovery and Validation

One of the primary goals of functional genomics in Autism Spectrum Disorder (ASD) is to identify therapeutic targets. High-throughput CRISPR-based screens systematically uncover genes whose perturbation alters neuronal phenotypes, providing candidate targets for drug development. A typical pipeline involves: (1) performing a large-scale pooled screen to identify genes that, when perturbed, produce a desired phenotype; (2) validating these "hits" in a lower-throughput, arrayed format using high-fidelity functional assays such as multi-electrode arrays (MEA); and (3) using isogenic cell lines to confirm a direct causal link between the gene and the observed phenotype. This approach ensures that candidate targets are both impactful and specific to neuronal biology. In ASD, actionable targets might include synaptic proteins, signaling molecules, or epigenetic regulators that are drug-accessible. Hits from screens—such as RhoA, identified for its role in synaptic dysfunction—can then be validated and prioritized for downstream therapeutic development. AI-based prioritization can accelerate the selection of targets most likely to translate from bench to clinical trial. Where pooled organoid screens have been applied (e.g., "CHOOSE"), these platforms exemplify how complex human neural models can anchor CRISPR-based discoveries in structural and developmental contexts.

13.2. Small-Molecule and Biologic Development Informed by Perturbation Phenotypes

Functional genomics screens not only identify targets but also directly inform small-molecule and biologic development. Perturbation phenotypes from CRISPRi or CRISPRa can guide therapeutic strategies: for instance, if CRISPRi of gene X produces a beneficial phenotype, small molecules mimicking that effect can be sought; conversely, if knockout causes detrimental phenotypes, inhibitors or activators can be used to restore function. Phenotypic readouts, including synapse number, neuronal activity, and circuit-level alterations, serve as relevant endpoints. For example, a mouse model of ASD revealed hyperactivity in the reticular thalamic nucleus as a driver of autism-like behaviors, which was reversible with repurposed epilepsy drugs, demonstrating that functional genomics can inform actionable, circuit-level interventions. Biologics such as growth factors or antibodies can also be tested in cellular models if they target CRISPR-identified pathways. Additionally, chemogenomic screens combining CRISPR perturbations with chemical libraries allow systematic identification of compounds that mimic or suppress specific genetic effects, exemplified by small molecules like Rhosin targeting RhoA.

13.3. Gene-Therapy Design Considerations

Gene therapy in ASD is challenging due to genetic heterogeneity, but certain gene classes are promising targets. Haploinsufficient genes (Category 1, e.g., CHD8, ADNP, SCN2A) can potentially benefit from viral vector-mediated delivery (e.g., AAV) to restore normal gene dosage. AAV-based gene replacement has been successful in other monogenic neurodevelopmental disorders such as Angelman syndrome. Loss-of-function missense variants may be addressed using wild-type cDNA

 doi:10.20944/preprints202511.1997.v1

delivery or RNA-based therapies, including antisense oligonucleotides for exon skipping or splicing modulation. Dominant-negative mutations affecting synaptic receptors are more challenging, but approaches such as CRISPRa to boost the normal allele or CRISPR/Cas13 to selectively silence the mutant transcript are being explored.

The CRISPR screens further inform which ASD genes possess functional domains amenable to drug targeting or which pathways converge on "druggable" nodes like mTOR or neurotransmitter receptors. While polygenic forms of ASD complicate gene therapy approaches, monogenic or syndromic cases (e.g., MECP2 mutations) are more tractable. CRISPR-based perturbation studies can reveal which genes, when up- or downregulated, have the strongest effects on neural function, guiding candidate selection for gene therapy. Prime and base editors enable precise, single-variant corrections, broadening the therapeutic possibilities. Because double-strand break–inducing nucleases may activate a p53 checkpoint in human pluripotent stem cells (hPSCs), strategies such as ribonucleoprotein (RNP) delivery, base or prime editing, and stringent clone screening are recommended to mitigate editing toxicity.

13.4. Personalized Medicine: Patient iPSC Panels and Isogenic Variant Correction

Patient-derived induced pluripotent stem cells (iPSCs) provide a platform for personalized therapy development. Panels of iPSCs carrying diverse ASD-linked variants—such as those from the SFARI iPSC library—allow systematic testing of variant pathogenicity. Isogenic correction, where a patient-specific mutation is corrected in iPSCs, can reveal causality by comparing corrected and uncorrected lines. Conversely, introducing a variant of unknown significance into control iPSCs allows assessment of its functional impact. This "variant-to-function" approach refines diagnostics by distinguishing benign polymorphisms from pathogenic mutations and informs individualized therapeutic strategies. For example, hyperexcitability in neurons carrying a channelopathy mutation could guide treatment with specific channel-blocking drugs. Using patient iPSC panels in combination with CRISPR-based screens connects genetic diagnosis directly to tailored interventions. To strengthen causality, edit-and-rescue strategies—loss-of-function perturbation followed by cDNA rescue or CRISPRa—are recommended as the standard for isogenic validation (Dixit, *et al*., 2016).

13.5. Regulatory Pathway for Therapeutic Translation

Translating findings from CRISPR-based functional genomics screens into clinical applications necessitates careful navigation of regulatory frameworks. For gene and cell therapies, regulatory agencies such as the FDA require rigorous demonstration of safety and efficacy in preclinical models. CRISPR screens contribute mechanistic insights, facilitate the identification of biomarkers of response, and support the establishment of predictive endpoints. For example, correction of a synaptic phenotype *in vitro* should ideally be associated with corresponding behavioral improvements in relevant animal models. Early engagement with regulatory authorities regarding novel endpoints, including molecular or physiological biomarkers derived from CRISPR screens, can help streamline approval pathways. Additionally, companies are developing regulatory-ready platforms, such as the submission of GMP-grade human induced pluripotent stem cell (hiPSC) lines and the pursuit of designations like Regenerative Medicine Advanced Therapy (RMAT), which can accelerate the development of promising regenerative products. Ensuring assay relevance to clinical outcomes, coupled with collaboration between clinical and industry partners, is critical for translating functional genomics findings into safe and effective ASD therapies. Inferences from these approaches indicate that AI-prioritized CRISPR hits can accelerate the pipeline from screens to clinical trials. Functional genomics enables the identification of both gene-specific and circuit-level targets suitable for small molecules, biologics, and gene therapies, while patient-specific iPSC models and isogenic corrections support precision medicine strategies. Regulatory engagement further ensures the translational feasibility, safety, and clinical relevance of these novel therapeutic interventions.

14. Challenges, Limitations, and Open Questions

 doi:10.20944/preprints202511.1997.v1

14.1. Model Limitations

Despite immense progress, current *in vitro* models have inherent limitations and do not perfectly replicate the human brain. Human pluripotent stem cell (hPSC)-derived neurons, even after months in culture, generally resemble embryonic or early postnatal neurons [116]. They often lack full dendritic complexity, myelination, or mature synaptic properties. Organoids improve this situation but still capture only early developmental stages. Furthermore, organoids lack a functional vascular system and do not fully recapitulate long-term maturation, limiting their capacity to model later developmental stages or key processes such as synaptic pruning in the context of an adult brain. Environmental cues present *in vivo*—including blood supply, immune system interactions, sensory inputs, and endocrine signals—are absent or incomplete. Co-cultures and organ-on-chip systems can partially address these gaps, but no *in vitro* model fully mimics the *in vivo* milieu. Consequently, researchers must interpret results with caution: a gene showing no effect *in vitro* may still play an important role *in vivo*, and vice versa. For disorders such as ASD, which manifest behaviorally, linking cellular phenotypes back to brain function represents a major challenge. Hybrid organoid-animal models may partially address these gaps by combining the advantages of both systems.

14.2. Technical Constraints for Scaling Pooled CRISPR in Organoids

While pooled CRISPR screens are highly scalable in 2D cultures, several technical hurdles limit their application in 3D organoids [117]. First, uniform delivery of gRNA libraries is challenging: lentiviral penetration into three-dimensional tissues is limited, and achieving single-copy infection across thousands of organoids is nontrivial. Organoids derived from a single pool often display mosaicism, with each organoid containing multiple different perturbations, complicating interpretation. The CHOOSE approach partially overcomes this by barcoding, but scaling to hundreds of organoids remains complex. Second, organoid assays, such as single-cell RNA-sequencing, are expensive on a per-cell basis, which limits throughput. Third, inherent variability between organoid batches—including differences in cell composition—can mask the effects of perturbations. These limitations currently restrict organoid CRISPR screens to targeted gene sets rather than genome-wide approaches. Moreover, the absence of vascularization and long-term maturation further constrains the fidelity and scalability of these models.

14.3. Interpreting Pleiotropic Gene Effects and Cell Non-Autonomous Phenotypes

A pervasive challenge in ASD research is the pleiotropy of candidate genes. Many ASD-associated genes have multiple roles across neuronal and non-neuronal cell types. Perturbation of such a gene can produce phenotypes either through cell-autonomous mechanisms or via altered neuron-glia signaling in mixed cultures. Some genes, particularly synaptic adhesion molecules, function at the system level, meaning a knockout may only manifest a phenotype in the presence of a normal partner cell. Dissecting direct versus indirect effects often requires complementary arrayed co-culture experiments or single-cell-type isolations. Deconvoluting these complex, multi-level effects to identify causal mechanisms for disease phenotypes demands sophisticated computational tools and multi-modal datasets.

14.4. Statistical and Computational Challenges (False Positives/Negatives, Multiple Testing)

High-throughput and single-cell CRISPR screens generate massive datasets, which introduce significant statistical and computational challenges [118]. Multiple hypothesis testing inflates the risk of false positives, while stringent correction methods such as Benjamini-Hochberg false discovery rate (FDR) adjustment can increase false negatives, potentially missing real hits. Complex covariance structures, including batch effects and cell-cycle variation, further complicate differential expression analyses. Technical noise inherent to single-cell RNA-seq, such as dropout events and low read counts, presents additional challenges for accurate signal detection. Algorithms like MIMOSCA or SCEPTRE provide built-in controls, but no method is perfect. Cross-validation through independent experiments is essential to confirm hits. The computational demands are high, as single-cell CRISPR

 doi:10.20944/preprints202511.1997.v1

datasets can encompass millions of cells, necessitating optimized pipelines and substantial computing resources. Open-source tools such as MAGIC, Seurat, and Scanpy facilitate analysis, but expertise in bioinformatics remains critical [119].

14.5. Prioritizing Hits for Translational Follow-Up

High-throughput screens frequently generate extensive lists of candidate genes, posing a significant challenge in determining which hits are most suitable for translational follow-up. Prioritization of these candidates typically considers criteria such as effect size, reproducibility across multiple guides and replicates, and biomedical tractability, including the presence of known drug targets. Genes previously implicated in ASD through human genetic studies are prioritized for validation, whereas novel genes have the potential to uncover previously unrecognized biological pathways. Integration with complementary data types—including co-expression networks, protein-protein interaction maps, and patient-derived datasets—facilitates the identification of the most plausible candidate genes [120]. Practical considerations, such as the availability of reagents, ease of cloning, and the existence of animal models, also influence decision-making. Furthermore, community-wide initiatives that cross-reference CRISPR screen hits with patient variants or clinical data provide additional guidance for selecting targets for downstream translational studies.

Despite these advances, modeling polygenic risk and prioritizing hits for therapeutic intervention continue to present significant challenges. Emerging hybrid organoid-animal models offer promising strategies to address current limitations in maturation, vascularization, and system-level interactions, thereby improving the translational relevance of candidate gene validation.

15. Future Directions and Roadmap

15.1. Integrating Spatially Resolved Perturbomics and Multi-Omics

The next generation of ASD functional genomics will rely on the seamless integration of spatially resolved perturbomics, multi-omics, and high-throughput electrophysiology. Spatial transcriptomics will become routine, linking genetic perturbations to tissue architecture, while multi-omics approaches—including simultaneous scRNA-seq, scATAC-seq, and protein quantification in the same cells—will provide richer, multi-dimensional phenotypes. Advanced imaging techniques, such as light-sheet microscopy of whole organoids and *in vivo* two-photon imaging, could eventually capture functional readouts alongside CRISPR perturbations. Robotics and high-throughput platforms will facilitate parallel electrophysiology or live imaging, enabling comprehensive functional screens. A plausible roadmap envisions a converged "spatial perturb-seq," running screens that profile transcriptomes, epigenomes, and activity (via calcium imaging or voltage reporters) in an intact 3D neural tissue [121]. Importantly, proof-of-concept studies already demonstrate spatially resolved CRISPR screens that link genomic perturbations with tissue structure [51]. Machine learning and AI will be essential to analyze these vast datasets, uncovering patterns beyond human perception, and allowing a shift from single-cell analyses to understanding complex neural circuits.

15.2. Scaled Patient-Derived iPSC Biobanks and Standardized CRISPR Libraries for Population-Scale Functional Genomics

To capture the full spectrum of ASD genetic risk, large-scale, community-driven biobanks of patient-derived iPSCs are critical. Each line would have a known genotype and could be systematically combined with standardized CRISPR libraries based on curated resources such as the SFARI Gene database [122]. This approach would allow population-scale functional genomics, assessing how genetic background influences gene function (gene-by-gene interactions). For feasibility, the community should agree on standard reference lines (e.g., a few "universal" iPSC lines) and common guide RNA libraries. Initiatives such as SFARI and the Human Cell Atlas (HCA) are already moving in this direction. Such standardization would accelerate cross-study comparisons

 doi:10.20944/preprints202511.1997.v1

and enable meta-analyses, extending modeling beyond individual patients to the entire range of ASD risk alleles.

15.3. Use of Base and Prime Editors to Directly Model Patient Variants at Scale

Future screens will focus increasingly on variant-level modeling rather than gene-level knockouts. Base and prime editors will enable the creation of libraries where each cell carries a distinct ASD-linked mutation, allowing high-throughput interrogation of variant effects on neuronal physiology. For example, tiling prime-editing guide RNAs across exons of genes such as *SCN2A* could facilitate parallel functional assessment of thousands of clinically reported variants, including variants of uncertain significance (VUS). In a prospective scenario, sequencing a patient genome could reveal a novel missense mutation; pre-existing variant libraries would allow immediate functional testing, directly informing pathogenicity. These advances, combined with high-content phenotyping, could significantly improve clinical interpretation of ASD-linked variants. Notably, base/prime editing reduces double-strand break–triggered *TP53* activation observed in hPSCs, a known limitation of nuclease-based editing [123].

15.4. AI/ML Integration for Prioritization and Causal Inference

The immense datasets generated by multi-modal ASD screens are beyond the capacity of traditional analysis. AI and machine learning will be crucial for prioritizing genes or variants for experimental follow-up, predicting functional consequences, and inferring causal relationships in complex biological networks. Deep learning approaches, including autoencoders and graph neural networks, can integrate multi-omic perturbation data to model developmental programs, cell fate outcomes, and the functional impact of perturbations. As large public datasets accumulate, including perturbation atlases, AI-driven approaches will be indispensable for generating insights that humans alone cannot synthesize [117]. Recent advances further highlight how AI-based trajectory inference can resolve cell fate outcomes in CRISPR-perturbed neural systems [45].

15.5. Proposed Community Blueprint: Standard Reference Lines, Standard Readouts, and Common Data Formats

A coordinated community "blueprint" will be critical for advancing functional genomics in autism spectrum disorder (ASD). This blueprint entails consensus on standard reference cell lines—analogous to HEK293 for kidney research and a limited number of induced pluripotent stem cell (iPSC) lines for neuronal studies—along with standardized differentiation protocols to minimize inter-laboratory variability. Core phenotypic readouts, including synapse number, neuronal firing rate, and key gene-expression signatures, are essential to ensure consistency across studies. Adoption of standardized data formats, such as the Human Cell Atlas's Loom or AnnData for perturb-seq datasets, will facilitate data sharing and interoperability. Governance through collaborative consortia, modeled after initiatives like ENCODE or the Human Cell Atlas, will ensure adherence to these standards. Collectively, these efforts provide a roadmap for causal genomics, with the potential to transform ASD research and ultimately improve patient care [113]. Integration of spatial perturbomics, multi-omics profiling, high-throughput electrophysiology, scaled iPSC biobanks, variant-level modeling, and AI-driven analytical frameworks, all within a community-standardized blueprint, will enable comprehensive, causal functional genomics in ASD. Such an approach is poised to accelerate the translation from genotype to phenotype and drive the development of targeted therapeutic strategies.

16. Conclusion

CRISPR-enabled functional genomics in human stem cell models is transforming the study of autism spectrum disorder (ASD) biology. By systematically perturbing ASD-associated genes in relevant cell types and analyzing molecular and cellular phenotypes, these platforms are bridging the gap between genetic association and biological mechanism. Early studies have already uncovered

 doi:10.20944/preprints202511.1997.v1

novel roles for known genes, such as ADNP in microglia and ARID1B in progenitor fate, while revealing vulnerable cell populations, highlighting the power of these approaches in identifying causal relationships in human-relevant contexts.

The convergence of CRISPR tools—including nucleases and base editors—with advanced cellular models such as neurons, glia, and organoids, alongside rich readouts including single-cell omics, high-resolution imaging, and electrophysiology, has created an unprecedented data stream. This integration is enabling systematic mapping of gene-to-phenotype relationships, moving the field beyond the search for individual disease-causing genes toward the identification of convergent molecular pathways and circuit-level dysfunctions that can be targeted for therapeutic intervention.

Despite the inherent complexity of ASD—including its heterogeneity, pleiotropy, and numerous variants of unknown significance—these sophisticated platforms offer an unprecedented opportunity to unravel the disorder's intricate pathobiology. The technological roadmap ahead emphasizes the integration of spatial, multi-omic, and high-throughput electrophysiological readouts, the creation of large-scale standardized biobanks, and the strategic application of advanced computational methods to analyze and interpret complex datasets.

While significant technical and ethical challenges remain—ranging from limitations of current *in vitro* models to the complexities of multi-modal data analysis—the rapid progress of the field is undeniable. The ultimate goal is to leverage these insights for the benefit of individuals with ASD by refining diagnoses, uncovering novel therapeutic targets, and informing precision medicine strategies. CRISPR-stem cell platforms, therefore, revolutionize ASD functional genomics, offering unparalleled promise for bridging genetics, mechanisms, and therapies, and paving the way for personalized and effective interventions for the millions of individuals and families affected by autism.

Funding: The author(s) acknowledge the funds received by Uddalak Das from the **Department of Biotechnology** (Grant No: DBTHRDPMU/JRF/BET-24/I/2024-25/376), the **Council of Scientific and Industrial Research** (Grant No: 24J/01/00130), and the **Indian Council of Medical Research** (Grant No: 3/1/3/BRET-2024/HRD (L1)).

Authorship Contribution Statement: S.T. Gopukumar: Writing Review & Editing; **Madhumita Saha: V**alidation, Writing Original Draft, Writing Review & Editing, Visualization, Supervision, Project Administration**; Tanveen Kaur Soni:** Writing Original Draft, Writing Review & Editing, Visualization**; Samer Shamshad:** Writing Original Draft, Writing Review & Editing**; Sunil Tadakod:** Writing - Review & Editing**; Alok Sharma:** Writing Review & Editing**; Uddalak Das:** Conceptualization, Project Administration, Supervision, Funding Acquisition, Review & Editing. All authors critically revised the paper and approved the final version.

Declaration of Competing Interest: The author(s) report no conflict of interest.

Acknowledgments: None.

Ethical Statements: None.

Declaration of generative AI and AI-assisted technologies in the writing process: The writing of this review paper involved the use of generative AI and AI-assisted technologies only to enhance the clarity, coherence, and overall quality of the manuscript. The authors acknowledge the contributions of AI in the writing process while ensuring that the final content reflects the author's own insights and interpretations of the literature. All interpretations and conclusions drawn in this manuscript are the sole responsibility of the author.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders [Internet]. Fifth Edition. American Psychiatric Association; 2013 [cited 2025 Sep 6]. Available from: https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596
2. 2. Centers for Disease Control and Prevention. (2022). Prevalence and characteristics of autism spectrum disorder among children aged 8 years — Autism and Developmental Disabilities Monitoring

 doi:10.20944/preprints202511.1997.v1

Network, 16 sites, United States, 2022. Morbidity and Mortality Weekly Report Surveillance Summaries, 71(8), 1-23.

3. Lai MC, Lombardo MV, Baron-Cohen S. Autism. The Lancet. 2014 Mar;383(9920):896–910.
4. Lavelle TA, Weinstein MC, Newhouse JP, Munir K, Kuhlthau KA, Prosser LA. Economic Burden of Childhood Autism Spectrum Disorders. Pediatrics. 2014 Mar 1;133(3):e520–9.
5. Lukmanji S, Manji SA, Kadhim S, Sauro KM, Wirrell EC, Kwon CS, Jetté N. The co-occurrence of epilepsy and autism: A systematic review. Epilepsy Behav. 2019 Sep;98:238–48.
6. Sandin S, Lichtenstein P, Kuja-Halkola R, Hultman C, Larsson H, Reichenberg A. The Heritability of Autism Spectrum Disorder. JAMA. 2017 Sep 26;318(12):1182.
7. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L, Stevens C, Reichert J, Mulhern MS, Artomov M, Gerges S, Sheppard B, Xu X, Bhaduri A, Norman U, Brand H, Schwartz G, Nguyen R, Guerrero EE, Dias C, Betancur C, Cook EH, Gallagher L, Gill M, Sutcliffe JS, Thurm A, Zwick ME, Børglum AD, State MW, Cicek AE, Talkowski ME, Cutler DJ, Devlin B, Sanders SJ, Roeder K, Daly MJ, Buxbaum JD, Aleksic B, Anney R, Barbosa M, Bishop S, Brusco A, Bybjerg-Grauholm J, Carracedo A, Chan MCY, Chiocchetti AG, Chung BHY, Coon H, Cuccaro ML, Curró A, Dalla Bernardina B, Doan R, Domenici E, Dong S, Fallerini C, Fernández-Prieto M, Ferrero GB, Freitag CM, Fromer M, Gargus JJ, Geschwind D, Giorgio E, González-Peñas J, Guter S, Halpern D, Hansen-Kiss E, He X, Herman GE, Hertz-Picciotto I, Hougaard DM, Hultman CM, Ionita-Laza I, Jacob S, Jamison J, Jugessur A, Kaartinen M, Knudsen GP, Kolevzon A, Kushima I, Lee SL, Lehtimäki T, Lim ET, Lintas C, Lipkin WI, Lopergolo D, Lopes F, Ludena Y, Maciel P, Magnus P, Mahjani B, Maltman N, Manoach DS, Meiri G, Menashe I, Miller J, Minshew N, Montenegro EMS, Moreira D, Morrow EM, Mors O, Mortensen PB, Mosconi M, Muglia P, Neale BM, Nordentoft M, Ozaki N, Palotie A, Parellada M, Passos-Bueno MR, Pericak-Vance M, Persico AM, Pessah I, Puura K, Reichenberg A, Renieri A, Riberi E, Robinson EB, Samocha KE, Sandin S, Santangelo SL, Schellenberg G, Scherer SW, Schlitt S, Schmidt R, Schmitt L, Silva IMW, Singh T, Siper PM, Smith M, Soares G, Stoltenberg C, Suren P, Susser E, Sweeney J, Szatmari P, Tang L, Tassone F, Teufel K, Trabetti E, Trelles MDP, Walsh CA, Weiss LA, Werge T, Werling DM, Wigdor EM, Wilkinson E, Willsey AJ, Yu TW, Yu MHC, Yuen R, Zachi E, Agerbo E, Als TD, Appadurai V, Bækvad-Hansen M, Belliveau R, Buil A, Carey CE, Cerrato F, Chambert K, Churchhouse C, Dalsgaard S, Demontis D, Dumont A, Goldstein J, Hansen CS, Hauberg ME, Hollegaard MV, Howrigan DP, Huang H, Maller J, Martin AR, Martin J, Mattheisen M, Moran J, Pallesen J, Palmer DS, Pedersen CB, Pedersen MG, Poterba T, Poulsen JB, Ripke S, Schork AJ, Thompson WK, Turley P, Walters RK. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020 Feb;180(3):568-584.e23.
8. Werling DM, Geschwind DH. Sex differences in autism spectrum disorders: Curr Opin Neurol. 2013 Apr;26(2):146–53.
9. Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, Won H, Van Bakel H, Varghese M, Wang Y, Shieh AW, Haney J, Parhami S, Belmont J, Kim M, Moran Losada P, Khan Z, Mleczko J, Xia Y, Dai R, Wang D, Yang YT, Xu M, Fish K, Hof PR, Warrell J, Fitzgerald D, White K, Jaffe AE, PsychENCODE Consortium, Peters MA, Gerstein M, Liu C, Iakoucheva LM, Pinto D, Geschwind DH, Ashley-Koch AE, Crawford GE, Garrett ME, Song L, Safi A, Johnson GD, Wray GA, Reddy TE, Goes FS, Zandi P, Bryois J, Jaffe AE, Price AJ, Ivanov NA, Collado-Torres L, Hyde TM, Burke EE, Kleiman JE, Tao R, Shin JH, Akbarian S, Girdhar K, Jiang Y, Kundakovic M, Brown L, Kassim BS, Park RB, Wiseman JR, Zharovsky E, Jacobov R, Devillers O, Flatow E, Hoffman GE, Lipska BK, Lewis DA, Haroutunian V, Hahn CG, Charney AW, Dracheva S, Kozlenkov A, Belmont J, DelValle D, Francoeur N, Hadjimichael E, Pinto D, Van Bakel H, Roussos P, Fullard JF, Bendl J, Hauberg ME, Mangravite LM, Peters MA, Chae Y, Peng J, Niu M, Wang X, Webster MJ, Beach TG, Chen C, Jiang Y, Dai R, Shieh AW, Liu C, Grennan KS, Xia Y, Vadukapuram R, Wang Y, Fitzgerald D, Cheng L, Brown M, Brown M, Brunetti T, Goodman T, Alsayed M, Gandal MJ, Geschwind DH, Won H, Polioudakis D, Wamsley B, Yin J, Hadzic T, De La Torre Ubieta L, Swarup V, Sanders SJ, State MW, Werling DM, An JY, Sheppard B, Willsey AJ, White KP, Ray M, Giase G, Kefi A, Mattei E, Purcaro M, Weng Z, Moore J, Pratt H, Huey J, Borrman T, Sullivan PF, Giusti-Rodriguez P, Kim Y, Sullivan P, Szatkiewicz J, Rhie SK, Armoskus C, Camarena A, Farnham PJ, Spitsyna VN, Witt H, Schreiner S, Evgrafov OV, Knowles JA, Gerstein M, Liu S, Wang D, Navarro FCP, Warrell J, Clarke D,

 doi:10.20944/preprints202511.1997.v1

Emani PS, Gu M, Shi X, Xu M, Yang YT, Kitchen RR, Gürsoy G, Zhang J, Carlyle BC, Nairn AC, Li M, Pochareddy S, Sestan N, Skarica M, Li Z, Sousa AMM, Santpere G, Choi J, Zhu Y, Gao T, Miller DJ, Cherskov A, Yang M, Amiri A, Coppola G, Mariani J, Scuderi S, Szekely A, Vaccarino FM, Wu F, Weissman S, Roychowdhury T, Abyzov A. Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. Science. 2018 Dec 14;362(6420):eaat8127.

10. Sanders SJ. Next-Generation Sequencing in Autism Spectrum Disorder. Cold Spring Harb Perspect Med. 2019 Aug;9(8):a026872.
11. Patterson PH. Maternal infection and immune involvement in autism. Trends Mol Med. 2011 Jul;17(7):389–94.
12. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Genome Aggregation Database Consortium, Aguilar Salinas CA, Ahmad T, Albert CM, Ardissino D, Atzmon G, Barnard J, Beaugerie L, Benjamin EJ, Boehnke M, Bonnycastle LL, Bottinger EP, Bowden DW, Bown MJ, Chambers JC, Chan JC, Chasman D, Cho J, Chung MK, Cohen B, Correa A, Dabelea D, Daly MJ, Darbar D, Duggirala R, Dupuis J, Ellinor PT, Elosua R, Erdmann J, Esko T, Färkkilä M, Florez J, Franke A, Getz G, Glaser B, Glatt SJ, Goldstein D, Gonzalez C, Groop L, Haiman C, Hanis C, Harms M, Hiltunen M, Holi MM, Hultman CM, Kallela M, Kaprio J, Kathiresan S, Kim BJ, Kim YJ, Kirov G, Kooner J, Koskinen S, Krumholz HM, Kugathasan S, Kwak SH, Laakso M, Lehtimäki T, Loos RJF, Lubitz SA, Ma RCW, MacArthur DG, Marrugat J, Mattila KM, McCarroll S, McCarthy MI, McGovern D, McPherson R, Meigs JB, Melander O, Metspalu A, Neale BM, Nilsson PM, O'Donovan MC, Ongur D, Orozco L, Owen MJ, Palmer CNA, Palotie A, Park KS, Pato C, Pulver AE, Rahman N, Remes AM, Rioux JD, Ripatti S, Roden DM, Saleheen D, Salomaa V, Samani NJ, Scharf J, Schunkert H, Shoemaker MB, Sklar P, Soininen H, Sokol H, Spector T, Sullivan PF, Suvisaari J, Tai ES, Teo YY, Tiinamaija T, Tsuang M, Turner D, Tusie-Luna T, Vartiainen E, Vawter MP, Ware JS, Watkins H, Weersma RK, Wessman M, Wilson JG, Xavier RJ, Neale BM, Daly MJ, MacArthur DG. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020 May 28;581(7809):434–43.
13. Talkowski ME, Minikel EV, Gusella JF. Autism Spectrum Disorder Genetics: Diverse Genes with Diverse Clinical Outcomes. Harv Rev Psychiatry. 2014 Mar;22(2):65–75.
14. Abrahams BS, Arking DE, Campbell DB, Mefford HC, Morrow EM, Weiss LA, Menashe I, Wadkins T, Banerjee-Basu S, Packer A. SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol Autism. 2013 Dec;4(1):36.
15. Geschwind DH. Genetics of autism spectrum disorders. Trends Cogn Sci. 2011 Sep;15(9):409–16.
16. Premeaux TA, Mediouni S, Leda A, Furler RL, Valente ST, Fine HA, Nixon DF, Ndhlovu LC. Next-Generation Human Cerebral Organoids as Powerful Tools To Advance NeuroHIV Research. Burdo TH, Prasad VR, editors. mBio. 2021 Aug 31;12(4):e00680-21.
17. Zeidan J, Fombonne E, Scorah J, Ibrahim A, Durkin MS, Saxena S, Yusuf A, Shih A, Elsabbagh M. Global prevalence of autism: A systematic review update. Autism Res. 2022 May;15(5):778–90.
18. Buescher AVS, Cidav Z, Knapp M, Mandell DS. Costs of Autism Spectrum Disorders in the United Kingdom and the United States. JAMA Pediatr. 2014 Aug 1;168(8):721.
19. Li C, Fleck JS, Martins-Costa C, Burkard TR, Themann J, Stuempflen M, Peer AM, Vertesy Á, Littleboy JB, Esk C, Elling U, Kasprian G, Corsini NS, Treutlein B, Knoblich JA. Single-cell brain organoid screening identifies developmental defects in autism. Nature. 2023 Sep 14;621(7978):373–80.
20. Colizzi M, Lasalvia A, Ruggeri M. Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care? Int J Ment Health Syst. 2020 Dec;14(1):23.
21. Nakanishi M, Anderson MP, Takumi T. Recent genetic and functional insights in autism spectrum disorder. Curr Opin Neurol. 2019 Aug;32(4):627–34.

 doi:10.20944/preprints202511.1997.v1

22. Alonso-Gonzalez A, Rodriguez-Fontenla C, Carracedo A. De novo Mutations (DNMs) in Autism Spectrum Disorder (ASD): Pathway and Network Analysis. Front Genet. 2018 Sep 21;9:406.
23. Vysotskiy M, Zhong X, Miller-Fleming TW, Zhou D, Autism Working Group of the Psychiatric Genomics Consortium^, Bipolar Disorder Working Group of the Psychiatric Genomics Consortium^, Schizophrenia Working Group of the Psychiatric Genomics Consortium^, Cox NJ, Weiss LA. Integration of genetic, transcriptomic, and clinical data provides insight into 16p11.2 and 22q11.2 CNV genes. Genome Med. 2021 Dec;13(1):172.
24. More RP, Warrier V, Brunel H, Buckingham C, Smith P, Allison C, Holt R, Bradshaw CR, Baron-Cohen S. Identifying rare genetic variants in 21 highly multiplex autism families: the role of diagnosis and autistic traits. Mol Psychiatry. 2023 May;28(5):2148–57.
25. Yousaf A, Waltes R, Haslinger D, Klauck SM, Duketis E, Sachse M, Voran A, Biscaldi M, Schulte-Rüther M, Cichon S, Nöthen M, Ackermann J, Koch I, Freitag CM, Chiocchetti AG. Quantitative genome-wide association study of six phenotypic subdomains identifies novel genome-wide significant variants in autism spectrum disorder. Transl Psychiatry. 2020 Jul 5;10(1):215.
26. iPSYCH-Broad Autism Group, Psychiatric Genomics Consortium Autism Group, Weiner DJ, Wigdor EM, Ripke S, Walters RK, Kosmicki JA, Grove J, Samocha KE, Goldstein JI, Okbay A, Bybjerg-Grauholm J, Werge T, Hougaard DM, Taylor J, Skuse D, Devlin B, Anney R, Sanders SJ, Bishop S, Mortensen PB, Børglum AD, Smith GD, Daly MJ, Robinson EB. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. Nat Genet. 2017 Jul;49(7):978–85.
27. Shaw KA, Williams S, Patrick ME, Valencia-Prado M, Durkin MS, Howerton EM, Ladd-Acosta CM, Pas ET, Bakian AV, Bartholomew P, Nieves-Muñoz N, Sidwell K, Alford A, Bilder DA, DiRienzo M, Fitzgerald RT, Furnier SM, Hudson AE, Pokoski OM, Shea L, Tinker SC, Warren Z, Zahorodny W, Agosto-Rosa H, Anbar J, Chavez KY, Esler A, Forkner A, Grzybowski A, Agib AH, Hallas L, Lopez M, Magaña S, Nguyen RHN, Parker J, Pierce K, Protho T, Torres H, Vanegas SB, Vehorn A, Zhang M, Andrews J, Greer F, Hall-Lande J, McArthur D, Mitamura M, Montes AJ, Pettygrove S, Shenouda J, Skowyra C, Washington A, Maenner MJ. Prevalence and Early Identification of Autism Spectrum Disorder Among Children Aged 4 and 8 Years — Autism and Developmental Disabilities Monitoring Network, 16 Sites, United States, 2022. MMWR Surveill Summ. 2025 Apr 17;74(2):1–22.
28. Gudmundsson S, Singer-Berk M, Watts NA, Phu W, Goodrich JK, Solomonson M, Genome Aggregation Database Consortium, Rehm HL, MacArthur DG, O'Donnell-Luria A. Variant interpretation using population databases: Lessons from gnomAD. Hum Mutat. 2022 Aug;43(8):1012–30.
29. Austin-Tse CA, Jobanputra V, Perry DL, Bick D, Taft RJ, Venner E, Gibbs RA, Young T, Barnett S, Belmont JW, Boczek N, Chowdhury S, Ellsworth KA, Guha S, Kulkarni S, Marcou C, Meng L, Murdock DR, Rehman AU, Spiteri E, Thomas-Wilson A, Kearney HM, Rehm HL, Medical Genome Initiative*. Best practices for the interpretation and reporting of clinical whole genome sequencing. Npj Genomic Med. 2022 Apr 8;7(1):27.
30. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, Ebbert A, Riley ZL, Royall JJ, Aiona K, Arnold JM, Bennet C, Bertagnolli D, Brouner K, Butler S, Caldejon S, Carey A, Cuhaciyan C, Dalley RA, Dee N, Dolbeare TA, Facer BAC, Feng D, Fliss TP, Gee G, Goldy J, Gourley L, Gregor BW, Gu G, Howard RE, Jochim JM, Kuan CL, Lau C, Lee CK, Lee F, Lemon TA, Lesnar P, McMurray B, Mastan N, Mosqueda N, Naluai-Cecchini T, Ngo NK, Nyhus J, Oldre A, Olson E, Parente J, Parker PD, Parry SE, Stevens A, Pletikos M, Reding M, Roll K, Sandman D, Sarreal M, Shapouri S, Shapovalova NV, Shen EH, Sjoquist N, Slaughterbeck CR, Smith M, Sodt AJ, Williams D, Zöllei L, Fischl B, Gerstein MB, Geschwind DH, Glass IA, Hawrylycz MJ, Hevner RF, Huang H, Jones AR, Knowles JA, Levitt P, Phillips JW, Šestan N, Wohnoutka P, Dang C, Bernard A, Hohmann JG, Lein ES. Transcriptional landscape of the prenatal human brain. Nature. 2014 Apr;508(7495):199–206.
31. Woodward AA, Urbanowicz RJ, Naj AC, Moore JH. Genetic heterogeneity: Challenges, impacts, and methods through an associative lens. Genet Epidemiol. 2022 Dec;46(8):555–71.

 doi:10.20944/preprints202511.1997.v1

32. Fiksinski AM, Schneider M, Zinkstok J, Baribeau D, Chawner SJRA, Vorstman JAS. Neurodevelopmental Trajectories and Psychiatric Morbidity: Lessons Learned From the 22q11.2 Deletion Syndrome. Curr Psychiatry Rep. 2021 Mar;23(3):13.
33. Kruth KA, Grisolano TM, Ahern CA, Williams AJ. SCN2A channelopathies in the autism spectrum of neuropsychiatric disorders: a role for pluripotent stem cells? Mol Autism. 2020 Dec;11(1):23.
34. Tang D, Freudenberg J, Dahl A. Factorizing polygenic epistasis improves prediction and uncovers biological pathways in complex traits. Am J Hum Genet. 2023 Nov;110(11):1875–87.
35. Ardhanareeswaran K, Mariani J, Coppola G, Abyzov A, Vaccarino FM. Human induced pluripotent stem cells for modelling neurodevelopmental disorders. Nat Rev Neurol. 2017 May;13(5):265–78.
36. Nehme R, Barrett LE. Using human pluripotent stem cell models to study autism in the era of big data. Mol Autism. 2020 Dec;11(1):21.
37. Nomura T. Interneuron Dysfunction and Inhibitory Deficits in Autism and Fragile X Syndrome. Cells. 2021 Oct 1;10(10):2610.
38. Xu J, Wen Z. Brain Organoids: Studying Human Brain Development and Diseases in a Dish. Li X, editor. Stem Cells Int. 2021 Sep 9;2021:1–16.
39. Kshirsagar A, Mnatsakanyan H, Kulkarni S, Guo J, Cheng K, Ofria LD, Bohra O, Sagar R, Mahairaki V, Badr CE, Kathuria A. Multi-Region Brain Organoids Integrating Cerebral, Mid-Hindbrain, and Endothelial Systems. Adv Sci. 2025 Sep;12(33):e03768.
40. Zhao X, Xu Z, Xiao L, Shi T, Xiao H, Wang Y, Li Y, Xue F, Zeng W. Review on the Vascularization of Organoids and Organoids-on-a-Chip. Front Bioeng Biotechnol. 2021 Apr 12;9:637048.
41. Cha JH, Kim K, Cho IJ. Beyond Structure: Next-Generation Electrophysiological Platforms for Functional Brain Organoids. Int J Stem Cells. 2025 Aug 30;18(3):215–36.
42. Demmings MD, Da Silva Chagas L, Traetta ME, Rodrigues RS, Acutain MF, Barykin E, Datusalia AK, German-Castelan L, Mattera VS, Mazengenya P, Skoug C, Umemori H. (Re)building the nervous system: A review of neuron–glia interactions from development to disease. J Neurochem. 2025 Jan;169(1):e16258.
43. Da Silva GG, Sacomani DP, De Carvalho BG, Porcionatto MA, Gobbi A, Lima RS, De La Torre LG. Microfluidic Systems to Mimic the Blood–Brain Barrier: from Market to Engineering Challenges and Perspectives. ACS Biomater Sci Eng. 2025 Jul 14;11(7):3789–815.
44. Sampaio AR, Maia RF, Ciardulli MC, Santos HA, Sarmento B. Organ-on-chip platforms for nanoparticle toxicity and efficacy assessment: Advancing beyond traditional *in vitro* and *in vivo* models. Mater Today Bio. 2025 Aug;33:102053.
45. Meng X, Yao D, Imaizumi K, Chen X, Kelley KW, Reis N, Thete MV, Arjun McKinney A, Kulkarni S, Panagiotakos G, Bassik MC, Pașca SP. Assembloid CRISPR screens reveal impact of disease genes in human neurodevelopment. Nature. 2023 Oct 12;622(7982):359–66.
46. Volpato V, Webber C. Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. Dis Model Mech. 2020 Jan 1;13(1):dmm042317.
47. Sanjana NE. Genome-scale CRISPR pooled screens. Anal Biochem. 2017 Sep;532:95–9.
48. Alda-Catalinas C, Eckersley-Maslin MA, Reik W. Pooled CRISPR-activation screening coupled with single-cell RNA-seq in mouse embryonic stem cells. STAR Protoc. 2021 Jun;2(2):100426.
49. Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, Marjanovic ND, Dionne D, Burks T, Raychowdhury R, Adamson B, Norman TM, Lander ES, Weissman JS, Friedman N, Regev A. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. Cell. 2016 Dec;167(7):1853-1866.e17.
50. Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. Nat Med. 2018 Jul;24(7):927–30.
51. Binan L, Jiang A, Danquah SA, Valakh V, Simonton B, Bezney J, Manguso RT, Yates KB, Nehme R, Cleary B, Farhi SL. Simultaneous CRISPR screening and spatial transcriptomics reveal intracellular, intercellular, and functional transcriptional circuits. Cell. 2025 Apr;188(8):2141-2158.e18.
52. Prilutsky D, Palmer NP, Smedemark-Margulies N, Schlaeger TM, Margulies DM, Kohane IS. iPSC-derived neurons as a higher-throughput readout for autism: promises and pitfalls. Trends Mol Med. 2014 Feb;20(2):91–104.

 doi:10.20944/preprints202511.1997.v1

53. Ferreira Da Silva J, Salic S, Wiedner M, Datlinger P, Essletzbichler P, Hanzl A, Superti-Furga G, Bock C, Winter G, Loizou JI. Genome-scale CRISPR screens are efficient in non-homologous end-joining deficient cells. Sci Rep. 2019 Oct 31;9(1):15751.
54. Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford GE, Gersbach CA. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat Methods. 2015 Dec;12(12):1143–9.
55. Kluesner MG, Lahr WS, Lonetree C lin, Smeester BA, Qiu X, Slipek NJ, Claudio Vázquez PN, Pitzen SP, Pomeroy EJ, Vignes MJ, Lee SC, Bingea SP, Andrew AA, Webber BR, Moriarity BS. CRISPR-Cas9 cytidine and adenosine base editing of splice-sites mediates highly-efficient disruption of proteins in primary and immortalized cells. Nat Commun. 2021 Apr 23;12(1):2437.
56. Kantor A, McClements M, MacLaren R. CRISPR-Cas9 DNA Base-Editing and Prime-Editing. Int J Mol Sci. 2020 Aug 28;21(17):6240.
57. Jeffries MA. Epigenetic editing: How cutting-edge targeted epigenetic modification might provide novel avenues for autoimmune disease therapy. Clin Immunol. 2018 Nov;196:49–58.
58. Zhu G, Zhou X, Wen M, Qiao J, Li G, Yao Y. CRISPR–Cas13: Pioneering RNA Editing for Nucleic Acid Therapeutics. BioDesign Res. 2024;6:0041.
59. DeBoever C, Li H, Jakubosky D, Benaglio P, Reyna J, Olson KM, Huang H, Biggs W, Sandoval E, D'Antonio M, Jepsen K, Matsui H, Arias A, Ren B, Nariai N, Smith EN, D'Antonio-Chronowska A, Farley EK, Frazer KA. Large-Scale Profiling Reveals the Influence of Genetic Variation on Gene Expression in Human Induced Pluripotent Stem Cells. Cell Stem Cell. 2017 Apr;20(4):533-546.e7.
60. Wang X, Lalli M, Thopte U, Buxbaum JD. A scalable, high-throughput neural development platform identifies shared impact of ASD genes on cell fate and differentiation [Internet]. 2024 [cited 2025 Sep 6]. Available from: http://biorxiv.org/lookup/doi/10.1101/2024.09.25.614184
61. Navarro M, Simpson TI. SFARI Genes and where to find them; classification modelling to identify genes associated with Autism Spectrum Disorder from RNA-seq data [Internet]. 2021 [cited 2025 Sep 6]. Available from: http://biorxiv.org/lookup/doi/10.1101/2021.01.29.428754
62. Shin T, Song JHT, Kosicki M, Kenny C, Beck SG, Kelley L, Qian X, Bonacina J, Papandile F, Antony I, Gonzalez D, Scotellaro J, Bushinsky EM, Andersen RE, Maury E, Pennacchio LA, Doan RN, Walsh CA. Rare variation in noncoding regions with evolutionary signatures contributes to autism spectrum disorder risk [Internet]. 2023 [cited 2025 Sep 6]. Available from: http://medrxiv.org/lookup/doi/10.1101/2023.09.19.23295780
63. Holland KD, Bouley TM, Horn PS. Comparison and optimization of *in silico* algorithms for predicting the pathogenicity of sodium channel variants in epilepsy. Epilepsia. 2017 Jul;58(7):1190–8.
64. Boggess SC, Gandhi V, Tsai MC, Marzette E, Teyssier N, Chou JYY, Hu X, Cramer A, Yadanar L, Shroff K, Jeong CG, Eidenschenk C, Hanson JE, Tian R, Kampmann M. A Massively Parallel CRISPR-Based Screening Platform for Modifiers of Neuronal Activity [Internet]. 2024 [cited 2025 Sep 6]. Available from: http://biorxiv.org/lookup/doi/10.1101/2024.02.28.582546
65. Read A, Gao S, Batchelor E, Luo J. Flexible CRISPR library construction using parallel oligonucleotide retrieval. Nucleic Acids Res. 2017 Jun 20;45(11):e101–e101.
66. Hutcheson RL, Hayes M, Sugden B. Optimal LentiCRISPR-Based System for Sequential CRISPR/Cas9 Screens. ACS Synth Biol. 2022 Jul 15;11(7):2259–66.
67. Merten OW, Hebben M, Bovolenta C. Production of lentiviral vectors. Mol Ther - Methods Clin Dev. 2016;3:16017.
68. Fuentes CM, Schaffer DV. Adeno-associated virus-mediated delivery of CRISPR-Cas9 for genome editing in the central nervous system. Curr Opin Biomed Eng. 2018 Sep;7:33–41.
69. Matsoukas IG. Prime Editing: Genome Editing for Rare Genetic Diseases Without Double-Strand Breaks or Donor DNA. Front Genet. 2020 Jun 9;11:528.
70. Verstraelen P, Van Dyck M, Verschuuren M, Kashikar ND, Nuydens R, Timmermans JP, De Vos WH. Image-Based Profiling of Synaptic Connectivity in Primary Neuronal Cell Culture. Front Neurosci. 2018 Jun 26;12:389.

 doi:10.20944/preprints202511.1997.v1

71. Pogacar Z, Groot K, Jochems F, Dos Santos Dias M, Mulero-Sánchez A, Morris B, Roosen M, Wardak L, De Conti G, Velds A, Lieftink C, Thijssen B, Beijersbergen RL, Bernards R, Leite De Oliveira R. Genetic and compound screens uncover factors modulating cancer cell response to indisulam. Life Sci Alliance. 2022 Sep;5(9):e202101348.
72. Feldman D, Funk L, Le A, Carlson RJ, Leiken MD, Tsai F, Soong B, Singh A, Blainey PC. Pooled genetic perturbation screens with image-based phenotypes. Nat Protoc. 2022 Feb;17(2):476–512.
73. Saunders RA, Allen WE, Pan X, Sandhu J, Lu J, Lau TK, Smolyar K, Sullivan ZA, Dulac C, Weissman JS, Zhuang X. Perturb-Multimodal: A platform for pooled genetic screens with imaging and sequencing in intact mammalian tissue. Cell. 2025 Aug;188(17):4790-4809.e22.
74. Cerneckis J, Cai H, Shi Y. Induced pluripotent stem cells (iPSCs): molecular mechanisms of induction and applications. Signal Transduct Target Ther. 2024 Apr 26;9(1):112.
75. Dueck H, Khaladkar M, Kim TK, Spaethling JM, Francis C, Suresh S, Fisher SA, Seale P, Beck SG, Bartfai T, Kuhn B, Eberwine J, Kim J. Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation. Genome Biol. 2015 Jun 9;16(1):122.
76. Imaizumi K, Sone T, Ibata K, Fujimori K, Yuzaki M, Akamatsu W, Okano H. Controlling the Regional Identity of hPSC-Derived Neurons to Uncover Neuronal Subtype Specificity of Neurological Disease Phenotypes. Stem Cell Rep. 2015 Dec;5(6):1010–22.
77. Liu J, Tran V, Vemuri VNP, Byrne A, Borja M, Kim YJ, Agarwal S, Wang R, Awayan K, Murti A, Taychameekiatchai A, Wang B, Emanuel G, He J, Haliburton J, Oliveira Pisco A, Neff NF. Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing. Life Sci Alliance. 2023 Jan;6(1):e202201701.
78. Ishahak M, Han RH, Annamalai D, Woodiwiss T, McCornack C, Cleary RT, DeSouza PA, Qu X, Dahiya S, Kim AH, Millman JR. Genetically Engineered Brain Organoids Recapitulate Spatial and Developmental States of Glioblastoma Progression. Adv Sci. 2025 Mar;12(10):2410110.
79. Yao D, Binan L, Bezney J, Simonton B, Freedman J, Frangieh CJ, Dey K, Geiger-Schuller K, Eraslan B, Gusev A, Regev A, Cleary B. Compressed Perturb-seq: highly efficient screens for regulatory circuits using random composite perturbations [Internet]. 2023 [cited 2025 Sep 6]. Available from: http://biorxiv.org/lookup/doi/10.1101/2023.01.23.525200
80. Broussard GJ, Liang R, Tian L. Monitoring activity in neural circuits with genetically encoded indicators. Front Mol Neurosci [Internet]. 2014 Dec 5 [cited 2025 Sep 6];7. Available from: http://journal.frontiersin.org/article/10.3389/fnmol.2014.00097/abstract
81. Ahmed M, Muffat J, Li Y. Understanding neural development and diseases using CRISPR screens in human pluripotent stem cell-derived cultures. Front Cell Dev Biol. 2023 Apr 10;11:1158373.
82. Larsch J, Ventimiglia D, Bargmann CI, Albrecht DR. High-throughput imaging of neuronal activity in *Caenorhabditis elegans*. Proc Natl Acad Sci [Internet]. 2013 Nov 5 [cited 2025 Sep 6];110(45). Available from: https://pnas.org/doi/full/10.1073/pnas.1318325110
83. Morris JA, Sun JS, Sanjana NE. Next-generation forward genetic screens: uniting high-throughput perturbations with single-cell analysis. Trends Genet. 2024 Feb;40(2):118–33.
84. Kavaliauskaite G, Madsen JGS. Automatic quality control of single-cell and single-nucleus RNA-seq using valiDrops. NAR Genomics Bioinforma. 2023 Oct 11;5(4):lqad101.
85. Manz KM, Siemann JK, McMahon DG, Grueter BA. Patch-clamp and multi-electrode array electrophysiological analysis in acute mouse brain slices. STAR Protoc. 2021 Jun;2(2):100442.
86. Lu C, Wei Y, Abbas M, Agula H, Wang E, Meng Z, Zhang R. Application of Single-Cell Assay for Transposase-Accessible Chromatin with High Throughput Sequencing in Plant Science: Advances, Technical Challenges, and Prospects. Int J Mol Sci. 2024 Jan 25;25(3):1479.
87. Ogbeide S, Giannese F, Mincarelli L, Macaulay IC. Into the multiverse: advances in single-cell multiomic profiling. Trends Genet. 2022 Aug;38(8):831–43.
88. D'haene E, Vergult S. Interpreting the impact of noncoding structural variation in neurodevelopmental disorders. Genet Med. 2021 Jan;23(1):34–46.
89. Baysoy A, Tian X, Zhang F, Renauer P, Bai Z, Shi H, Li H, Tao B, Yang M, Enninful A, Gao F, Wang G, Zhang W, Tran T, Patterson NH, Bao S, Dong C, Xin S, Zhong M, Rankin S, Guy C, Wang Y, Connelly JP,

 doi:10.20944/preprints202511.1997.v1

Pruett-Miller SM, Chi H, Chen S, Fan R. Spatially Resolved *in vivo* CRISPR Screen Sequencing via Perturb-DBiT [Internet]. 2024 [cited 2025 Sep 6]. Available from: http://biorxiv.org/lookup/doi/10.1101/2024.11.18.624106

90. Song L, Chen W, Hou J, Guo M, Yang J. Spatially resolved mapping of cells associated with human complex traits. Nature. 2025 May 22;641(8064):932–41.
91. Qin W, Cho KF, Cavanagh PE, Ting AY. Deciphering molecular interactions by proximity labeling. Nat Methods. 2021 Feb;18(2):133–43.
92. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L. Integrated Genomic and Proteomic Analyses of a Systematically Perturbed Metabolic Network. Science. 2001 May 4;292(5518):929–34.
93. Hempel CM, Sivula M, Levenson JM, Rose DM, Li B, Sirianni AC, Xia E, Ryan TA, Gerber DJ, Cottrell JR. A System for Performing High Throughput Assays of Synaptic Function. Dryer SE, editor. PLoS ONE. 2011 Oct 5;6(10):e25999.
94. Vahsen BF, Gray E, Candalija A, Cramb KML, Scaber J, Dafinca R, Katsikoudi A, Xu Y, Farrimond L, Wade-Martins R, James WS, Turner MR, Cowley SA, Talbot K. Human iPSC co-culture model to investigate the interaction between microglia and motor neurons. Sci Rep. 2022 Jul 23;12(1):12606.
95. Wu X, Yang X, Dai Y, Zhao Z, Zhu J, Guo H, Yang R. Single-cell sequencing to multi-omics: technologies and applications. Biomark Res. 2024 Sep 27;12(1):110.
96. Cai L, Ma X, Ma J. Integrating scRNA-seq and scATAC-seq with inter-type attention heterogeneous graph neural networks. Brief Bioinform. 2024 Nov 22;26(1):bbae711.
97. Schraivogel D, Steinmetz LM, Parts L. Pooled Genome-Scale CRISPR Screens in Single Cells. Annu Rev Genet. 2023 Nov 27;57(1):223–44.
98. Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J, Schuster LC, Kuchler A, Alpar D, Bock C. Pooled CRISPR screening with single-cell transcriptome readout. Nat Methods. 2017 Mar;14(3):297–301.
99. Rosati D, Palmieri M, Brunelli G, Morrione A, Iannelli F, Frullanti E, Giordano A. Differential gene expression analysis pipelines and bioinformatic tools for the identification of specific biomarkers: A review. Comput Struct Biotechnol J. 2024 Dec;23:1154–68.
100. Marks PW, Witten CM, Califf RM. Clarifying Stem-Cell Therapy's Benefits and Risks. N Engl J Med. 2017 Mar 16;376(11):1007–9.
101. Gilis J, Perin L, Malfait M, Van Den Berge K, Assefa AT, Verbist B, Risso D, Clement L. Differential detection workflows for multi-sample single-cell RNA-seq data [Internet]. 2023 [cited 2025 Sep 6]. Available from: http://biorxiv.org/lookup/doi/10.1101/2023.12.17.572043
102. Lun ATL, Marioni JC. Overcoming confounding plate effects in differential expression analyses of single-cell RNA-seq data. Biostatistics. 2017 Jul 1;18(3):451–64.
103. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014 Apr;32(4):381–6.
104. Guo Y, Zhao X. CRISPR-based genetic screens in human pluripotent stem cells derived neurons and brain organoids. Cell Tissue Res. 2025 Jan;399(1):1–8.
105. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008 Dec;9(1):559.
106. Kiselev VY, Andrews TS, Hemberg M. Challenges in unsupervised clustering of single-cell RNA-seq data. Nat Rev Genet. 2019 May;20(5):273–82.
107. Hie B, Bryson B, Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nat Biotechnol. 2019 Jun;37(6):685–91.
108. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019 Apr 29;10(1):1930.
109. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, Bodenmiller B, Campbell P, Carninci P, Clatworthy M, Clevers H, Deplancke B, Dunham I, Eberwine J, Eils R, Enard W, Farmer A, Fugger L, Göttgens B, Hacohen N, Haniffa M, Hemberg M, Kim S, Klenerman P, Kriegstein A, Lein E, Linnarsson S,

 doi:10.20944/preprints202511.1997.v1

Lundberg E, Lundeberg J, Majumder P, Marioni JC, Merad M, Mhlanga M, Nawijn M, Netea M, Nolan G, Pe'er D, Phillipakis A, Ponting CP, Quake S, Reik W, Rozenblatt-Rosen O, Sanes J, Satija R, Schumacher TN, Shalek A, Shapiro E, Sharma P, Shin JW, Stegle O, Stratton M, Stubbington MJT, Theis FJ, Uhlen M, Van Oudenaarden A, Wagner A, Watt F, Weissman J, Wold B, Xavier R, Yosef N, Human Cell Atlas Meeting Participants. The Human Cell Atlas. eLife. 2017 Dec 5;6:e27041.

110. Jin Y, Marquardt S. Dual sgRNA-based Targeted Deletion of Large Genomic Regions and Isolation of Heritable Cas9-free Mutants in Arabidopsis. BIO-Protoc [Internet]. 2020 [cited 2025 Sep 6];10(20). Available from: https://bio-protocol.org/e3796
111. Wang YX, Feige P, Brun CE, Hekmatnejad B, Dumont NA, Renaud JM, Faulkes S, Guindon DE, Rudnicki MA. EGFR-Aurka Signaling Rescues Polarity and Regeneration Defects in Dystrophin-Deficient Muscle Stem Cells by Increasing Asymmetric Divisions. Cell Stem Cell. 2019 Mar;24(3):419-432.e6.
112. Lancaster MA, Renner M, Martin CA, Wenzel D, Bicknell LS, Hurles ME, Homfray T, Penninger JM, Jackson AP, Knoblich JA. Cerebral organoids model human brain development and microcephaly. Nature. 2013 Sep 19;501(7467):373–9.
113. Lovell-Badge R, Anthony E, Barker RA, Bubela T, Brivanlou AH, Carpenter M, Charo RA, Clark A, Clayton E, Cong Y, Daley GQ, Fu J, Fujita M, Greenfield A, Goldman SA, Hill L, Hyun I, Isasi R, Kahn J, Kato K, Kim JS, Kimmelman J, Knoblich JA, Mathews D, Montserrat N, Mosher J, Munsie M, Nakauchi H, Naldini L, Naughton G, Niakan K, Ogbogu U, Pedersen R, Rivron N, Rooke H, Rossant J, Round J, Saitou M, Sipp D, Steffann J, Sugarman J, Surani A, Takahashi J, Tang F, Turner L, Zettler PJ, Zhai X. ISSCR Guidelines for Stem Cell Research and Clinical Translation: The 2021 update. Stem Cell Rep. 2021 Jun;16(6):1398–408.
114. Ludwig TE, Andrews PW, Barbaric I, Benvenisty N, Bhattacharyya A, Crook JM, Daheron LM, Draper JS, Healy LE, Huch M, Inamdar MS, Jensen KB, Kurtz A, Lancaster MA, Liberali P, Lutolf MP, Mummery CL, Pera MF, Sato Y, Shimasaki N, Smith AG, Song J, Spits C, Stacey G, Wells CA, Zhao T, Mosher JT. ISSCR standards for the use of human stem cells in basic research. Stem Cell Rep. 2023 Sep;18(9):1744–52.
115. Hu X, Jing M, Wang Y, Liu Y, Hua Y. Functional analysis of a novel de novo SCN2A variant in a patient with seizures refractory to oxcarbazepine. Front Mol Neurosci. 2023 Apr 20;16:1159649.
116. Jiangping H, Lihui L, Jiekai C, Center for Cell Lineage and Atlas (CCLA), Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory), Guangzhou 510320, China, Key Laboratory of Regenerative Biology of the Chinese Academy of Sciences and Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China. Practical bioinformatics pipelines for single-cell RNA-seq data analysis. Biophys Rep. 2022;8(3):158–69.
117. Macaulay IC, Voet T. Single Cell Genomics: Advances and Future Perspectives. Maizels N, editor. PLoS Genet. 2014 Jan 30;10(1):e1004126.
118. Vaidyanathan S, Salahudeen AA, Sellers ZM, Bravo DT, Choi SS, Batish A, Le W, Baik R, De La O S, Kaushik MP, Galper N, Lee CM, Teran CA, Yoo JH, Bao G, Chang EH, Patel ZM, Hwang PH, Wine JJ, Milla CE, Desai TJ, Nayak JV, Kuo CJ, Porteus MH. High-Efficiency, Selection-free Gene Repair in Airway Stem Cells from Cystic Fibrosis Patients Rescues CFTR Function in Differentiated Epithelia. Cell Stem Cell. 2020 Feb;26(2):161-171.e4.
119. Menyhart O, Weltz B, Győrffy B. MultipleTesting.com: A tool for life science researchers for multiple hypothesis testing correction. Delcea C, editor. PLOS ONE. 2021 Jun 9;16(6):e0245824.
120. Hens K, Noens I, Peeters H, Steyaert J. The ethics of patenting autism genes. Nat Rev Genet. 2018 May;19(5):247–8.
121. Qu H, Fang X. A Brief Review on the Human Encyclopedia of DNA Elements (ENCODE) Project. Genomics Proteomics Bioinformatics. 2013 Jun 1;11(3):135–41.
122. Arpi MNT, Simpson TI. SFARI genes and where to find them; modelling Autism Spectrum Disorder specific gene expression dysregulation with RNA-seq data. Sci Rep. 2022 Jun 16;12(1):10158.
123. Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, Theriault K, Kommineni S, Chen J, Sondey M, Ye C, Randhawa R, Kulkarni T, Yang Z, McAllister G, Russ C, Reece-Hoyes J, Forrester W, Hoffman GR, Dolmetsch R, Kaykas A. p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. Nat Med. 2018 Jul;24(7):939–46.

 doi:10.20944/preprints202511.1997.v1