OXFORD

TRENDS AND PERSPECTIVES

Strategies to Advance Drug Discovery in Rare Monogenic Intellectual Disability Syndromes

Nuwan C. Hettige, Karla Manzano-Vargas, Malvin Jefri, Carl Ernst

Department of Human Genetics, McGill University, Montreal, QC, Canada (Mr Hettige, Ms Manzano-Vargas, and Dr Ernst); McGill Group for Suicide Studies, Douglas Mental Health University Institute, Montreal, QC, Canada (Mr Hettige, Ms Manzano-Vargas, Mr Jefri, and Dr Ernst); Department of Psychiatry (Dr Ernst), and Integrated Program in Neuroscience Mr Jefri and Dr Ernst), McGill University, Montreal, QC, Canada.

N.C.H. and K.M.-V. contributed equally to the article.

Correspondence: Carl Ernst, PhD, Douglas Mental Health University Institute, 6875 Boulevard LaSalle, Montreal, QC, Canada H4H 1R3 (carl.ernst@mcgill.ca).

Abstract

Some intellectual disability syndromes are caused by a mutation in a single gene and have been the focus of therapeutic intervention attempts, such as Fragile X and Rett Syndrome, albeit with limited success. The rate at which new drugs are discovered and tested in humans for intellectual disability is progressing at a relatively slow pace. This is particularly true for rare diseases where so few patients make high-quality clinical trials challenging. We discuss how new advances in human stem cell reprogramming and gene editing can facilitate preclinical study design and we propose new workflows for how the preclinical to clinical trajectory might proceed given the small number of subjects available in rare monogenic intellectual disability syndromes.

Keywords: induced pluripotent stem cells, drug discovery, high throughput screening, intellectual disability

Introduction

Developmental disorders are characterized by deficits in development and include a wide variety of conditions including attention deficit disorder, intellectual disability (ID), and autism spectrum disorders (ASDs). In the United States, between the years 2006 and 2008, approximately 1 in 6 children were diagnosed with a developmental disability (Boyle et al., 2011), suggesting a significant burden on affected families and society. More specifically, ID, defined as deficits in intellectual and adaptive functioning before age 18, has an estimated prevalence of 14.3 in 1000 individuals (Chiurazzi and Pirozzi, 2016), affecting approximately 1% of the global population (Maulik et al., 2011). The causes of ID can vary, but a large portion of cases can be attributed to pathological genetic mutations, affecting either multiple or single genes. Disorders in which a single gene is mutated are classified as monogenic, while disorders where chromosomal regions are missing or duplicated, for example, are defined as multigenic.

Monogenic ID syndromes have been associated with several clinical diagnoses, including Fragile X (FXS) and Rett Syndrome (RS), but as a whole affect a large proportion of the ID population. This genetic heterogeneity poses a challenge to developing therapeutics due to the varying underlying causes of individual syndromes. Treatment options are limited: usually, individuals with ID are offered palliative care for related features (e.g., treatment for seizure control) or behavioral therapy such as one-on-one learning, but there is essentially no treatment for ID itself. The notable exception to this are in the rare cases where a gene critical in a metabolic pathway is identified as mutated (Jaggumantri et al., 2015). Dietary control, usually by limiting certain amino acids, can lead to reversal of ID depending on how

early the deficiency is detected. There are also some instances where treatments are prescribed when well-known molecules are used, because they have previously been attempted on patients with similar mutations, with some success (van Karnebeek et al., 2016).

The current model of therapeutic design for monogenic ID syndromes highlights the need to develop new strategies. Previous attempts at treating FXS and RS relied mostly on repurposing drugs approved for other uses (Figure 1). These are usually tested in a small number of subjects with very specific outcome measures, often not associated with ID. For example, recent clinical trials for RS have attempted to use ketamine, which is also investigated as a treatment for depression, to improve breathing and behavioral symptoms, though with little success (NCT02562820). In addition, FXS trials have tested mGluR5 antagonists (NCT00718341, NCT01253629, NCT01357239) and an acetylcholinesterase inhibitor (Kesler et al., 2009) in an attempt to reduce repetitive behaviors and improve memory deficits associated with the disorder. All preclinical work for these studies have been based on data from rodent models of disease. The timeline of these studies is shown in Figure 1, with recent patents for FXS and RS treatments presented in supplementary Table 1. The purpose of providing both a timeline (Figure 1) and table of patents for FXS and RS is to survey what has been tried or will be tried in the coming years. We believe drastic changes need to happen to develop more high-risk/high-reward therapies for FXS and RS, but also all monogenic ID syndromes, most of which remain in a preclinical phase.

Two approaches to drug development are currently viewed as the most plausible in the context of ID: one where therapies target a single causative gene (e.g., what has been produced in FXS and RS), and the other, where convergent molecular pathways are identified allowing for the same drug to be used across syndromes. Both would be done in the context of the underlying genetics of disease. Using a demethylation strategy for FXS would be an example of the first category, as is being attempted with 5-aza-cytidine. While this chemical would affect methylation of the whole genome, the intent is to demethylate the hypermethylated region of *FMR1*. This is a drug treatment that would presumably only work in the context of FXS. Treating deficits in protein synthesis with metformin is an example of the second category. Despite being caused by different genes, protein deficits may be seen in FXS (Gkogkas et al., 2014) and other distinct ID syndromes (Buffington et al., 2014; Huber et al., 2015). Thus, metformin might be a treatment for any ID where there is evidence of protein synthesis alterations.

Breakthroughs in high-throughput sequencing, gene editing technology, and small molecule screening have provided investigators an unprecedented opportunity to combine these technologies to develop therapeutic interventions for many ID syndromes. The goal of this paper is to conceptualize what drug development might look like for rare monogenic ID syndromes in the context of this rapidly changing technological framework.

Preclinical Phase

Animal models of genetic syndromes, whether knock-in of human variants or knock-outs of endogenous orthologs, remain critical in the drug development pipeline. Genetic engineering improvements such as CRISPR have drastically improved the efficiency of creating mouse knock-in/knock-out models and has allowed for transgenic work beyond mice. There is currently a shift on studies
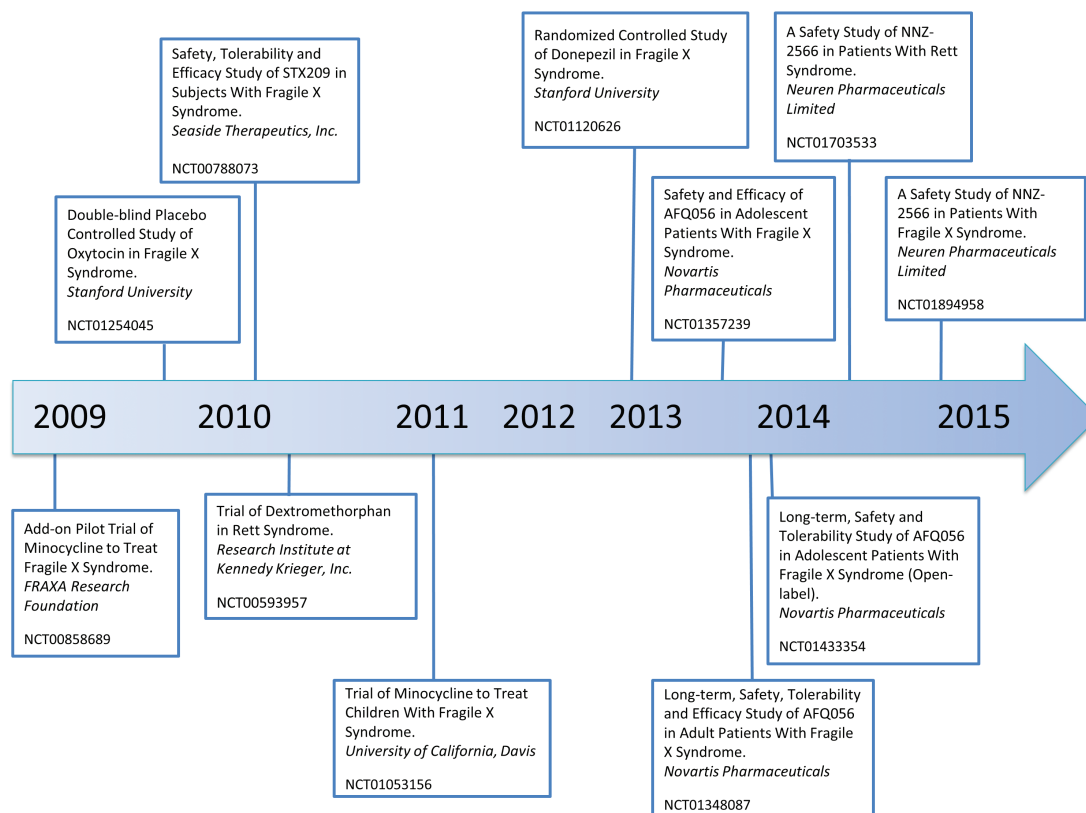

Figure 1. Completed clinical trials in Fragile X syndrome (FXS) and Rett syndrome (RS). Studies are organized by date of completion.

of neurodevelopmental disorders whereby nonhuman primates are being used to model human monogenic ID syndromes (Liu et al., 2016), assuming the more evolutionarily related the species is to humans, the more accurate the model. While *Macaque* has been used to express human *MECP2* (Liu et al., 2016) and mutant *SHANK3* (Y. Zhou, G. Feng, unpublished observations), further studies may involve marmoset since litter sizes are larger (1–3 offspring/female) and generation times are shorter. Nonhuman primate ID models can also be maintained for many generations. We expect many genetic syndromes affecting the brain to have a primate model within the next 10 years, where it currently takes about 7 years to establish a model, at least in *Macaque*.

Induced pluripotent stem cells (iPSCs) have led to new insights into neurological disease modeling. Patient cells can be made into neurons from any somatic cell, and genetic mutations can be edited to create isogenic control lines. Recent work from our laboratory has described this in a rare neurodevelopmental disease context, where we can move from urine collection to neuronal subtypes in 75 days (Bell et al., 2017). By converting somatic cells to neurons, previously unattainable patient cells are now available directly from the patients themselves (Hallett et al., 2015), eliminating intra- and interspecies variation. This technique offers the opportunity to accurately model patient neurons to what would be expected in their natural in vivo environment (Marchetto et al., 2011). Combined with gene editing technology such as CRISPR/Cas9, iPSC models can now be generated to recapitulate the exact mutation of the patient (Dow, 2015; Smith et al., 2015). Isogenic control lines can also be generated as an accurate model of the disease cell lacking the causative mutation. Using isogenic cells (either patient corrected cells or mutated control lines) significantly reduces the noise caused by varying genetic backgrounds in patient-derived cells. To date, most iPSC studies involving ID patient cells use very few cell lines (Chailangkarn et al., 2016), with as little as 3 patients in a study. Extreme variation in cell line transcriptional patterns exist, even from controls, particularly for "sister" cell lines (Germain and Testa, 2017). This suggests that patient-derived cells will remain important but perhaps secondary to studies in more carefully controlled experiments using isogenic knock-in or knock-out cell lines.

Despite the ability to create accurate iPSC disease models and isogenic controls, identifying screenable cellular phenotypes associated with disease remains a challenge. What does one look for in culture with respect to a disease associated with ID? Obvious possibilities are spine density or synaptic connectivity; however, these pose significant challenges when attempting to screen many thousands of molecules because of the high resolution required for accurate assessment and stochastic growth in culture. Ideally, a cellular read-out would be relevant to disease, have a large effect size compared with controls, be scalable, and allow for within cell endogenous controls.

High throughput screening (HTS) is the idea of assessing small molecules for alteration of a cellular phenotype associated with disease. Paired with a robust cellular assay predictive of pathophysiology, thousands of small molecules can be tested for their efficacy in reversing an in vitro phenotype. To do this, 3 assumptions are made. First, an accurate cell model of the disease is available. In today's technological framework, the assumption is that these may be human iPSC-derived cells, though this need not always be the case. Second, a quantifiable assay has been developed. Often the most challenging step, these assays require clear specificity and measurable change from negative and positive controls. They also need to be simple enough that the resolution of detection is meaningful. For example, a dendritic outgrowth assay may not be possible

to perform accurately given the fine detail of measurements required. Third, libraries of small molecules are available that might improve the cellular phenotype. Currently, several libraries are publicly available, and most drug companies are willing to provide their compounds developed for other diseases for repurposing (Kim et al., 2014).

What might a preclinical HTS pipeline look like for a rare ID syndrome? We have recently developed a pipeline to assess a gene dosage syndrome (shown in Figure 2 is the *KMT2D* gene, but any gene associated with loss of dosage could be used). Notably, gene dosage syndromes potentially caused by de novo structural variants disrupting a known ID gene may account for 29% of cases, compared with 31% due to de novo point mutations (Gilissen et al., 2014), making treatment for any gene dosage syndrome valuable to develop. Briefly, our pipeline requires the reprogramming and differentiation of patient somatic cells to a neural progenitor cell state, where patient disease is caused by a deletion of a single gene (not point mutations). As a control, we use sex-matched sibling cells derived under identical conditions to patient cells. We perform reprogramming and gene editing simultaneously, allowing us to acquire clonal iPSC populations (Bell et al., 2017). At the somatic cell stage, we have taken advantage of the 2A system (Ryan et al., 1991; de Felipe et al., 2010), which gets edited into the stop codon of the disease of interest with a reporter gene, such as GFP, in both the healthy and diseased sibling. Because 2A peptide-containing proteins are self-cleaved within the 2A peptide, their introduction allows for the ligation of genes to produce multiple proteins from a single transcript (i.e., GFP and a gene of interest) to allow for simultaneous expression. This results in an assay that produces one molecule of GFP, for example, for every protein of interest in a 1:1 ratio. Knocking this into patient and sibling controls should recapitulate disease, whereby disease cells show ~50% GFP expression compared with control cells. Small molecules in HTS assays can then be assessed for their ability to increase GFP signal in disease cells to levels more similar to control cells (Figure 2). The premise is that the underlying feature of disease is a loss of overall dosage and that increasing output from the wild-type allele may compensate to improve disease. This means that molecules that affect the regulation of gene expression are screened rather than those that target protein function. Also, only patients with gene deletion or complete loss of function (i.e., not dominant negative mutations) and with one intact allele can be assessed. Finally, one can knock-in a different reporter such as RFP into both copies of a gene that is not expected to change, providing a way to screen out molecules that are promiscuous or that create systemic changes.

Clinical Phase

Because of the genetic heterogeneity and variable expressivity of many ID syndromes, a gene-first approach to clustering individual patients is ideal for developing treatments. Framing therapeutic intervention as a treatment for mutations in gene X is preferable to a treatment for a clinical diagnosis. For example, mutations in *SHANK3* have been observed in patients with schizophrenia (Gauthier et al., 2010) and ASDs (Durand et al., 2007). While these different mutations may affect the SHANK3 protein differently to lead to different effects on the brain, one might group both schizophrenia and ASD patients in a *SHANK3* cohort and develop molecular treatments for the *SHANK3* mutation.

Gene-first patient cohorts, defined as patients with similar but not identical clinical features and with predicted causative mutations in the same gene, are now accumulating. For example,
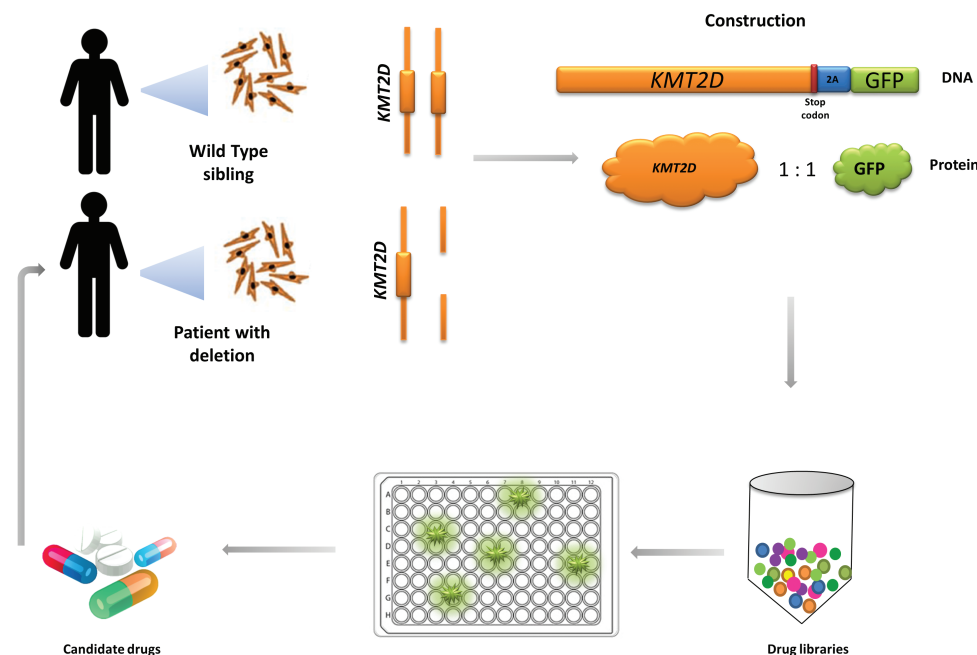
Figure 2. An iPSC/gene editing drug discovery strategy for monogenic intellectual disability (ID) syndromes caused by reduced gene dosage and where one functional allele is still present. Neural progenitor cells (NPCs) derived from renal epithelial cells (RECs) collected from the urine of a patient with a heterozygous gene deletion and a sex-matched sibling carrying both wild-type alleles, here using *KMT2D* (whose dosage loss causes Kabuki Syndrome) as an example. The gene editing strategy is to knock-in a 2A element with GFP (also called *cis*-acting hydrolase element, CHYSEL) immediately after the stop codon in both the affected sibling and the healthy sibling. This produces one molecule of GFP for every molecule of KMT2D translated, so that the GFP signal is stochiometric to the amount of KMT2D protein. Control genes can also be targeted with different fluorescent genes, such as RFP (not shown) using the same strategy. The system provides a ready assay for drug screening because patient cells should produce about 50% of GFP levels of control cells. High throughput screening (HTS) of any small molecule can then be performed to look for molecules that boost the GFP signal to levels more similar to control cells. Importantly, once a pipeline is established, any gene dosage syndrome could be assessed, alleviating the need to find different phenotypes for different monogenic diseases. Essentially, the system takes advantage of the genetic underpinnings of the disease and targets the regulatory machinery of a gene of interest rather than the protein product of said gene.

Kleefstra syndrome, caused by mutations in *EHMT1*, now has over 100 subjects enrolled in an online database run by a family foundation (GeneSpark.org). Furthermore, *GRIN2B* syndrome has over 20 patients reported in the literature to date. Several other rare mutations resemble this trend, meaning that cohorts will increase as more people with ID are assessed at the genetic level. Highly motivated families usually lead the push to create gene-first cohorts, allowing for the connection of a community and to act as a resource for researchers willing to collaborate. Assembling gene-first cohorts provides a single gene anchor with which to consider therapeutic intervention and clinical trials.

The road from small molecule identification to human clinical trials is a long one. Molecules that have not been previously assessed for safety must first go through rigorous testing in animals and phase 1 human trials. This long timeline largely explains why only molecules with known safety profiles have been used in FXS and RS (Figure 1). However, this drastically limits the molecules that can be tested. Further criteria in selecting effective molecules involve assessing the efficacy of their delivery to the CNS. Systemic drug treatment for ID syndromes remains a challenge due to the blood-brain barrier and the blood-cerebrospinal fluid barrier, which restricts the passive diffusion of many drugs into the brain. To reach effective drug concentrations in the CNS, invasive and noninvasive strategies have been developed to enhance drug delivery, which must be taken into account when identifying a potential drug of interest. Given that brain formation will consolidate with increasing age, the period in development in which a drug is administered should also be considered. Administering drug treatments during early infancy would theoretically be ideal to mitigate ID; however, oftentimes

it is unknown how a compound may affect development and lead to undesirable outcomes. Many drugs selected for repurposing, with known safety profiles, have not been assessed in children or toddlers. Clearly, additional limitations remain that must be addressed when designing an effective clinical trial.

Sample size, dose, and length of time are all critical design aspects to be carefully considered for a successful clinical trial. Monogenic ID syndromes have the advantage of patient stratification by genotype but the associated disadvantage of having small sample sizes. Drug screening in human clinical trials will likely carry forward with small sample sizes, given that most monogenic ID disorders are rare. One solution is a continuous crossover design trial, as described for another rare genetic disorder (Khasnavis et al., 2016), which we modify and illustrate here. In this design, each patient is continuously exposed to either drug or placebo, over many trials, allowing patient response to be assessed continuously (Figure 3). While this study design addresses the problem of small sample size, it also comes with several limitations. Because subjects are continuously exposed and not exposed to a drug, patients and families who score behavior may easily recognize side effects associated with active medication and absent from placebo. Also, constant switching from active medication to placebo may potentially induce an adverse response among patients. Furthermore, there is the risk of a carryover effect in which the effect of the drug persists after a patient is switched to placebo, thereby complicating data interpretation. Lastly, the higher cost and length of time required for a continuous crossover design is a major limitation that remains to be addressed. Nevertheless, the advantages of this approach are numerous: small sample sizes are powered by
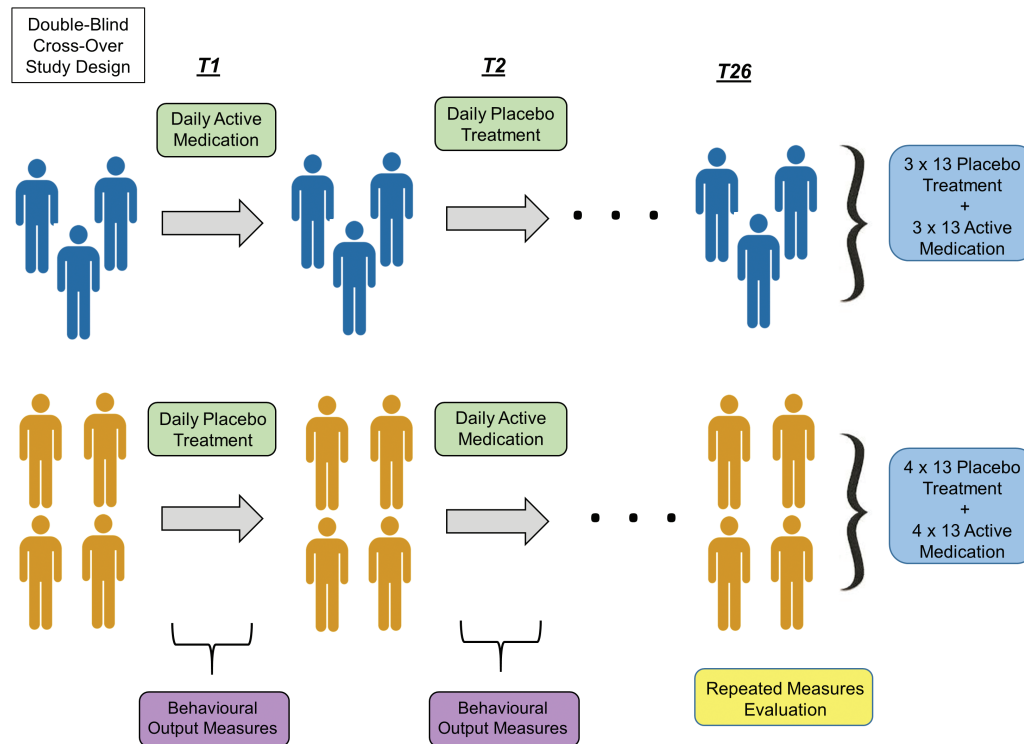
Figure 3. A randomized, repeated, within-subjects cross-over drug testing strategy for monogenic intellectual disability (ID) syndromes with small sample sizes. Most monogenic ID syndromes are rare, affecting few subjects, so novel study designs are essential to test drugs and ensure any negative results are truly negative. In the example shown, 7 subjects are recruited to test the effectiveness of a drug. Subjects are randomized to active medication (top row) or placebo (bottom row). In the first 2-week time point (T1), subjects receive identical regiments of either placebo or active medication, and parents are asked to score behavioral output measures. At the end of T1, subjects previously receiving placebo receive active medication for T2, while subjects that received active medication receive placebo for T2. Behavioral output measures are identical to T1 and all subsequent time points. To generate power, subjects are continuously crossed-over between active medication and placebo, and this need not be done sequentially, provided an equal number of blocks of placebo and active medication are completed. For an active medication to be considered effective, subjects should show improved outcomes at all or most time points where they were taking active medication. If increased power is needed (which can be assessed by preliminary investigation of effect sizes in early time points), the study can be prolonged. In the figure shown, a 1-year period is depicted (26 time points of 2 weeks each). Study design assumes no washout period is required and no major side effects are immediately apparent during use or cessation of the active medication.

multiple exposures; patients are compared with their own placebo response; and the increased familiarity of family members who score behaviors with questionnaires over several trials may improve resolution (i.e., they may know what to look for during the many observations, while a double-blind design ensures they do not know if active medication is received).

Conclusion

There is a pressing need to develop treatments for ID syndromes given the many people affected. Gene-first patient cohorts need to be created, iPSC models need to be developed both from patients and using gene editing technologies, and nonhuman primate models using these gene editing technologies should be developed for genes that are unambiguously related to ID. Innovation and creativity with new testing designs need to be tried in clinical trials.

Acknowledgments

Statement of Interest

References

Bell S, Peng H, Crapper L, Kolobova I, Maussion G, Vasuta C, Yerko V, Wong TP, Ernst C (2017) A rapid pipeline to model rare neurodevelopmental disorders with simultaneous CRISPR/Cas9 gene editing. Stem Cells Transl Med 6:886–896.

Boyle CA, Boulet S, Schieve LA, Cohen RA, Blumberg SJ, Yeargin-Allsopp M, Visser S, Kogan MD (2011) Trends in the prevalence of developmental disabilities in US children, 1997–2008. Pediatrics 127:1034–1042.

Buffington SA, Huang W, Costa-Mattioli M (2014) Translational control in synaptic plasticity and cognitive dysfunction. Annu Rev Neurosci 37:17–38.

Chailangkarn T, et al. (2016) A human neurodevelopmental model for Williams syndrome. Nature 536:338–343.

Chiurazzi P, Pirozzi F (2016) Advances in understanding: genetic basis of intellectual disability. F1000Res doi: 10.12688/f1000research.7134.1.

de Felipe P, Luke GA, Brown JD, Ryan MD (2010) Inhibition of 2A-mediated 'cleavage' of certain artificial polyproteins bearing N-terminal signal sequences. Biotechnol J 5:213–223.

Dow LE (2015) Modeling disease in vivo with CRISPR/Cas9. Trends Mol Med 21:609–621.

Durand CM, et al. (2007) Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nat Genet 39:25–27.

Gauthier J, et al. (2010) De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. Proc Natl Acad Sci U S A 107:7863–7868.

Germain PL, Testa G (2017) Taming human genetic variability: transcriptomic meta-analysis guides the experimental design and interpretation of iPSC-based disease modeling. Stem Cell Reports 8:1784–1796.

Gilissen C, et al. (2014) Genome sequencing identifies major causes of severe intellectual disability. Nature 511:344–347.

Gkogkas CG, Khoutorsky A, Cao R, Jafarnejad SM, Prager-Khoutorsky M, Giannakas N, Kaminari A, Fragkouli A, Nader K, Price TJ, Konicek BW, Graff JR, Tzinia AK, Lacaille JC, Sonenberg N (2014) Pharmacogenetic inhibition of eIF4E-dependent Mmp9 mRNA translation reverses fragile X syndrome-like phenotypes. Cell Rep 9:1742–1755.

Hallett PJ, Deleidi M, Astradsson A, Smith GA, Cooper O, Osborn TM, Sundberg M, Moore MA, Perez-Torres E, Brownell AL, Schumacher JM, Spealman RD, Isacson O (2015) Successful function of autologous iPSC-derived dopamine neurons following transplantation in a non-human primate model of Parkinson's disease. Cell Stem Cell 16:269–274.

Huber KM, Klann E, Costa-Mattioli M, Zukin RS (2015) Dysregulation of mammalian target of rapamycin signaling in mouse models of autism. J Neurosci 35:13836–13842.

Jaggumantri S, Dunbar M, Edgar V, Mignone C, Newlove T, Elango R, Collet JP, Sargent M, Stockler-Ipsiroglu S, van Karnebeek CD (2015) Treatment of creatine transporter (SLC6A8) deficiency with oral s-adenosyl methionine as adjunct to l-arginine, glycine, and creatine supplements. Pediatr Neurol 53:360–363.

Kesler SR, Lightbody AA, Reiss AL (2009) Cholinergic dysfunction in fragile X syndrome and potential intervention: a preliminary 1H MRS study. Am J Med Genet A 149:403–407.

Khasnavis T, Reiner G, Sommerfeld B, Nyhan WL, Chipkin R, Jinnah HA (2016) A clinical trial of safety and tolerability for the selective dopamine D1 receptor antagonist ecopipam in patients with Lesch-Nyhan disease. Mol Genet Metab 117:401–406.

Kim J, Kim H, Park SB (2014) Privileged structures: efficient chemical "navigators" toward unexplored biologically relevant chemical spaces. J Am Chem Soc 136:14629–14638.

Liu Z, et al. (2016) Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. Nature 530:98–102.

Marchetto MC, Brennand KJ, Boyer LF, Gage FH (2011) Induced pluripotent stem cells (iPSCs) and neurological disease modeling: progress and promises. Hum Mol Genet 20:109–115.

Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S (2011) Prevalence of intellectual disability: a meta-analysis of population-based studies. Res Dev Disabil 32:419–436.

Ryan MD, King AM, Thomas GP (1991) Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. J Gen Virol 72:2727–2732.

Smith C, Abalde-Atristain L, He C, Brodsky BR, Braunstein EM, Chaudhari P, Jang YY, Cheng L, Ye Z (2015) Efficient and allele-specific genome editing of disease loci in human iPSCs. Mol Ther 23:570–577.

van Karnebeek CD, Bowden K, Berry-Kravis E (2016) Treatment of neurogenetic developmental conditions: from 2016 into the future. Pediatr Neurol 65:1–13.