Variations on a Theme: Crispr Models for 15q11-q13 Disorders and Beyond

Barbara J Bailus
Buck Institute for Research on Aging, Novato, CA 94945, USA

Corresponding author: Barbara J. Bailus, Buck Institute for Research on Aging, Novato, CA 94945, USA, Tel: 1-707-235-6514; E-mail: bjbailus@ucdavis.edu

Abstract

Objective: Genetic disorders involving the 15q11-q13 region in humans are complex and often exhibit a variety of mutation types that can affect multiple genes in the region. Angelman syndrome and Prader-Willi syndrome are neurodevelopmental disorders that are caused by mutations in this region. A variety of mutations, from point to large deletions are exhibited in the patient population. Although the mutations that cause several of these disorders are well understood and characterized the availability of in-vivo models that accurately reflect the high variation of mutations in individual patients has not been feasible. This review examines how applying CRISPR technology can result in more accurate models of 15q11-q13 disorders and other multiple gene disorders, including autism spectrum disorders.

Background: Previous methods of creating disease models have been cost and labor intensive making it impossible to accurately represent the variation in complex genetic disorders. The advancement of CRISPR technology has drastically changed the ease of producing in-vivo models for diseases where the mutation varies in individuals.

Methods: A review of relevant literature on Angelman syndrome, Prader-Willi syndrome, and CRISPR technology, and the implications of applying CRISPR technology to the autism field.

Results: CRISPR technology has the potential to drastically impact the 15q11-q13 disorders and autism field in the creation of more varied and accurate in-vivo models, which will advance our understanding of these diseases and potentially lead to better treatments.

Keywords: CRISPR; Prader-Willi syndrome; Angelman syndrome; Autism spectrum disorders; 15q11-q13; iPSC

Abbreviations:

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeat; TALE: Transcription Activator-Like Effector; ZF: Zinc Finger; AS: Angelman Syndrome; PWS: Prader-Willi syndrome; ASD: Autism Spectrum Disorder

Introduction

Disorders of the 15q11-q13 region have been difficult to model as there exists a wide variation in mutation types that have caused these disorders. Although the minimal mutation region is known for both Angelman Syndrome (AS) and Prader-Willi syndrome (PWS) few patients have these microdeletions, instead patients have a variety of mutations including point mutations, large deletions of multiple genes, imprint center control mutations and uniparental disomy [1,2]. This variation of mutations has been difficult to replicate in both animal and cellular models, instead there has been a heavy reliance on one predominant model being used for a majority of the studies [3,4]. A similar situation of multiple models being necessary for a disorder can be found in autism spectrum disorders (ASD), a percentage of individuals with 15q11-q13 disorders are also classified on the autism spectrum. Over 100 candidate genes are implicated in ASD, of which a combination of these genes may be mutated in a single individual [5].

Until recently the task of making multiple disease models with varied and exact deletion sizes or multiple gene mutations was extremely labor and cost intensive. The advent of CRISPR (clustered regularly interspaced short palindromic repeats) has the ability to dramatically reduce both time and cost in creating a vast variety of mutation sizes and multiple gene mutations [6]. This review will focus on the advantages the CRISPR technology offers in creating more varied and accurate models for 15q11-q13 associated conditions of Angelman Syndrome and Prader-Willi syndrome and the ability of the technology to be applied beyond into the ASD field, or any genetic disease.

Background

Angelman Syndrome and Prader-Willi syndrome are neurodevelopmental disorders of which a subpopulation are often placed on the autism spectrum [5,7-9]. Angelman syndrome affects 1:10,000 individuals and symptoms include ataxia, intellectual disability, sleep disorders and in some cases seizures [2,10,11]. Prader-Willi syndrome affects 1:10,000 individuals and symptoms include obesity, intellectual impairment, sleep disorders and hyperphagia [2,12]. Both diseases have been shown to occur with mutations ranging from a single point mutation to large gene spanning deletions of 4MB, making the application of a single model representative of each individual case insufficient [1,13]. The causative gene in Angelman syndrome is UBE3A, a ubiquitin ligase protein that adds ubiquitin as a post-
translational modification to target proteins for degradation [14,15]. In Prader-Willi syndrome the causative mutation is in the SNORD116 cluster, a group of small nucleolar RNAs that help with protein folding [12,16,17]. Cellular iPSC models have been derived from a small number of AS and PWS patients, but represent only a fraction of the mutation variability in the patient population [18]. Patient derived cells can be an excellent source of information, but the genetic background of each patient is different making it difficult to study the exact changes that occur based only on a combination of mutations or the size of a specific deletion. Ideal cellular models would have an identical genetic background and only a mutation variance. The introduction of induced pluripotent stem cells (iPSCs) has made it possible to create many cellular types with an identical genetic background, in an effort to understand the effect of the mutations across various cell types [19-21]. The ability to fully utilize isogenic iPSC models was limited until recently by the time consuming methods with which to introduce new mutations into an iPSC line [22,23].

The mouse models of Angelman and Prader-Willi disease focus on recapitulating the disease phenotypes, sometimes at the expense of an accurate mutation representation. The most well characterized Angelman syndrome mouse model has a deletion in the N-terminal section of the protein, effectively deleting many of the known isoforms of Ube3a [3]. This model does not reflect on a genetic level the large deletions seen in 70% of the patients, the point mutations or imprinting control region mutations seen in another 20% of the patients [13]. In many Angelman patients multiple genes are deleted in this region, attempts to mimic this large deletion have not been very successful, resulting in poor viability or inexact targeting with traditional homologous recombination technology [24-26]. Prader-Willi syndrome models have also been similarly limited, with several of the models only having a portion of the large deletion, though this had been due in part to extremely early mortality in large deletion models [4,26]. All of the current models were time and cost consuming to create and are limited in their ability to represent the full spectrum of variation of mutations in actual patients. This limitation has been problematic for both understanding the disease mechanism and range of symptoms and developing effective treatments. CRISPR technology can help with establishing new and more varied models of 15q11-q13 disorders and the same technology could be applied to ASD.

CRISPR is the latest generation in the genome engineering tools that have been discovered and characterized in the last twenty years [27]. CRISPR technology is a gene editing nuclease system, found in many bacteria and archaea [28]. The most well studied of the CRISPR nucleases is Cas9 which is targeted to a specific site in the genome by a guide RNA (gRNA) [6]. The guide RNA (gRNA), specifies 18-21bp in the DNA, making it possible to target a single unique site in the genome [29]. Unlike the previous genome engineering tools, zinc fingers and TALENs (transcription activator-like effector nucleases), CRISPR does not require any specialized expertise to use, targets any site in the genome with high efficiency, can be made to order, and has minimal off target effects [30]. Recent studies have shown that the off-target effects of CRISPR are minimal with as little as one off target site that was cleaved [31,32]. This low off-target cut rate is especially important when introducing a mutation for a disease model, to ensure that only one mutation is introduced into the genome. CRISPR has been used to successfully make several iPSC disease models and mammalian models including the generation of a mouse model with multiple specific mutations in a single generation [33-36]. Zinc fingers exhibit a significant rate of off-target binding, making them less ideal tools for disease model creation [37-39]. TALENS appear to be more specific than ZFN, however they are complicated to create and unlike the CRISPR system two TALENs are needed per desired mutation site, requiring several plasmids/viral vectors for a multiple gene disorder model [39-43]. CRISPR offers three distinct advantages when compared to previous genome engineering tools, high accuracy, ease of design and construction, the ability to fit multiple gRNAs on the plasmid/viral vector with the Cas9 nuclease drastically increasing the efficiency of creating a multiple gene disorder as only one plasmid/vector is required for the cell or embryo [44,45].

**Discussion**

An ideal disease model has the gene causing mutation and no other background mutations, in this way the disease phenotype observed is due only to the single gene mutation. The previous generation of such models was very low efficiency and often utilized a selectable marker, adding an artificial element to the disease model, and potentially even the protein [22,46]. Using patient derived cell lines would allow for a representation of the variety of mutations, but there would be large amount of background noise across the various models as each individual patient’s genome is different with potentially other mutations in different genes. For a cellular model iPSCs are the ideal model in which to do a series of deletions of varying lengths and a variety of mutations reflective of the patient population, minus the genetic background noise of the population. Using iPSCs, isogenic lines could be established from a wild-type individual so that in the disease model iPSC derived tissues the only variance would be mutations and the size of the mutation if applicable. An additional benefit to utilizing iPSCs is their ability to be reprogrammed into any cell type, allowing for observation of the disease phenotype across many cell types, which all share an identical genome. The CRISPR system has been used with high efficiency and accuracy to generate many iPSC cell models, including other neurological diseases [35,47,48]. The creation of several models of various 15q11-q13 or ASD mutations would be a simple extension of the current technology. Large deletions could be done by having two gRNAs, which would target appropriately spaced genetic sites, creating large deletions [49]. To reflect multiple gene disorders, the Cas9 nuclease can be used with multiple gRNAs, to create a model in a single step experiment. The CRISPR system has been shown to be highly efficient and accurate, with minimal off target effects and induced mutation rates as high as 15% [32,50]. The created models could be screened for off-target events by a whole genome sequence, which can be done in a time efficient and cost effective manner, ensuring that the model has only one insertion event [31,32,51].

The CRISPR system has been used to efficiently generate a wide variety of disease models in different model organisms, including primates by inserting the Cas9 DNA, gRNA and donor DNA in the embryo [33,52,53]. The 15q11-q13 region disorders and autism spectrum disorders could benefit from this methodology as several forms of autism are caused by multiple mutations in different genes [5,54-56]. Occasionally, in humans the deletion region encompasses genes that are not found in the same region on the mouse chromosome. Until CRISPR, the concept of creating mutations across multiple sites in the genome, on different chromosomes required many generations of mice and different founder lines. The CRISPR system allows for the generation of multiple gene mutations in a single generation, by designing and injecting the multiple gRNAs into the
desired embryo. Alternatively, if deletion of such a large region proved lethal or unstable then multiple mutations could be made in the genes of interest simultaneously, creating a model in a single generation [26,33]. These models would also help elucidate the interplay between multiple gene mutations that cause autism. This technology makes the establishment of these varied models efficient and achievable in any laboratory. Continued modifications of the CRISPR system have reduced off-target mutations to almost background levels, ensuring the model phenotypes accurately reflect the molecular underpinnings of the desired disease [51,57-60].

In an effort to streamline the screening of viable disease treatments, cell lines with specifically tagged proteins could be made with the CRISPR system. These lines could be utilized in the same way the current Ube3a-YFP lines have been utilized, as a high throughput screening method for small molecule drugs that activate paternal Ube3a, in an effort to find therapeutics for Angelman syndrome [46,61]. This use of Ube3a-YFP was pivotal in identifying a class of drugs, topoisomerases, which activated paternal Ube3a [61,62]. The activation of paternal Ube3a lead to a better understanding of how paternal Ube3a was silenced and if activation of paternal Ube3a could compensate for screening efficiency and could even help elucidate protein interactions. Before the optimization of CRISPR technology the tagging of over 100 ASD candidate genes would have been a very time consuming and labor intensive task, now it is a goal that could be accomplished by a single laboratory [5]. A study done on a much larger scale did a whole genome CRISPR knockout screening to identify genes involved in cancer, this study targeted 18,080 genes with 64,751 gRNAs [63]. The scale on which this was done has made the creation of an entire library of isogenic iPSCs for 15q11-q13 mutations and every ASD candidate gene is a realistic goal, with the optimization of CRISPR technology. An establishment of this library will make massive parallel drug screening possible, making the screening of potential therapies rapid and straightforward.

Conclusion

Increased availability of varying deletion length models or multiple gene mutation models will dramatically impact both the basic understanding of 15q11-q13 and autism spectrum disorders, enhancing the potential to find effective treatments for the diseases. Increased availability of varied deletion size and multiple gene mutation disease models with minimal genetic background noise in isogenic iPSCs, tissues and animal models will increase understanding of the underlying molecular mechanism, which may expand options of potential therapies. It is possible that the deletion size, or having multiple gene mutations could impact drug response and efficacy, something that could be tested in created models. This strategy is by no means an all-inclusive model, but the implementation of CRISPR technology would allow for a more representative set of models to advance the molecular and behavioral understanding of 15q11-q13 disorders and ASD, while increasing the ability to evaluate the efficacy of potential therapies.

Acknowledgements

The author would like to thank Henriette O’Geen for her careful reading of the manuscript, Lisa Ellerby, David J Segal and Mahru An for insightful conversations. BJB was supported by an NSF graduate fellowship (0707429) and a grant to UC Davis from the Howard Hughes Medical Institute through the Med into Grad Initiative (56005706).

References


