Efficient GNE myopathy disease modeling with mutation specific phenotypes in human pluripotent stem cells by base editors

Running title: Mutation specific GNE modeling by BEs

Ju-Chan Park[1*], Jumee Kim[1*], Hyun-Ki Jang[2,6], Seung-Yeon Lee[1], Keun-Tae Kim[1], Seokwoo Park[3], Hyun Sik Lee[4], Hee-Jung Choi[4], Soon-Jung Park[5], Sung-Hwan Moon[5], Sangsu Bae[2,6] and Hyuk-Jin Cha[1#]

[1]College of Pharmacy, Seoul National University, Seoul, Republic of Korea, [2]Institute for Convergence of Basic Sciences, Hanyang University, Seoul, Republic of Korea, [3]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, South Korea, [4]Department of Biological Sciences, Seoul National University College of Medicine, Seoul, South Korea, [5]Stem Cell Research Institute, T&R Biofab Co. Ltd, Siheung, Republic of Korea, [6]Department of Chemistry, Hanyang University, Seoul, Republic of Korea

* These authors contributed equally

To whom correspondence should be addressed to

Prof. Hyuk-Jin Cha, Ph.D.

College of Pharmacy, Seoul National University

1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

Tel.: +82-2-880-7825; Fax: +82-2-880-9122; E-mail: hjcha93@snu.ac.kr

Keywords: Gene editing, Base editor, human pluripotent stem cells, disease modeling, GNE myopathy, hypo-sialylation

Abstract

Isogenic pairs of cell lines derived from human pluripotent stem cells (hPSCs) enable the precise assessment of mutation-specific phenotypes through differentiation to target cells, as this method of disease modeling excludes the bias of genetic variation. However, the extremely low efficiency of precise gene editing based on homology-directed repair (HDR) with Cas9 in hPSCs remains a technical hurdle for this approach. Herein, we took advantage of currently available base editors (BEs) in hPSCs to epitomize the isogenic disease model from hPSCs with a pathophysiological indicator. Using this method, we established 14 hPSCs that harbor point mutations on the *GNE* gene, including four different mutations found in GNE myopathy patients. Because BEs activated p53 to a lesser degree than Cas9, we observed a higher editing efficiency with BEs. Four different mutations in the epimerase or kinase domains of *GNE* revealed mutation-specific hyposialylation, which was closely correlated to pathological clinical phenotypes. These mutation-specific hyposialylation patterns were evident in GNE protein structure modeling. Furthermore, treatment with a drug candidate currently under clinical trials showed a mutation-specific drug response in GNE myopathy disease models. These data suggest that isogenic disease models from hPSCs using BEs could serve as a useful tool for mimicking the pathophysiology of GNE myopathy and for predicting drug responses.

Introduction

Since the establishment of induced pluripotent stem cells (iPSCs), great advances in both autologous stem cell therapy and disease modeling from patient-specific iPSCs has been achieved [1]. The "disease-in-a-dish" model [2] enables the identification of unknown disease mechanisms and drug screening once the disease phenotype indicator has been determined [3]. However, the phenotypes arising from the disparate genetic backgrounds of the control and disease iPSCs are often more robust than those of the disease itself, which hinders translationally relevant comparison [4]. The establishment of isogenic pairs from disease to control human pluripotent stem cells (hPSCs) or vice versa using precise gene editing technology was proposed as a way to minimize this bias [4, 5].

Considering that the majority (58%) of genetic diseases result from point mutations [6], precise gene editing technologies such as base substitution based on homology-directed repair with Cas9 (hereafter referred to as HDR) [7, 8] and base editing (BE) [9] have been applied in human iPSCs to correct genetic mutations in patient iPSCs. However, it is generally accepted that the gene editing efficiency of Cas9 *via* HDR in hPSCs is extremely low compared to other cell lines [10], which remains a technical challenge for hPSC-based disease modeling. Thus, various approaches have been developed to improve editing efficiency [11, 12]. One recent study demonstrated that p53-dependent cell death, induced by a DNA double-strand break (DSB) that results from the endonuclease activity of Cas9, is responsible for its low gene editing efficiency in hPSCs [13].

While HDR operates in tandem with DSB, base substitution by base editors (BEs) is achieved through deamination of the target base by deaminase conjugated with nickase Cas9 (nCas9), without inducing DSB [6, 14, 15]. Target sequence accessibility

(hereafter referred to as "target-ability") of BEs is determined by the protospacer adjacent motif (PAM) sequence that contains the canonical trinucleotide NGG (NGG-PAM). As the target-ability of BEs is restricted by the PAM sequence, BEs based on the NG-PAM sequence were recently developed to overcome this limitation [16, 17].

GNE myopathy (OMIM #605820), a rare autosomal recessive degenerative skeletal muscle disorder, is caused by bi-allelic loss of function (LOF) mutations of the *GNE* gene. The enzymatic activity of the GNE protein depends in turn upon N-acetylmannosamine (ManNAc) kinase activity and the epimerase activity of UDP-N-acetylglucosamine, which is rate-limiting in sialic acid biosynthesis. The reduced overall epimerase activity of the GNE protein leads to hyposialylation, which is thought to cause the progressive muscle weakness characteristic of GNE myopathy[18]. There are no FDA-approved therapies for GNE myopathy [19]. The only drug candidate in clinical trials is N-acetylmannosamine (ManNAc), the intermediate product in the synthesis of sialic acid [20]. Due to the unavailability of human GNE disease muscle cells, immortalized cell lines [21] or mouse models [22] were alternatively used to study the mechanism. However, these models are insufficient to mimic pathophysiology in human muscle cells [23]. Thus, an appropriate "GNE myopathy in a dish" model is required to characterize human GNE pathophysiology.

To establish the isogenic disease model, we operationally defined the GNE disease model as attenuated sialic acid production in GNE mutants, to serve as a pathophenotype indicator. First, we demonstrated that BEs [6, 14] were more efficient than HDR in hPSCs due to lesser induction of p53-dependent cell death. Using this method, we produced 14 hPSCs with point mutations on the *GNE* gene. Isogenic pairs of hESCs with four different mutations and myocytes derived from these mutant hESCs revealed mutation-specific hyposialylation. Additionally, the 3D structure modeling of

each GNE mutant was consistent with altered GNE enzymatic activity, correlating with different degrees of hyposialylation. The differing responses of each mutant upon drug candidate treatment suggests that the mutation status of GNE patients is an important factor to consider in clinical trials of GNE myopathy drug candidates.

Results

Target-ability of point mutations causing GNE myopathy by base editors

BEs, such as the adenine base editor (ABE) and cytosine base editor (CBE), substitute the target base through the action of the deaminases TadA or APOBEC1, respectively, conjugated with nCas9, which induces single-strand breaks (SSB) (Fig. 1A). Both BEs and HDR are potential strategies for the production of isogenic pairs of disease models with hPSCs. Thus, prior to producing hPSCs for GNE myopathy modeling, we first investigated the "target-ability" of mutations causing GNE myopathy from the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/) [24]. Among the total 63 mutations found in GNE patients that are categorized as "pathogenic" or "likely pathogenic," 34 mutations (54%) occur by a transition point mutation (purine to purine or pyrimidine to pyrimidine) and are targetable with ABE or CBE (Fig. 1B and Table S1). Eight mutations (12.7%) are transversion mutations (purine to pyrimidine or vice versa), which are targetable with either the prime editor (PE) or HDR. Among transition point mutations, for establishing isogenic pairs for a GNE model (from mutant to wild-type or vice versa), 73.5% of wild-type (25 out of 34 mutations) and 70.6% of mutants (24 out of 34 mutations) are targetable by BEs (Figs. 1C and D). As mentioned above, 25.6% of wild-type (nine mutations) and 29.4% of mutants (10 mutations) cannot be edited due to the lack of a PAM sequence. The target-ability of BEs has greatly expanded due to advances in BE techniques using NG-PAM [16, 17] (e.g., NG-ABE or NG-CBE) (Figs. 1C and D). Interestingly, the relatively frequent spontaneous deamination of cytosine to uracil in nature may account for the differential utility of CBE (for C to T mutation) and ABE (for T to C mutation) for producing mutants and wild-type lines, respectively (Figs. 1C and D).

Efficiency of base editing compared to HDR in hESCs

Considering that GNE myopathy occurs by loss of function (LOF) mutations of either the epimerase or kinase domains [25], it is necessary to examine the functional effect of mutation in these domains of the GNE protein. For this, we chose four mutations of the epimerase (R160Q and I329T) and kinase (I588T and V272M) domains that would be readily achievable by ABE, CBE, and HDR (Fig. 2A). Compared to BEs, the HDR technique is more broadly targetable to all point mutations regardless of whether they are transversion or transition mutations [6]. Thus, we first compared the editing efficiencies of all currently available techniques, HDR [26], ABE (ABEmax), NG-ABE (NG-ABEmax), CBE (BE4max) and NG-CBE (NG-BE4max) [17, 27] to generate point mutations of target sequences in hESCs. The efficiency of base substitution for R160Q by CBE was close to 37% (Fig. 2B), while those of NG-CBE and HDR were 4.1% and 0.93%, respectively (Figs. 2B and S1A). Similarly, 32.5% base substitution was achieved by ABE for I588T, while NG-ABE yielded 11.2% (Fig. S1B) and HDR yielded only 1% (Fig. 2C) substitution. For the V727M mutation, on-target base substitution efficiencies were 55.8% and 10.1% with CBE and NG-CBE (Fig. 2D and S1C), respectively, while HDR achieved 3.0% editing efficiency (Fig. 2D). Base substitution for I329T with HDR, ABE, and NG-ABE showed similar results, wherein the efficiency of HDR was much lower (4.5% for I329T) than that of ABE (8.6%) and NG-ABE (6.3%) (Figs. 2E and S1D). However, the existence of bystander sequences in the target window reduced the observed percentage of base substitution without bystander mutation in V727M (Fig. S1E). Although two mutants with bystander mutation occurred in the I329T target, one was a silent mutation (I329I) and the other was due to codon degeneracy (Fig. S1F). The GNE mutant hESCs with bystander mutations are listed in Table S2.

As low HDR efficiency compared to indel by Cas9 for GNE mutations (Fig. S1G) demands laborious clonal selection in hPSCs [28], the inefficiency of HDR remains a serious technical limitation to be resolved for disease modeling in hPSCs despite the advantage of broader target-ability. Furthermore, we observed that NGG-BEs (ABE and CBE) were more efficient than NG-BEs (NG-ABE and NG-CBE) as previously reported [29].

P53 response by Cas9 determines editing efficiency

The extremely low gene editing efficiency with Cas9 in hPSCs [10, 11] results from p53-dependent cell death upon DSB caused by the endonuclease activity of Cas9 [13]. hPSCs are particularly susceptible to p53-dependent cell death once DSB occurs, which is thought to be a protective mechanism to rigorously maintain genome integrity during development [30]. Thus, we hypothesized that the low editing efficiency of HDR compared to BEs (Figs. 2B - E) may result from different p53 responses. To explore this hypothesis, we first established *TP53* KO hESCs by targeting exon 3 of *TP53* (Fig. S2A). After introducing sgRNA and Cas9 in hESCs, nutlin3, which is an inhibitor of MDM2, was applied to stabilize p53, causing massive cell death, as expected (Fig. S2B). After nutlin3 treatment, the few clones that survived were *TP53* KO hESCs (Fig. S2C). The p53-dependent response of Cas9 or BEs was assessed by the measuring the level of *PPM1D*, a typical target gene of p53 [31]. When introducing Cas9 with sgRNA into hESCs as previously described [13], BEs only marginally affected the p53 response (Fig. 2F). Consistent with the p53 response, significantly more hESCs remained and survived when BEs were used, compared to the Cas9 approach, while all *TP53* KO hESCs tolerated the stress of gene editing (Fig. 2G). The effects of gene editing techniques in WT and *TP53* KO hESCs were monitored in real-time (Movies S1 and

S2). Due to the higher rate of survival during the gene editing process in *TP53* KO hESCs than in wild-type hESCs, the gene editing efficiency of KO with both Cas9 and BEs was improved in *TP53* KO hESCs (Fig. 2H).

Establishment of GNE myopathy disease models through base editors

While BEs are limited relative to HDR by their requirement for PAM and their target-ability only to transition mutations, they yield higher gene editing efficiency than HDR and are currently the most feasible approach for base substitution in hPSCs. Accordingly, we successfully established 14 GNE mutants in total in both hESCs and hiPSCs using BEs (Tables S2). Briefly, based on H9 hESCs, two LOF mutants in the epimerase domain of GNE generated with CBE (R160Q) and ABE (I329T) (Figs. 3A and B) and two LOF mutants in the kinase domain using NG-ABE (I588T) and CBE (V727M) were established (Figs. 3C and D). Other GNE mutant cell lines (I329T, R160Q, and I588T) were established in CHA3-hESCs (Fig. S3D) and BJ-iPSCs (Figs. S3E and F) using NG-ABE, ABE, or CBE, as listed in Table S2. Of note, since the target windows of I329T and V727M carry bystander bases, five mutants with bystander mutations including I329I, a silent mutation, were also established in V727M (Figs. S3B and C) and I329T (Fig. S3D).

Mutation-specific phenotype of GNE myopathy disease models

The four established GNE mutant hESCs displayed typical cellular morphology and comparable OCT4 expression (Fig. 4A). As the putative underlying cause of GNE myopathy is hyposialylation, we first assessed levels of sialic acid production using a sialic acid-specific *Sambucus nigra agglutinin* (SNA) lectin binding assay in four different GNE-hESCs. As shown in Figure 4B, the levels of lectin staining were

significantly diminished in all mutant hESCs to different degrees, relative to controls. The mean fluorescence intensity (MFI) of SNA lectin clearly indicated that the sialylation of the protein varied depending on the mutation (Figs. 4C and D). Of note, binding of SNA lectin indicates that the level of sialylation of protein, the final step for sialic acid biosynthesis, is a reliable indicator of GNE enzymatic activity. The pathogenic severity of GNE myopathy in a patient harboring the I329T mutation [25], could therefore be predicted by the extent of sialylation reduction, which was greatest in the I329T mutant (Fig. 4D). In this cell line, the level of sialic acid produced in GNE mutants relative to WT (Fig. 4D) can be used to determine pathogenicity. Thus, we compared pathological severity previously predicted by *in silico* platforms (PolyPhen2, SIFT, Align, and Pmut) [25] to the MFI of each GNE mutant. The V727M mutant, which was predicted to be largely "pathogenic," showed moderate hyposialylation (Fig. 4C). As the V727M mutation occurs in the compound heterozygous state, even with the prevalent distribution of this variant (rs121908627), this mutation itself would have a less serious effect on GNE myopathy. We would thus expect the I588T mutant, which is a variant of uncertain significance (VUS), to be "Pathogenic/Likely pathogenic" since hyposialylation levels in this mutant are comparable to those observed in the R160Q mutant, which is classified as "Pathogenic/Likely pathogenic" (Fig. 4D).

The hyposialylation of other GNE models derived from BJ-iPSC and hCHA3 hESCs were examined. Evident hyposialylation (Fig. S4A) and recovery by sialic acid treatment (Fig. S4B), compared to that of the silent I329I mutation established in hCHA3 cells, ruled out the possibility of baseline effects of gene editing itself on hyposialylation. The comparable hyposialylation of R160Q and I588T was also reproduced in GNE mutants derived from BJ-iPSCs (Figs. S4C and D). Importantly, the levels of sialic acid differed in a cell type-specific manner depending on genetic

background (Fig. 4E). This supports the necessity of an isogenic pair for precise comparison as highlighted previously [4].

Prior to inducing the differentiation of GNE-hESCs into myocytes, we examined whether hyposialylation would affect "pluripotency." To examine this effect, I329T hESCs were used, as they showed the most severe hyposialylation. Neither alkaline phosphatase activity, which is a well-characterized pluripotency indicator, nor the mRNA expression of typical pluripotency markers were altered (Figs. S5A and B). Additionally, teratomas from the I329T mutant developed normally (Fig. S5C). Thus, we concluded that hyposialylation due to GNE mutation had minimal effects on the pluripotency of hESCs.

Structural analysis predicts mutation-specific hyposialylation

The varying degrees of hyposialylation among the GNE mutants inspired us to analyze each mutation at the molecular level. We used the GNE structure to predict the molecular consequences of each mutation on the enzymatic activity of GNE. The epimerization reaction of GNE requires tetramerization through the N-terminal domain, which is stabilized upon ligand binding (Fig. 5A) [32, 33]. R160 stabilizes the C-terminal loop via interaction with the main chains of D409, Q411, and E412 (Fig. 5B). This loop interacts with the residues involved in the N-terminal dimerization interface. The R160Q mutant contains a shorter side chain that would require loop reorganization were they to interact, which could possibly affect epimerase oligomerization (Fig. 5B). However, the hyposialylation effect of R160Q was less adverse than that of the I329T mutation. I329 participates in forming an extensive network of hydrophobic interactions (Fig 5C). Importantly, L249, located adjacent to I329, interacts directly with uridine diphosphate (UDP) (Fig. 5C). The I329T mutation, which introduces a

smaller and less hydrophobic residue, would weaken the Van der Waals interactions within the hydrophobic core, affecting the structural integrity near the substrate-binding site. Meanwhile, the kinase domain of GNE displays a typical bi-lobal architecture and contains a type I zinc-binding motif characteristic of the repressor, open reading frame, kinase (ROK) family (Fig. 5A) [34, 35]. Full-length GNE requires dimerization to permit ManNAc kinase activity [35]. I588 is located at the dimerization interface of the kinase C-terminal lobe (Fig. 5D); mutating this residue to less-hydrophobic threonine (Thr) affects kinase dimerization and thus its activity. Finally, V727 is buried in a hydrophobic core [34]. The modeling of the V727M mutation suggests that the bulkier methionine (Met) may induce structural changes at this site to avoid steric clash (Fig. 5E). However, since this residue is positioned away from the catalytic site, the V727M mutation is expected to cause a less severe phenotype than the I329T mutation.

Structural analysis of GNE mutants shows that the I329T mutation, which can cause structural instability, is located near the active site. As a result, this mutation is expected to have the most adverse effect on GNE epimerase activity. Indeed, the I329T mutant shows the most pronounced hyposialylation (Figs. 4C and D).

Mutation-specific drug response of GNE myopathy disease models

To compensate for the hyposialylation caused by reduced GNE enzymatic activity in GNE myopathy patients, administration of ManNAc is being investigated in clinical trials for GNE patients (NCT02346461) following a safety profile study [20], despite the unsatisfactory outcome of a sialic acid clinical trial [36]. We therefore hypothesized that ManNAc would elicit a distinctive response in GNE patients with a mutation in the domain of ManNAc kinase because ManNAc is a direct substrate of this enzyme (Fig. 6A). To explore this possibility, we examined the recovery of hyposialylation of GNE mutants in the epimerase domain (R160Q and I329T) and

kinase domain (I588T and V727M) after treatment with ManNAc. As predicted, the recovery of sialic acid production after ManNAc treatment occurred in GNE mutants harboring mutations in the epimerase domains (R160Q and I329T), while the effect in the kinase mutant (I588T) was marginal (Figs. 6B and C). The dose-dependent response curve clearly revealed the mutant-specific ManNAc response (Figs. 6D and E). We noticed that sialic acid production was reduced in V727M, for unknown reasons (Fig. 6E). These data suggest that the ManNAc response differs according to the mutation type of GNE.

GNE myopathy modeling in myocytes derived from GNE-hESCs

Next, to verify mutation-specific responses in myocytes, WT and GNE-hESCs were differentiated into skeletal myocytes (hereafter referred to as myocytes) through multiple steps (Fig. 7A). The expression levels of typical markers for myocytes (e.g., *MyoG*, *MYH2*, and *MYH7*) and pluripotency (e.g., *NANOG* or *SOX2*) were used to evaluate myocyte differentiation (Fig. 7B). Differentiated myocytes possessed typical elongated, rod-shaped morphology and expressed myosin heavy chain (MHC) (Fig. 7C). Myocytes differentiated from GNE-hESCs also showed typical cellular morphology, similar to that of WT (Fig. 7D). Unfortunately, we did not observe unusual structural alterations such as rimmed vacuoles, a typical clinicopathologic feature of GNE myopathy, in GNE myocytes (Fig. 7D). This may have resulted from the technical limitations associated with fully mature myocyte differentiation from hPSCs [37]. Notably, the mutation-specific hyposialylation observed in GNE-hESCs (Fig. 4D) was also more pronounced in GNE-myocytes relative to controls (Fig. 7E). Consistent with results in GNE-hESCs, ManNAc treatment restored sialylation in I329T GNE myocytes harboring epimerase mutations but not V727M GNE myocytes with kinase

mutations (Fig. 7F). While understanding of the functionality of skeletal myocytes derived from hPSCs is limited, the functionality of cardiomyocytes differentiated from hESCs (hESC-CMs) has been well-characterized [38]. Thus, we examined whether GNE mutation affects the functionality of cardiomyocytes, as cardiomyopathy has sometimes been reported in GNE patients [39]. hESC-CMs with higher Troponin T2, Cardiac type (*TNNT2*) and lower *NANOG* expression than hESCs (Fig. S6A) showed clear beating behavior (Movie S3). GNE mRNA expression was comparable regardless of mutation status (Fig. 6B). Consistent mutation-specific hyposialyation and drug responses were also observed in hESCs-CMs with different GNE mutations (Figs. S6C and S6D). The hyposialylation of hESC-CMs with the I329T mutation (Fig. S3D) was evident when compared with hESCs-CM with the I329I silent mutation (Fig. S6E). The comparison of beating behavior of hESC-CM from WT or I329T hESCs revealed no significant changes resulting from GNE mutation (Movies S4A and B).

Discussion

Base editors for disease modeling in hPSCs

Despite the high editing efficiency of BEs compared to HDR in hPSCs, the limited target accessibility of BEs relative to HDR would be disadvantageous in disease modeling using hPSCs. Through the development and implementation of NG-BEs for gene editing, coverage of BEs for GNE myopathy modeling encompassed 73.5% of transition point mutations in the present study (Fig. 1). We demonstrated that BEs were more efficient than HDR due to their lesser effect on p53 (Fig. S2). Along with PAM-less target sequences, which account for 26.5% of transition mutations in GNE myopathy, the production of undesirable mutations of bystander bases by BEs would be a significant drawback in hPSC-based disease modeling. Thus, the development of "PAM-less BEs" and "BEs with narrowed editing windows" would be advantageous for more efficient disease modeling with BE techniques, along with improved HDR.

Significance of isogenic disease model

Here, we established 14 GNE mutants from three hPSCs with different genetic backgrounds. Hyposialylation, a typical pathological phenotype of GNE myocytes, occurred in a mutation-specific manner regardless of hPSC cell type. However, because the basal level of sialic acid in the wild-type of the three hPSCs (H9, hCHA3: hESCs, BJ-iPSCs and SES8 [40]: iPSCs) differed relative to the mutant hPSCs, comparison of sialic acid between unmatched pairs could not allow us to precicelsy determine the effect of mutation (Fig. 4F). Previously, various disease models were established using iPSCs derived from patients, in which disease phenotype was compared with non-identical control iPSCs [41, 42]. This study confirmed the previously proposed [4] potential for bias introduced by the use of cells with an unmatched genetic background to hinder precise comparisons of pathological phenotypes. In addition, GNE is critical

for the early development of both skeletal and cardiac muscle, as demonstrated in GNE KO mouse ESCs [43]. Therefore, homozygous nonsense or frame-shifting mutations are not reported in GNE patients [44]. Following this reasoning, instead of the knockout model of GNE, GNE-hESCs with clinically diagnosed mutations could serve as a clinically relevant disease model.

Base editing technique for validating a variant of uncertain significance

Since large numbers of single nucleotide variants (SNVs) observed in GNE as well as other rare genetic diseases cause missense mutations, the functional consequences of mutations are often unpredictable, particularly in rare diseases with insufficient *a priori* observations and incomplete penetrance [45]. With isogenic disease models, one can classify these so-called "variants of uncertain significance" (VUS) [45] more accurately with the aid of functional information. In the present study, for example, since I588T showed a similar level of hyposialylation as R160Q, classified as "Pathogenic/Likely pathogenic," we inferred that I588T may also be pathogenic. The I588T variant was originally annotated as "uncertain significance" in the ClinVar database [24]. The efficient induction of point mutation with BEs (or PE in the near future) would enhance further studies seeking to validate the effect of VUS.

Consideration of mutation-specific drug response

We also demonstrated that epimerase mutants (R160Q and I329T) but not kinase mutants (I588T and V727M) showed recovery of sialylation after ManNAc treatment (Figs. 6D and 7G). Interestingly, while the I588T mutant showed no recovery, the V727M mutation resulted in reduced sialylation following ManNAc treatment, for unknown reasons. In light of these mutant-specific responses to ManNAc, it would be advisable to use patient stratification based on mutation status in clinical trials of ManNAc, to ensure accurate efficacy validation with ManNAc and similar drugs.

Future prospects for drug development based on isogenic disease models

A recent study developed a high throughput (HT) phenotypic screening system to measure CFTR protein in cystic fibrosis with an isogenic pair of iPSCs [46]. Similarly, HT phenotypic screening with isogenic GNE myopathy disease models using the level of sialylation as a phenotypic indicator could accelerate drug development for GNE myopathy. Furthermore, by taking advantage of isogenic disease models and a drug-transcriptome database such as the Connectivity MAP [47], drug candidates from the FDA-approved drug library that reverse pathogenic gene signatures can easily be predicted [48, 49]. The efficacy of these drug candidates can then be validated in the "GNE myopathy in a dish" model, making this strategy a strong potential option to develop drugs for GNE myopathy in the future.

Conclusion

NG-BEs covered the majority of pathogenic transversion point mutation and showed higher precise gene editing efficiency than HDR in hPSCs. The isogenic "GNE myopathy in a dish model," established in hPSCs by BEs with relatively high efficiency, revealed mutation-specific hyposialylation and drug responses. This work establishes important cellular tools and techniques for further mechanistic studies and development of drug screening systems.

Author contributions

HJ.C conceived the overall study design and led the experiments. JC.P and J.K mainly conducted the experiments, data analysis, and critical discussion of the results. HK.J and S.B contributed to sgRNA design and NGS analysis and provided gene editing techniques. SY.L, KT.K and S.P. produced mutant lines. HS.L and HJ.C performed 3D protein structure modeling. SJ.P and SH.M conducted the experiments with differentiated cardiomyocytes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by a grant from the National Research Foundation of Korea (2017M3A9B3061843 and 2020R1A2C2005914) and a grant from the Research of Korea Centers for Disease Control and Prevention (2020-ER6902-00).

Materials and Methods

Live cell ratio after gene editing in WT and TP53 KO hESCs

To trace the number of live cells during gene editing, cells were documented by JuLI stage for 36 hours after transfection. Cells were transfected with the same condition in 'Generation of TP53 knock-out and GNE mutant hPSCs'. The number of cells at the time point of 2 hours and 36 hours were counted by cell counter ImageJ plugin. To calculate the live cell ratio, the number of cells at the time point of 36 hours were divided by the number of cells at the time point of 2 hours.

Generation of *TP53* knock-out and *GNE* mutant hPSCs

For hPSC transfection, cells were detached with Accutase™ (561527, BD Bioscience) and washed with DPBS for three times and resuspended to concentration of 1 x 10^6 cell in 100 μl with Opti-MEM (31985070, Gibco). 2 μg of Cas9 or BE vectors (Cas9, BE4max, NG-BE4max, ABEmax and NG-ABEmax cloned in pCMV vector), 3 μg of sgRNA vector and 2 μg of ssODN (in case of HDR) was added to 100 μl of cell mixture. Electroporation was performed by NEPA-21 with 175 V of poring pulse and 2.5 mV of transfer pulse. To enrich *TP53* KO cells, 10 μM of nutlin3 was treated for 48 hours after 5 days from transfection.

Flow cytometry

Cells were detached with Accutase solution (561527, BD Bio-sciences) for hPSCs or with TrypLE (12604, Gibco) for myocytes, followed by 3 times DPBS wash and then analyzed through FACS Calibur I (BD Bioscience). For the determination of sialic acid, cells were stained with Fluorescein SNA (FL-1301, Vector Laboratories). CellQuest Pro soft-ware was used for FACS analysis.

Generation of skeletal muscle cells from hESC

hESCs were differentiated to skeletal muscle cells by using commercially available skeletal muscle differentiation medium kit (Amsbio). In brief, Cells were seeded onto a collagen I (Sigma)-coated dish at 5,000 cells /cm2 dish and maintained for 6 days in skeletal muscle induction medium (SKM01, Amsbio). At day 6, cells were dissociated with 0.05% trypsin, seeded onto a collagen I-coated dish at 5,000 cells /cm2 dish, and maintained for 6 days in myoblast cell culture medium (SKM02, Amsbio). At day 12, when myoblasts are confluent, the medium was changed to myotube cell culture medium (SKM03, Amsbio) and freshly changed every 3 days for more than 6 days. For drug response test, cells were treated with ManNAc (Sigma) at the indicated times and concentrations and the media with ManNAc was changed every day.

Statistical analysis

The quantitative data are expressed as the mean values ± standard deviation (SD). Student's unpaired t-tests was performed to analyze the statistical significance of each response variable using the PRISM. p values less than 0.05 were considered statistically significant (* < 0.05, ** <0.001 and ***<0.0001)

Reference

[1] (!!! INVALID CITATION !!! [1]).

[2] G. Tiscornia, E.L. Vivas, J.C. Izpisua Belmonte, Diseases in a dish: modeling human genetic disorders using induced pluripotent cells, Nat Med 17(12) (2011) 1570-6.

[3] (!!! INVALID CITATION !!! [3]).

[4] Florian T. Merkle, K. Eggan, Modeling Human Disease with Pluripotent Stem Cells: from Genome Association to Function, Cell Stem Cell 12(6) (2013) 656-668.

[5] K. Musunuru, Genome editing of human pluripotent stem cells to generate human cellular disease models, Disease Models & Mechanisms 6(4) (2013) 896-904.

[6] H.A. Rees, D.R. Liu, Base editing: precision chemistry on the genome and transcriptome of living cells, Nat Rev Genet 19(12) (2018) 770-788.

[7] J. Jackow, Z. Guo, C. Hansen, H.E. Abaci, Y.S. Doucet, J.U. Shin, R. Hayashi, D. DeLorenzo, Y. Kabata, S. Shinkuma, J.C. Salas-Alanis, A.M. Christiano, CRISPR/Cas9-based targeted genome editing for correction of recessive dystrophic epidermolysis bullosa using iPS cells, Proc Natl Acad Sci U S A (2019).

[8] A. Izmiryan, C. Ganier, M. Bovolenta, A. Schmitt, F. Mavilio, A. Hovnanian, Ex Vivo COL7A1 Correction for Recessive Dystrophic Epidermolysis Bullosa Using CRISPR/Cas9 and Homology-Directed Repair, Mol Ther Nucleic Acids 12 (2018) 554-567.

[9] M.J. Osborn, G.A. Newby, A.N. McElroy, F. Knipping, S.C. Nielsen, M.J. Riddle, L. Xia, W. Chen, C.R. Eide, B.R. Webber, H.H. Wandall, S. Dabelsteen, B.R. Blazar, D.R. Liu, J. Tolar, Base Editor Correction of COL7A1 in Recessive Dystrophic Epidermolysis Bullosa Patient-Derived Fibroblasts and iPSCs, J Invest Dermatol 140(2) (2020) 338-347 e5.

[10] P. Mali, L. Yang, K.M. Esvelt, J. Aach, M. Guell, J.E. DiCarlo, J.E. Norville, G.M. Church, RNA-guided human genome engineering via Cas9, Science 339(6121) (2013) 823-6.

[11] K.T. Kim, J.C. Park, H.K. Jang, H. Lee, S. Park, J. Kim, O.S. Kwon, Y.H. Go, Y. Jin, W. Kim, J. Lee, S. Bae, H.J. Cha, Safe scarless cassette-free selection of genome-edited human pluripotent stem cells using temporary drug resistance, Biomaterials 262 (2020) 120295.

[12] T.L. Maurissen, K. Woltjen, Synergistic gene editing in human iPS cells via cell cycle and DNA repair modulation, Nat Commun 11(1) (2020) 2876.

[13] R.J. Ihry, K.A. Worringer, M.R. Salick, E. Frias, D. Ho, K. Theriault, S. Kommineni, J. Chen, M. Sondey, C. Ye, R. Randhawa, T. Kulkarni, Z. Yang, G. McAllister, C. Russ, J. Reece-Hoyes, W. Forrester, G.R. Hoffman, R. Dolmetsch, A. Kaykas, p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells, Nat Med 24(7) (2018) 939-946.

[14] A.C. Komor, Y.B. Kim, M.S. Packer, J.A. Zuris, D.R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage, Nature 533(7603) (2016) 420-4.

[15] N.M. Gaudelli, A.C. Komor, H.A. Rees, M.S. Packer, A.H. Badran, D.I. Bryson, D.R. Liu, Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage, Nature 551(7681) (2017) 464-471.

[16] H. Nishimasu, X. Shi, S. Ishiguro, L. Gao, S. Hirano, S. Okazaki, T. Noda, O.O. Abudayyeh, J.S. Gootenberg, H. Mori, S. Oura, B. Holmes, M. Tanaka, M. Seki, H.

Hirano, H. Aburatani, R. Ishitani, M. Ikawa, N. Yachie, F. Zhang, O. Nureki, Engineered CRISPR-Cas9 nuclease with expanded targeting space, Science 361(6408) (2018) 1259-1262.
[17] T.P. Huang, K.T. Zhao, S.M. Miller, N.M. Gaudelli, B.L. Oakes, C. Fellmann, D.F. Savage, D.R. Liu, Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors, Nat Biotechnol 37(6) (2019) 626-631.
[18] S. Noguchi, Y. Keira, K. Murayama, M. Ogawa, M. Fujita, G. Kawahara, Y. Oya, M. Imazawa, Y. Goto, Y.K. Hayashi, I. Nonaka, I. Nishino, Reduction of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase activity and sialylation in distal myopathy with rimmed vacuoles, J Biol Chem 279(12) (2004) 11402-7.
[19] I. Nishino, N. Carrillo-Carrasco, Z. Argov, GNE myopathy: current update and future therapy, J Neurol Neurosurg Psychiatry 86(4) (2015) 385-92.
[20] X. Xu, A.Q. Wang, L.L. Latham, F. Celeste, C. Ciccone, M.C. Malicdan, B. Goldspiel, P. Terse, J. Cradock, N. Yang, S. Yorke, J.C. McKew, W.A. Gahl, M. Huizing, N. Carrillo, Safety, pharmacokinetics and sialic acid production after oral administration of N-acetylmannosamine (ManNAc) to subjects with GNE myopathy, Mol Genet Metab 122(1-2) (2017) 126-134.
[21] N.D. Pham, P.C. Pang, S. Krishnamurthy, A.M. Wands, P. Grassi, A. Dell, S.M. Haslam, J.J. Kohler, Effects of altered sialic acid biosynthesis on N-linked glycan branching and cell surface interactions, J Biol Chem 292(23) (2017) 9637-9651.
[22] Y.M. Chan, P. Lee, S. Jungles, G. Morris, J. Cadaoas, A. Skrinar, M. Vellard, E. Kakkis, Substantial deficiency of free sialic acid in muscles of patients with GNE myopathy and in a mouse model, PLoS One 12(3) (2017) e0173261.
[23] R.L. Perlman, Mouse models of human disease: An evolutionary perspective, Evol Med Public Health 2016(1) (2016) 170-6.
[24] M.J. Landrum, J.M. Lee, M. Benson, G. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, J. Hoover, W. Jang, K. Katz, M. Ovetsky, G. Riley, A. Sethi, R. Tully, R. Villamarin-Salomon, W. Rubinstein, D.R. Maglott, ClinVar: public archive of interpretations of clinically relevant variants, Nucleic Acids Res 44(D1) (2016) D862-8.
[25] F.V. Celeste, T. Vilboux, C. Ciccone, J.K. de Dios, M.C. Malicdan, P. Leoyklang, J.C. McKew, W.A. Gahl, N. Carrillo-Carrasco, M. Huizing, Mutation update for GNE gene variants associated with GNE myopathy, Hum Mutat 35(8) (2014) 915-26.
[26] L. Yang, M. Guell, S. Byrne, J.L. Yang, A. De Los Angeles, P. Mali, J. Aach, C. Kim-Kiselak, A.W. Briggs, X. Rios, P.Y. Huang, G. Daley, G. Church, Optimization of scarless human stem cell genome editing, Nucleic Acids Res 41(19) (2013) 9049-61.
[27] L.W. Koblan, J.L. Doman, C. Wilson, J.M. Levy, T. Tay, G.A. Newby, J.P. Maianti, A. Raguram, D.R. Liu, Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction, Nat Biotechnol 36(9) (2018) 843-846.
[28] H.L. Li, P. Gee, K. Ishida, A. Hotta, Efficient genomic correction methods in human iPS cells using CRISPR-Cas9 system, Methods 101 (2016) 27-35.
[29] Y.K. Jeong, J. Yu, S. Bae, Construction of non-canonical PAM-targeting adenosine base editors by restriction enzyme-free DNA cloning using CRISPR-Cas9, Sci Rep 9(1) (2019) 4939.
[30] U. Weissbein, N. Benvenisty, U. Ben-David, Quality control: Genome maintenance in pluripotent stem cells, J Cell Biol 204(2) (2014) 153-63.

[31] M. Fiscella, H. Zhang, S. Fan, K. Sakaguchi, S. Shen, W.E. Mercer, G.F. Vande Woude, P.M. O'Connor, E. Appella, Wip1, a novel human protein phosphatase that is induced in response to ionizing radiation in a p53-dependent manner, Proc Natl Acad Sci U S A 94(12) (1997) 6048-53.
[32] D. Ghaderi, H.M. Strauss, S. Reinke, S. Cirak, W. Reutter, L. Lucka, S. Hinderlich, Evidence for dynamic interplay of different oligomeric states of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase by biophysical methods, J Mol Biol 369(3) (2007) 746-58.
[33] S.C. Chen, C.H. Huang, S.J. Lai, C.S. Yang, T.H. Hsiao, C.H. Lin, P.K. Fu, T.P. Ko, Y. Chen, Mechanism and inhibition of human UDP-GlcNAc 2-epimerase, the key enzyme in sialic acid biosynthesis, Sci Rep 6 (2016) 23274.
[34] Y. Tong, W. Tempel, L. Nedyalkova, F. Mackenzie, H.W. Park, Crystal structure of the N-acetylmannosamine kinase domain of GNE, PLoS One 4(10) (2009) e7165.
[35] J. Martinez, L.D. Nguyen, S. Hinderlich, R. Zimmer, E. Tauberger, W. Reutter, W. Saenger, H. Fan, S. Moniot, Crystal structures of N-acetylmannosamine kinase provide insights into enzyme activity and inhibition, J Biol Chem 287(17) (2012) 13656-65.
[36] H. Lochmuller, A. Behin, Y. Caraco, H. Lau, M. Mirabella, I. Tournev, M. Tarnopolsky, O. Pogoryelova, C. Woods, A. Lai, J. Shah, T. Koutsoukos, A. Skrinar, H. Mansbach, E. Kakkis, T. Mozaffar, A phase 3 randomized study evaluating sialic acid extended-release for GNE myopathy, Neurology 92(18) (2019) e2109-e2117.
[37] N. Jiwlawat, E. Lynch, J. Jeffrey, J.M. Van Dyke, M. Suzuki, Current Progress and Challenges for Skeletal Muscle Differentiation from Human Pluripotent Stem Cells Using Transgene-Free Approaches, Stem Cells Int 2018 (2018) 6241681.
[38] C.L. Mummery, J. Zhang, E.S. Ng, D.A. Elliott, A.G. Elefanty, T.J. Kamp, Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview, Circ Res 111(3) (2012) 344-58.
[39] M.C. Malicdan, L. Mian, P. Leoyklang, F. Celeste, D. Despres, P. Zerfas, N. Carrillo-Carrasco, W.A. Gahl, GP 52, Neuromuscular Disorders 9(24) (2014) 810 %@ 0960-8966.
[40] T.H. Lee, S.H. Song, K.L. Kim, J.Y. Yi, G.H. Shin, J.Y. Kim, J. Kim, Y.M. Han, S.H. Lee, S.H. Lee, S.H. Shim, W. Suh, Functional recapitulation of smooth muscle cells via induced pluripotent stem cells from human aortic smooth muscle cells, Circ Res 106(1) (2010) 120-8.
[41] A. Yamashita, M. Morioka, H. Kishi, T. Kimura, Y. Yahara, M. Okada, K. Fujita, H. Sawai, S. Ikegawa, N. Tsumaki, Statin treatment rescues FGFR3 skeletal dysplasia phenotypes, Nature 513(7519) (2014) 507-11.
[42] G. Lee, E.P. Papapetrou, H. Kim, S.M. Chambers, M.J. Tomishima, C.A. Fasano, Y.M. Ganat, J. Menon, F. Shimizu, A. Viale, V. Tabar, M. Sadelain, L. Studer, Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs, Nature 461(7262) (2009) 402-6.
[43] I. Milman Krentsis, I. Sela, R. Eiges, V. Blanchard, M. Berger, M. Becker Cohen, S. Mitrani-Rosenbaum, GNE is involved in the early development of skeletal and cardiac muscle, PLoS One 6(6) (2011) e21389.
[44] O. Pogoryelova, J.A. Gonzalez Coraspe, N. Nikolenko, H. Lochmuller, A. Roos, GNE myopathy: from clinics and genetics to pathology and research strategies, Orphanet J Rare Dis 13(1) (2018) 70.

[45] G.M. Findlay, R.M. Daza, B. Martin, M.D. Zhang, A.P. Leith, M. Gasperini, J.D. Janizek, X. Huang, L.M. Starita, J. Shendure, Accurate classification of BRCA1 variants with saturation genome editing, Nature 562(7726) (2018) 217-222.
[46] S. Merkert, M. Schubert, R. Olmer, L. Engels, S. Radetzki, M. Veltman, B.J. Scholte, J. Zollner, N. Pedemonte, L.J.V. Galietta, J.P. von Kries, U. Martin, High-Throughput Screening for Modulators of CFTR Activity Based on Genetically Engineered Cystic Fibrosis Disease-Specific iPSCs, Stem Cell Reports 12(6) (2019) 1389-1403.
[47] O.S. Kwon, W. Kim, H.J. Cha, H. Lee, In silico drug repositioning: from large-scale transcriptome data to therapeutics, Arch Pharm Res 42(10) (2019) 879-889.
[48] O.S. Kwon, H. Lee, H.J. Kong, E.J. Kwon, J.E. Park, W. Lee, S. Kang, M. Kim, W. Kim, H.J. Cha, Connectivity map-based drug repositioning of bortezomib to reverse the metastatic effect of GALNT14 in lung cancer, Oncogene (2020).
[49] J. Liu, J. Lee, M.A. Salazar Hernandez, R. Mazitschek, U. Ozcan, Treatment of obesity with celastrol, Cell 161(5) (2015) 999-1011.

Figure legends

Figure. 1 Target-ability of point mutations causing GNE myopathy by base editors (A) Graphical description of three different gene editing technologies for generating point mutations (APOBEC1: cytidine deaminase, UGI: Uracil-DNA glycosylatse inhibitor, TadA: adenine deaminase, TadA*: evolved TadA) (B) Pie Chart of distribution of GNE myopathy associated mutations descripted as pathogenic or likely pathogenic mutations of 63 patients in ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/) (C and D) Pie chart of targetability of mutations for 'wild-type (WT) to mutant' (C) or 'mutant to wild-type (WT)' (D) modeling by base editors (left), graphical presentation of target-ability of each base editor (right)

Figure. 2 Efficiency of base editing compared to HDR in hESCs (A) Graphical description of site of mutations in the epimerase and kinase domain of human GNE protein (hGNE2) (left), List of mutations generated by each base editor (right) (B to E) Target deep sequencing results showing overall base conversion rate of each sequence by CBE (C to T) and ABE (A to G) (left), Graphical presentation of on-target editing efficiency of indicated gene editing tools (right) for R160Q (B), I588T (C), V727M (D) and I329T (E), Target base and PAM sequence is colored in red and green (3'-5') respectively. (F) mRNA expression of *PPM1D* in wild-type (WT) and *TP53* Knockout (TP53 KO) hESCs transfected with sgRNA targeting CCR5 (sgCCR5) and indicated gene editing tools (Cas9, CBE, NG-CBE, ABE and NG-ABE) compared to control (Cont: sgCCR5 only), *, $p<0.05$ (G) Graphical presentation of percentage of live cell ratio of wild-type (WT) and TP53 KO hESCs after transfection of sgCCR5 and indicated gene editing tools, ****, $p<0.0001$ (H) Gene editing efficiency of indicated

gene editing tools in WT and TP53 KO hESCs on CCR5. Editing efficiencies are analyzed by target deep sequencing.

Figure. 3 Establishment of GNE myopathy disease models through base editors (A to D) Sanger sequence of mutants of R160Q (A), I329T (B), I588T (C) and V727M (D) compared to wild-type (WT), Exons coding UDP-GlcNAc-2-epimerase and ManNAc kinase domain are colored in blue and in red respectively. Target amino acids and bases are colored in red.

Figure. 4 Mutation-specific phenotype of GNE myopathy disease models (A) Microscope images (top) and fluorescent images for OCT4 (bottom) of hESCs of WT and GNE mutants, DAPI for nuclear counterstaining (B) Fluorescent images of hESCs of WT and GNE mutants stained with SNA (green) and DAPI for counter nuclear staining (C) Histogram and (D) bar graph of mean fluorescence intensity (MFI) of SNA in WT and GNE mutants, ***, $p<0.001$, ****, $p<0.0001$ (E) List of predicted severity, clinical significance and mean fluorescence intensity of GNE mutants, predicted severe, medium and mild were colored in red, yellow and green respectively. (F) Mean fluorescence intensity of SNA of indicated hPSCs lines, ***, $p < 0.001$

Figure. 5 Structural analysis predicts mutation-specific hyposialylation (A) Structures of tetrameric epimerase and dimeric kinase domains of human GNE (PDB ID 4ZHT and 2YHW, respectively) are shown. Target mutations R160Q (B) and I329T (C) are found in the epimerase domain, while I588T (D) and V727M (E) are in the kinase domain. Mutagenesis and rotamer selection were performed using PyMOL, and mutated residues are indicated in grey. Inter-residue interactions within a distance of 4.1 Å are indicated with dashed lines.

Figure. 6 Mutation-specific drug response of GNE myopathy disease models (A) Graphical description of biosynthesis of sialic acid and sialyation and roles of GNE

protein as UDP-GlcNAc-2-epimerase and ManNAc kinase (B) Fluorescence microscopic images of GNE mutant hESCs stained with SNA after indicated dose of ManNAc treatment at day 2 and day 6 (C) Flow cytometry for SNA in GNE mutant hESCs at 2 days after indicated dose of ManNAc treatment (D and E) Relative mean fluorescent intensity of SNA at 2 (D) and 6 (E) days after indicated concentration of ManNAc treatment

Figure. 7 GNE myopathy modeling in myocytes derived from GNE-hESCs (A) Schematic diagram of myogenic differentiation from hESCs (B) mRNA expression of typical markers for skeletal myocytes (left) and pluripotency (right) in WT hESC and myocytes derived from WT hESCs (C) Fluorescence microscope images for myosin heavy chain (MHC) in myocytes derived from WT hESCs, DAPI for nuclear counterstaining (D) Microscope images of myocytes derived from WT and GNE mutants (scale bar = 80 μm) (E) mRNA expression of MYH7 of hESCs, myoblast and myocytes derived from indicated GNE mutant (F) Mean fluorescence intensity of SNA in indicated WT and mutant myocytes (G) Relative mean fluorescence intensity of SNA from myocytes harboring I329T and V727M mutant at 2 days after indicated concentration of ManNAc treatment

Graphical abstract

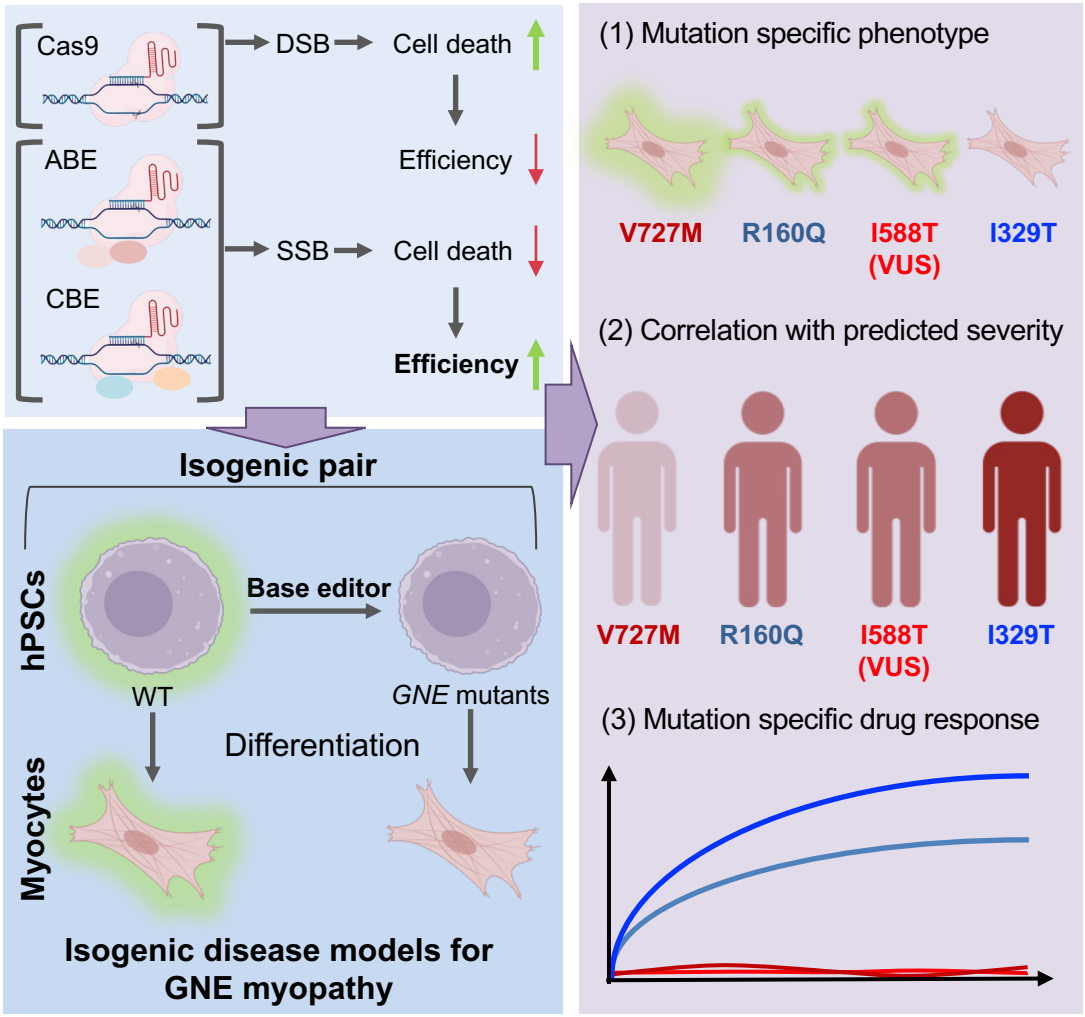


Figure. 1

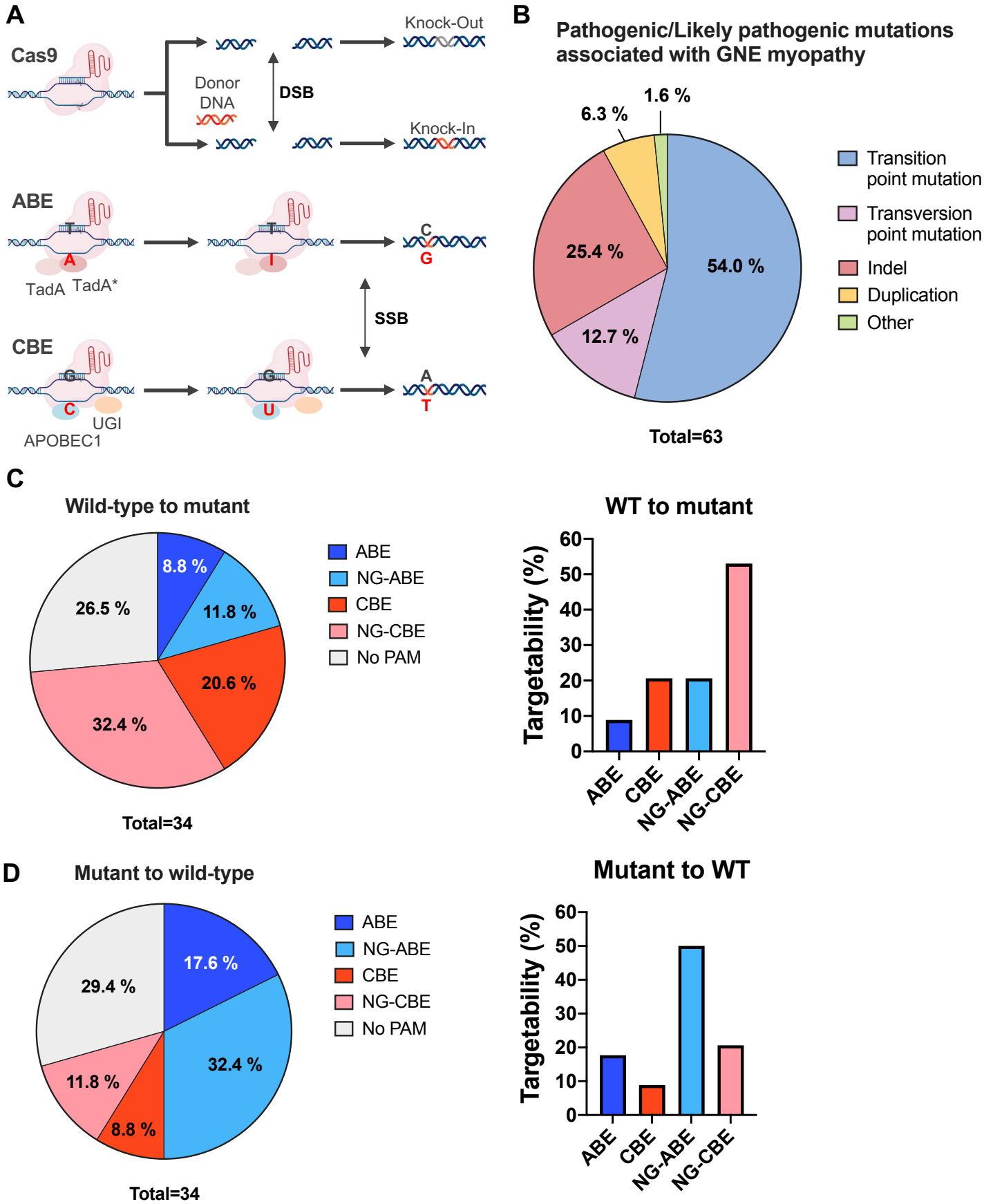
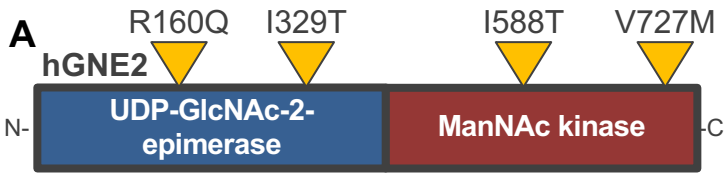
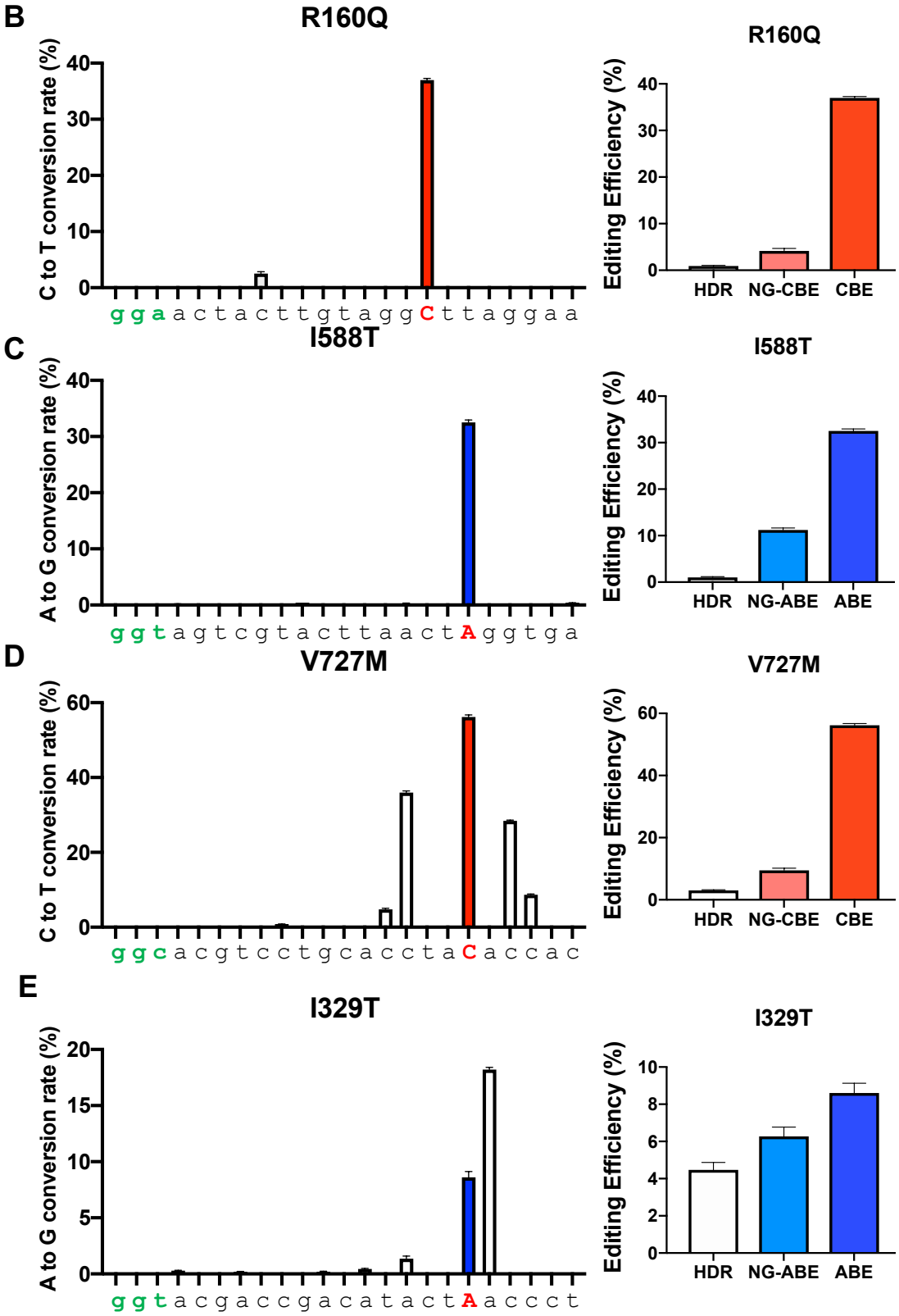
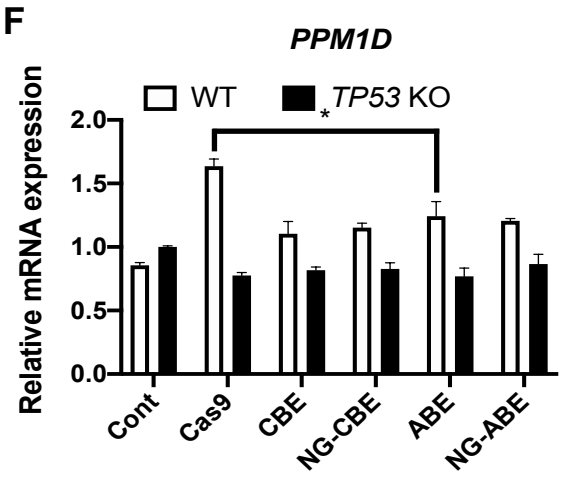

A
Cas9
Donor DNA
DSB
Knock-Out
Knock-In
ABE
A
I
C
G
TadA
TadA*
SSB
CBE
C
U
A
T
UGI
APOBEC1
B
Pathogenic/Likely pathogenic mutations associated with GNE myopathy
1.6 %
6.3 %
25.4 %
54.0 %
12.7 %
Transition point mutation
Transversion point mutation
Indel
Duplication
Other
Total=63
C
Wild-type to mutant
8.8 %
11.8 %
26.5 %
20.6 %
32.4 %
ABE
NG-ABE
CBE
NG-CBE
No PAM
Total=34
WT to mutant
Targetability (%)
60
50
40
30
20
10
0
ABE
CBE
NG-ABE
NG-CBE
D
Mutant to wild-type
17.6 %
29.4 %
32.4 %
11.8 %
8.8 %
ABE
NG-ABE
CBE
NG-CBE
No PAM
Total=34
Mutant to WT
Targetability (%)
60
50
40
30
20
10
0
ABE
CBE
NG-ABE
NG-CBE

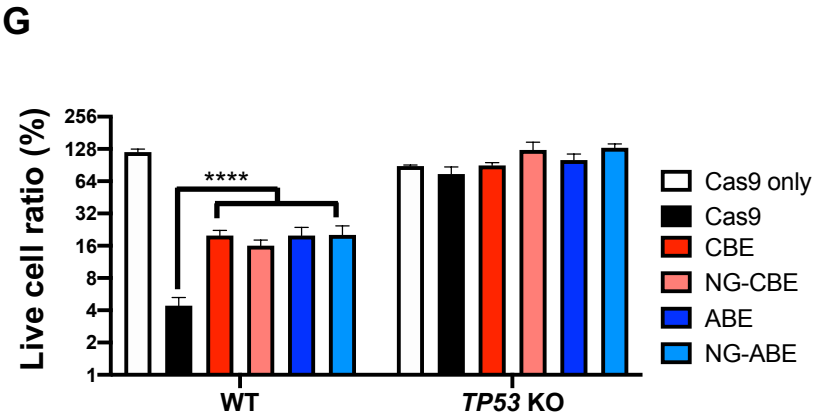
Figure. 2



	R160Q	I329T	I588T	V727M
Tool	CBE	ABE	ABE	CBE
Domain	Epimerase		Kinase	




F
PPM1D
WT
TP53 KO
*
Relative mRNA expression
2.0
1.5
1.0
0.5
0.0
Cont
Cas9
CBE
NG-CBE
ABE
NG-ABE


G
Live cell ratio (%)
256
128
64
32
16
8
4
2
1
****
WT
TP53 KO
Cas9 only
Cas9
CBE
NG-CBE
ABE
NG-ABE

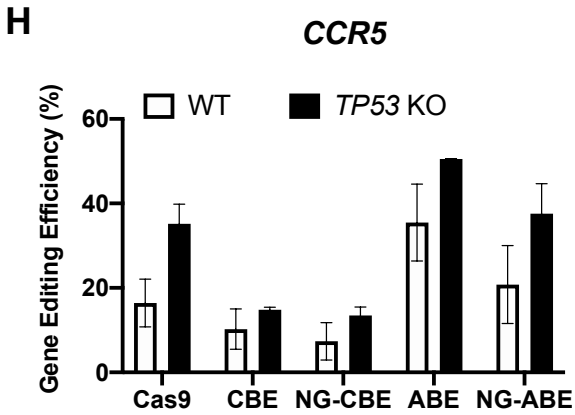

H
CCR5
WT
TP53 KO
Gene Editing Efficiency (%)
60
40
20
0
Cas9
CBE
NG-CBE
ABE
NG-ABE

Figure. 3

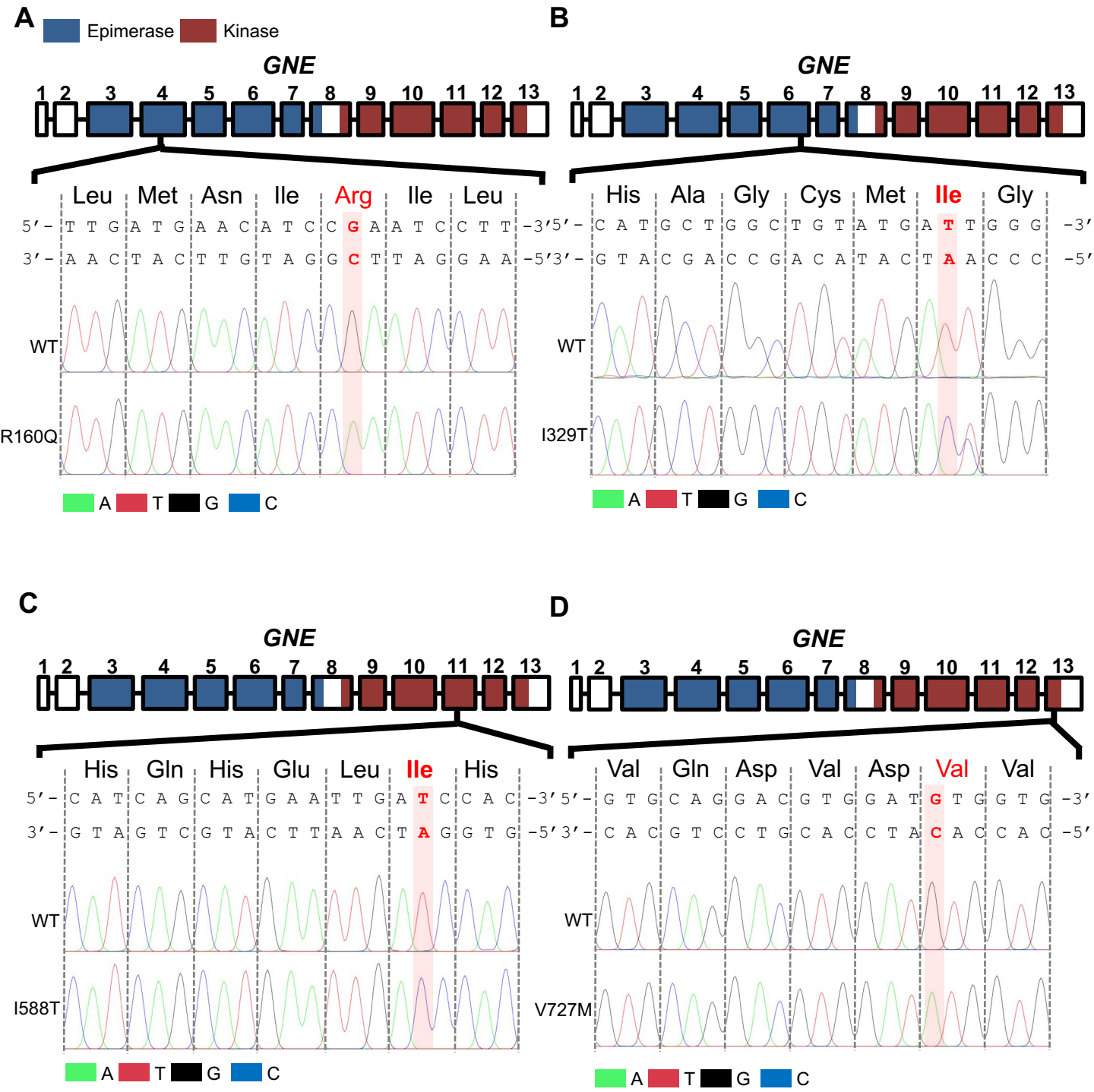
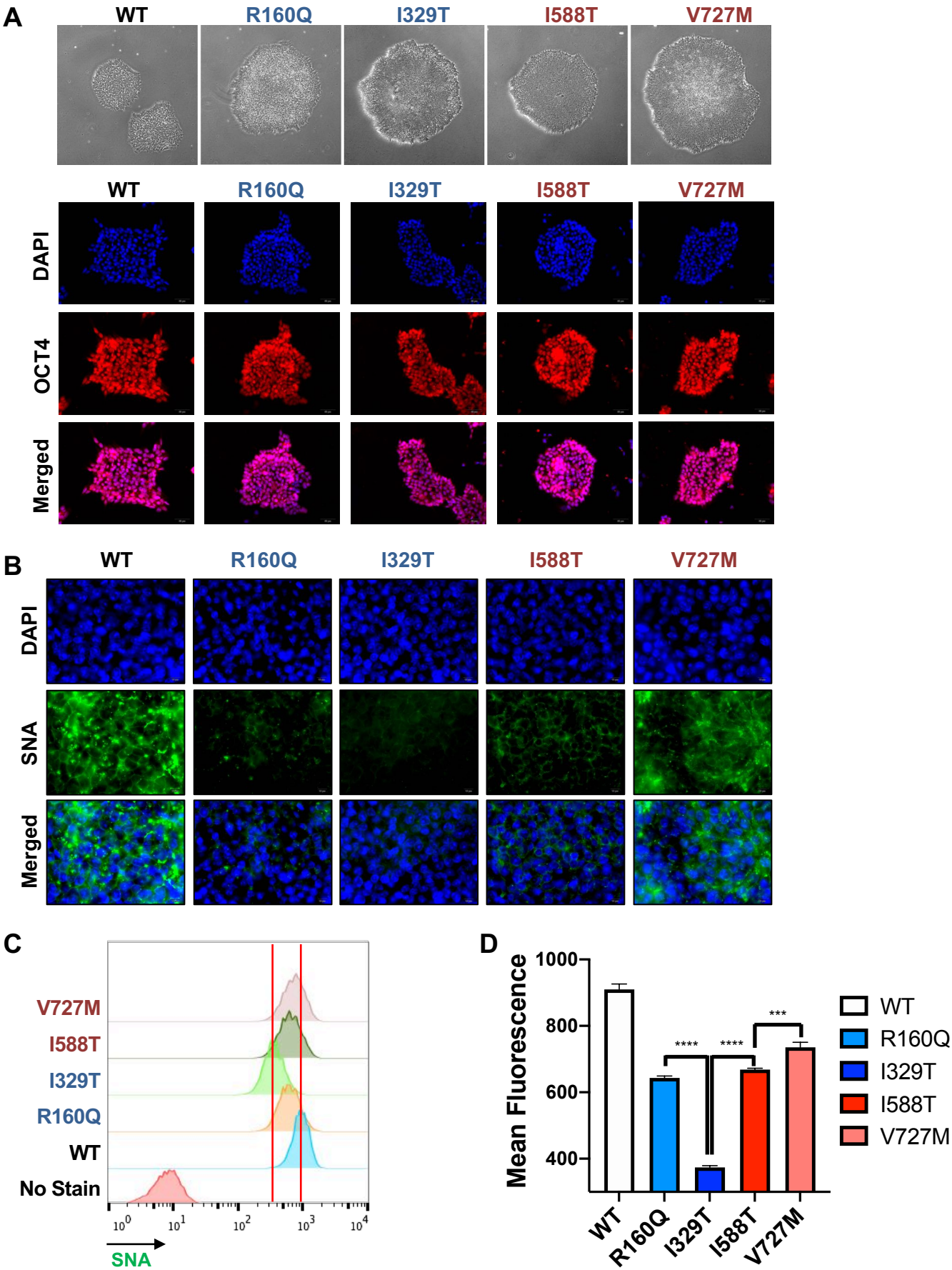

A
Epimerase
Kinase
GNE
1 2 3 4 5 6 7 8 9 10 11 12 13
Leu Met Asn Ile Arg Ile Leu
5′- T T G A T G A A C A T C C G A A T C C T T -3′
3′- A A C T A C T T G T A G G C T T A G G A A -5′
WT
R160Q
A T G C
B
GNE
1 2 3 4 5 6 7 8 9 10 11 12 13
His Ala Gly Cys Met Ile Gly
5′- C A T G C T G G C T G T A T G A T T G G G -3′
3′- G T A C G A C C G A C A T A C T A A C C C -5′
WT
I329T
A T G C
C
GNE
1 2 3 4 5 6 7 8 9 10 11 12 13
His Gln His Glu Leu Ile His
5′- C A T C A G C A T G A A T T G A T C C A C -3′
3′- G T A G T C G T A C T T A A C T A G G T G -5′
WT
I588T
A T G C
D
GNE
1 2 3 4 5 6 7 8 9 10 11 12 13
Val Gln Asp Val Asp Val Val
5′- G T G C A G G A C G T G G A T G T G G T G -3′
3′- C A C G T C C T G C A C C T A C A C C A C -5′
WT
V727M
A T G C

Figure. 4


A
WT
R160Q
I329T
I588T
V727M
DAPI
OCT4
Merged
B
SNA
C
No Stain
SNA
D
Mean Fluorescence
1000
800
600
400
****
****
***
WT
R160Q
I329T
I588T
V727M

E

Mutation (hGNE2)	Predicted severity					Clinical significance	Mean fluorescent intensity
	PolyPhen2	SIFT	Align	Pmut	Overall		
R160Q	Probably Damaging	Tolerated	35	Pathological	Medium	Pathogenic / Likely pathogenic	643.5
I329T	Probably Damaging	Damaging	65	Pathological	Severe	Likely pathogenic	374
I588T	Possibly Damaging	Damaging	65	Neutral	Medium	Uncertain significance	668.3
V727M	Probably Damaging	Damaging	15	Neutral	Medium	Conflicting (Pathogenic / Likely pathogenic / Likely benign / Uncertain significance)	735.3

Severe Medium Mild

F

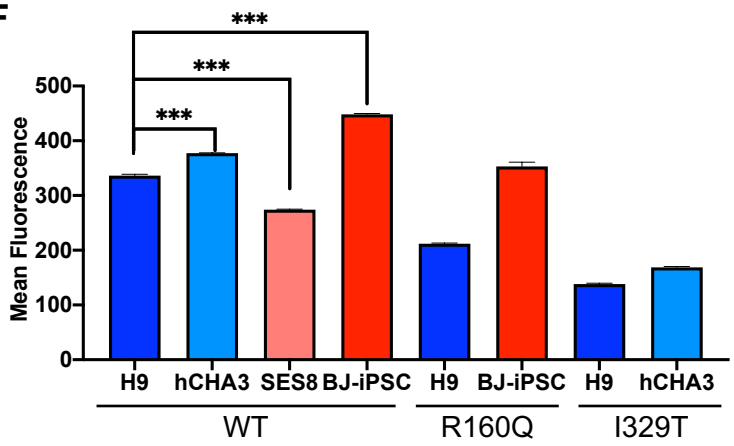


Figure. 5

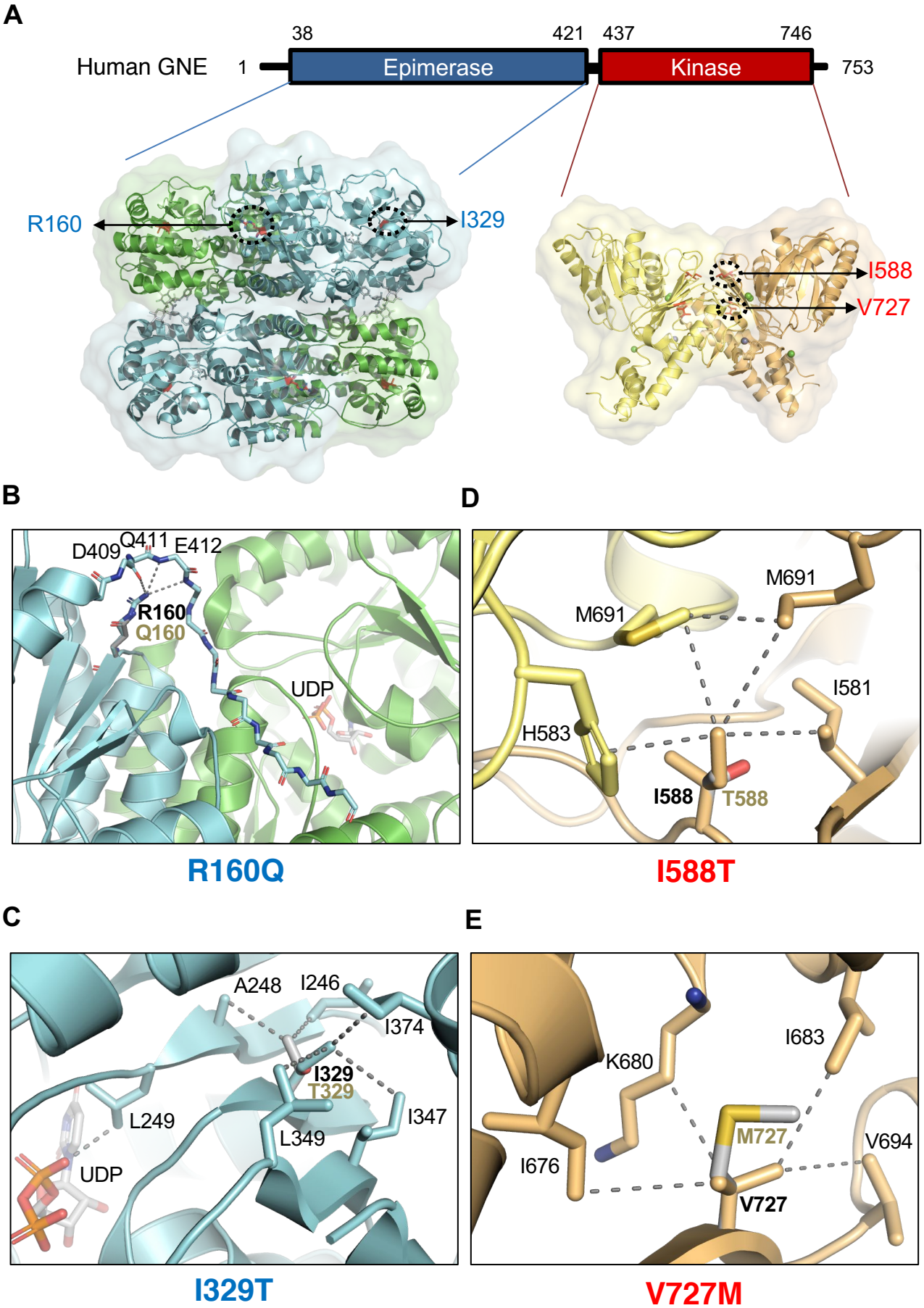
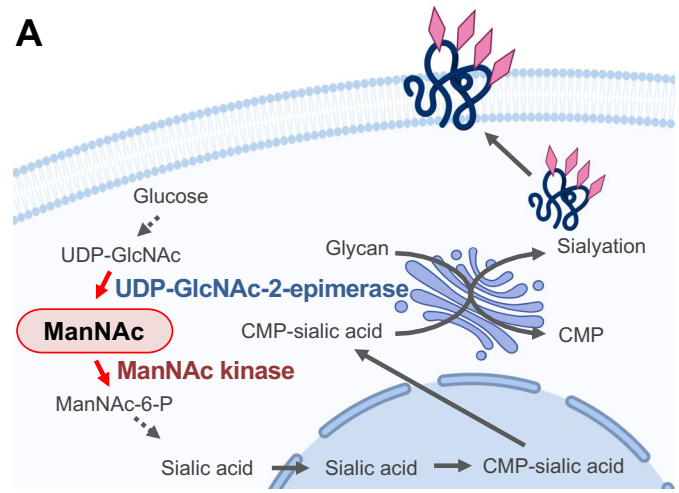
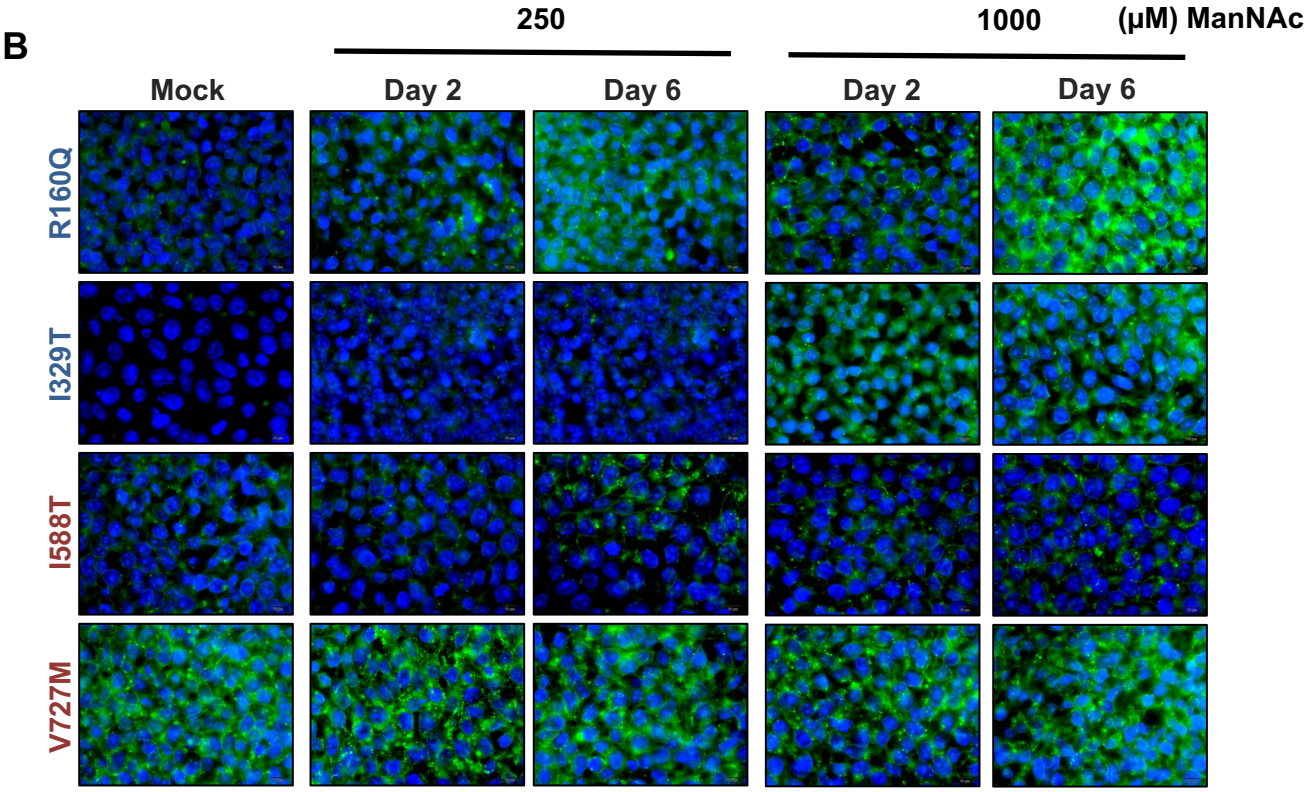

A
Human GNE
1
38
421
437
746
753
Epimerase
Kinase
R160
I329
I588
V727
B
D409
Q411
E412
R160
Q160
UDP
R160Q
D
M691
M691
I581
H583
I588
T588
I588T
C
A248
I246
I374
I329
T329
L249
L349
I347
UDP
I329T
E
K680
I683
M727
V694
I676
V727
V727M

Figure. 6


A
Glucose
UDP-GlcNAc
UDP-GlcNAc-2-epimerase
ManNAc
ManNAc kinase
ManNAc-6-P
Glycan
Sialyation
CMP-sialic acid
CMP
Sialic acid
Sialic acid
CMP-sialic acid


B
250
1000
(μM) ManNAc
Mock
Day 2
Day 6
Day 2
Day 6
R160Q
I329T
I588T
V727M

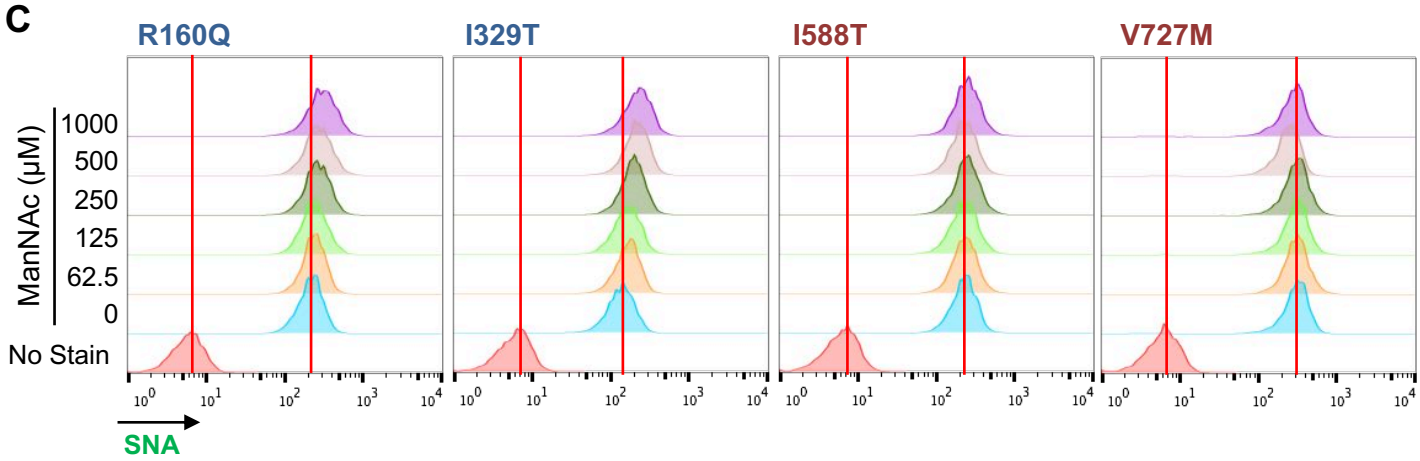

C
R160Q
I329T
I588T
V727M
ManNAc (μM)
1000
500
250
125
62.5
0
No Stain
SNA

Figure. 6

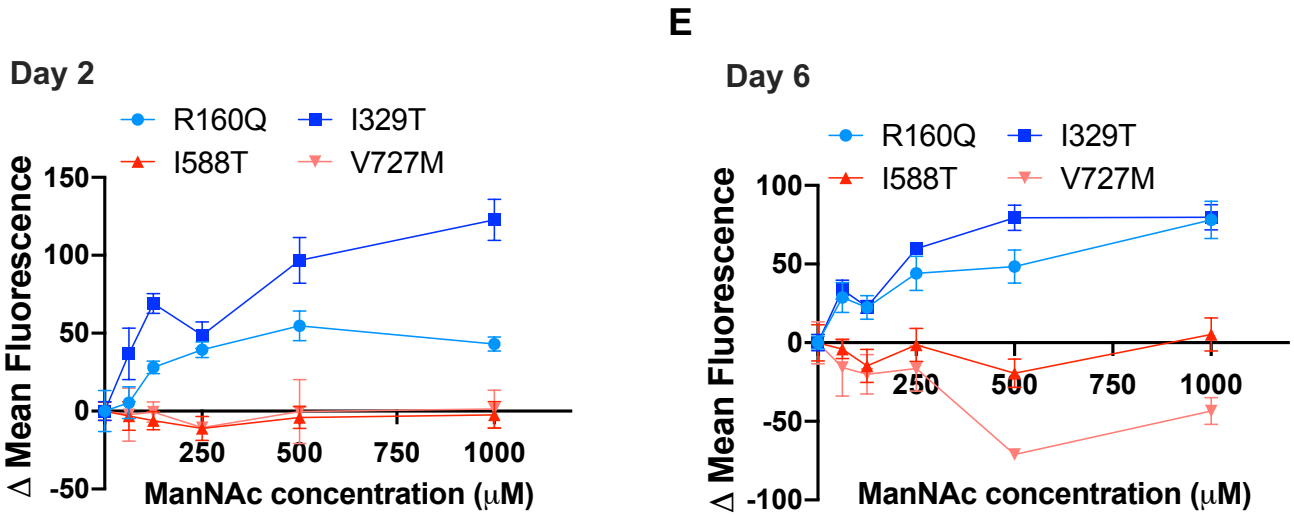
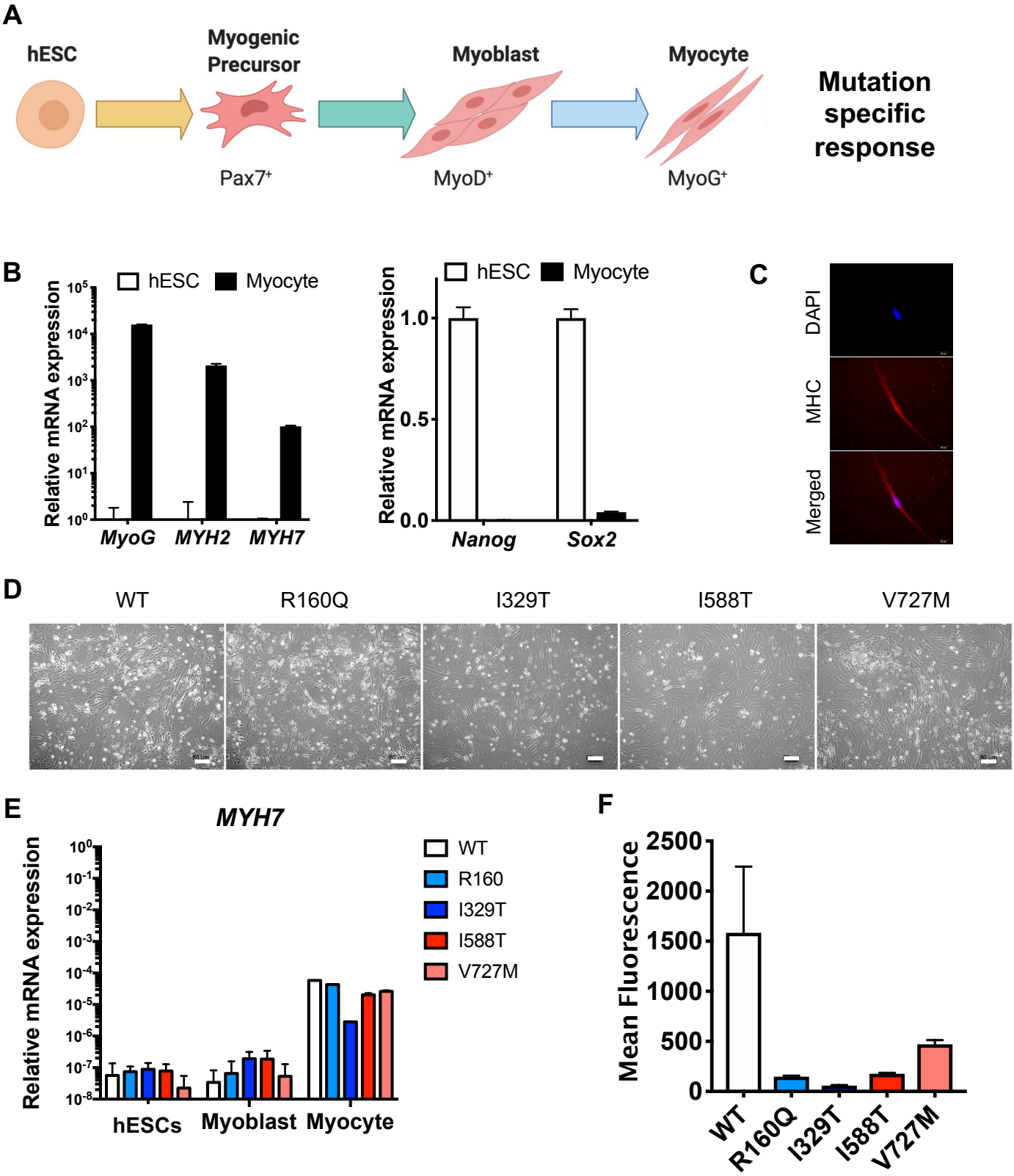
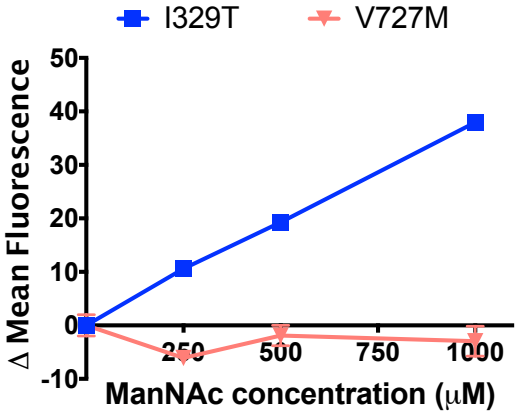

D
Day 2
R160Q
I329T
I588T
V727M
Δ Mean Fluorescence
150
100
50
0
-50
250
500
750
1000
ManNAc concentration (μM)
E
Day 6
R160Q
I329T
I588T
V727M
Δ Mean Fluorescence
100
50
0
-50
-100
250
500
750
1000
ManNAc concentration (μM)

Figure. 7


A
hESC
Myogenic Precursor
Myoblast
Myocyte
Pax7+
MyoD+
MyoG+
Mutation specific response
B
hESC
Myocyte
Relative mRNA expression
MyoG
MYH2
MYH7
Nanog
Sox2
C
DAPI
MHC
Merged
D
WT
R160Q
I329T
I588T
V727M
E
MYH7
Relative mRNA expression
WT
R160
I329T
I588T
V727M
hESCs
Myoblast
Myocyte
F
Mean Fluorescence
WT
R160Q
I329T
I588T
V727M

G


I329T
V727M
Δ Mean Fluorescence
50
40
30
20
10
0
-10
250
500
750
1000
ManNAc concentration (μM)