

# SCIENTIFIC REPORTS



OPEN

## CRISPR/Cas9-mediated genome editing in naïve human embryonic stem cells

Eva Z. Jacobs<sup>1</sup>, Sharat Warriar<sup>2</sup>, Pieter-Jan Volders<sup>1</sup>, Eva D'haene<sup>1</sup>, Eva Van Lombergen<sup>1</sup>, Lies Vantomme<sup>1</sup>, Margot Van der Jeught<sup>2</sup>, Björn Heindryckx<sup>2</sup>, Björn Menten<sup>1</sup> & Sarah Vergult<sup>1</sup>

The combination of genome-edited human embryonic stem cells (hESCs) and subsequent neural differentiation is a powerful tool to study neurodevelopmental disorders. Since the naïve state of pluripotency has favourable characteristics for efficient genome-editing, we optimized a workflow for the CRISPR/Cas9 system in these naïve stem cells. Editing efficiencies of respectively 1.3–8.4% and 3.8–19% were generated with the Cas9 nuclease and the D10A Cas9 nickase mutant. Next to this, wildtype and genome-edited naïve hESCs were successfully differentiated to neural progenitor cells. As a proof-of-principle of our workflow, two monoclonal genome-edited naïve hESCs colonies were obtained for *TUNA*, a long non-coding RNA involved in pluripotency and neural differentiation. In these genome-edited hESCs, an effect was seen on expression of *TUNA*, although not on neural differentiation potential. In conclusion, we optimized a genome-editing workflow in naïve hESCs that can be used to study candidate genes involved in neural differentiation and/or functioning.

Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of developing pre-implantation mouse or human blastocysts and are characterized by their ability for self-renewal and pluripotency<sup>1,2</sup>. Since it's shown that mouse pluripotent stem cells can also be derived from the mouse post-implantation blastocyst (epiblast stem cells, EpiSCs), two states of pluripotency are defined in mouse ESCs (mESCs) depicted as naïve and primed<sup>3–5</sup>. Naïve mESCs represent a more ground state of pluripotency in embryonic development and resemble cells of the pre-implantation epiblast *in vivo*. In contrast, EpiSCs resemble cells of the post-implantation epiblast, which is a later and already more advanced developmental stage, referred to as the primed state<sup>3,5</sup>. Naïve and primed mESCs show differences in several aspects such as colony morphology and active signalling-dependent pathways to maintain the pluripotent state, reflected in different transcriptomic and epigenetic characteristics<sup>3,5,6</sup>. In addition, naïve mESCs are able to survive single cell passaging, show a higher proliferation rate and have a higher clonogenic ability, compared to primed mESCs<sup>5</sup>.

Although human ESCs (hESCs) are also derived from the pre-implantation blastocyst, they differ from naïve mESCs in morphology, culture requirements, molecular profile, differentiation behaviour and clonogenicity<sup>5</sup>. Instead, these cells show more similarities to the post-implantation EpiSCs and thus conventional derived hESCs are considered as primed pluripotent stem cells<sup>5,7</sup>. In mouse, primed EpiSCs can be converted to the naïve state either by overexpression of pluripotency associated transcription factors or by applying naïve culture conditions. Hanna *et al.*<sup>8</sup> were the first to convert primed hESCs into mESC-like hESCs by ectopic induction of key pluripotency factors. Later, other groups defined different naïve media conditions to induce naïve pluripotency in conventional hESCs<sup>6,7,9–12</sup>.

The ability of human ESCs to differentiate into cells of the three germ layers makes them an attractive model system to study human development and disease *in vitro*<sup>13–17</sup>. To study gene function and the underlying regulatory pathways, genome editing has been shown to be of great avail<sup>18–20</sup>. The CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated proteins) Type II system is originally a prokaryotic RNA-mediated adaptive immune system used as a defense against invading foreign genomes and adapted for easy and efficient genome-editing in virtually every desired genome<sup>21,22</sup>. Almost any DNA sequence of interest can be targeted with the help of a customizable single guide RNA (sgRNA) containing 20 base pairs complementary to

<sup>1</sup>Center for Medical Genetics, Ghent University, Ghent University Hospital, Ghent, Belgium. <sup>2</sup>Ghent Fertility and Stem cell Team (G-FaST), Department for Reproductive Medicine, Ghent University Hospital, Ghent, Belgium. Correspondence and requests for materials should be addressed to S.V. (email: [sarah.vergult@ugent.be](mailto:sarah.vergult@ugent.be))

the target site. This target site needs an adjacent protospacer adjacent motif (PAM) for recognition by the sgRNA. The Cas9 endonuclease is guided by the sgRNA to the target locus and will induce a double-stranded break 3 base pairs (bp) upstream of the PAM sequence. Either these breaks will be repaired by non-homologous end joining (NHEJ) often leading to the formation of insertions/deletions (indels) near the breakpoints as this process is error-prone, or an exogenous DNA sequence can be inserted during homology-directed repair (HDR)<sup>21–23</sup>. To avoid undesired off-target effects, caused by sequence homology between the sgRNA and off-target sites, a D10A Cas9 nickase mutant with paired guide RNAs can be used to enhance specificity<sup>24</sup>.

Compared to mESCs, conventional primed hESCs are more difficult to genetically manipulate and render some practical obstacles for efficient genome-editing<sup>25</sup>. A bottleneck with the culture of conventional hESCs is their low proliferation rate<sup>7,26</sup>, their inability for single-cell passaging by trypsin digest<sup>7,25</sup> and the low single-cell cloning efficiency<sup>7</sup>. These characteristics complicate fast expansion of the culture and clonal isolation and selection after transfection<sup>25</sup>. This low single-cell cloning efficiency will also have an influence on the targeting efficiencies<sup>27</sup>. Another disadvantage of primed hESCs is the heterogeneity in differentiation potential<sup>6,26</sup>.

On the contrary, the naïve state enables fast proliferation and single-cell passaging. These are favourable characteristics to increase target efficiency, indicating this state is preferable for efficient genome-editing. It has indeed been reported that the naïve state is more amenable to genome-editing<sup>7,25–27</sup>. The study from Gafni *et al.*<sup>7</sup> also suggests that HDR will be more efficient in naïve ESCs due to their more open chromatin state.

In the present study, we show an optimized workflow for genome-editing of naïve hESC using the CRISPR/Cas9 system followed by neural differentiation. As a proof-of-principle we generated a knockdown with CRISPR/Cas9 of the long non-coding RNA (lncRNA) *TUNA* that has already been described to be involved in neural differentiation<sup>28</sup>. Since induced pluripotent stem cells (iPSCs) of patients with a (chromosomal) rearrangement involving a particular gene, are not always available or cannot be generated, this workflow can be used to study candidate genes involved in neural differentiation and/or functioning.

## Results

**Efficient genome-editing in naïve hESCs with CRISPR/Cas9 system.** To investigate the editing efficiency and characteristics of the CRISPR/Cas9 system in naïve hESCs, we nucleofected Cas9 nuclease as well as Cas9 nickase expressing plasmids targeting the coding gene *EMX1* and non-coding genes *TUNA* and *MEG3* in hESCs. The workflow is presented in Fig. 1. Each target site was successfully edited by both Cas9 proteins, however, we see big differences in editing efficiency, ranging from 1 to 19%.

For *TUNA*, the highest editing efficiency (19%) was accomplished with the Cas9 nickase, however, for *EMX1* and *MEG3* the highest editing efficiencies, respectively 8.4% and 7.8%, were generated using the nuclease combined with one sgRNA (Fig. 2). Based on these three loci and taken into account that there is a certain variation in sgRNA activity, both the wildtype Cas9 nuclease and mutant Cas9 nickase can edit the on-target site efficiently at similar editing rates.

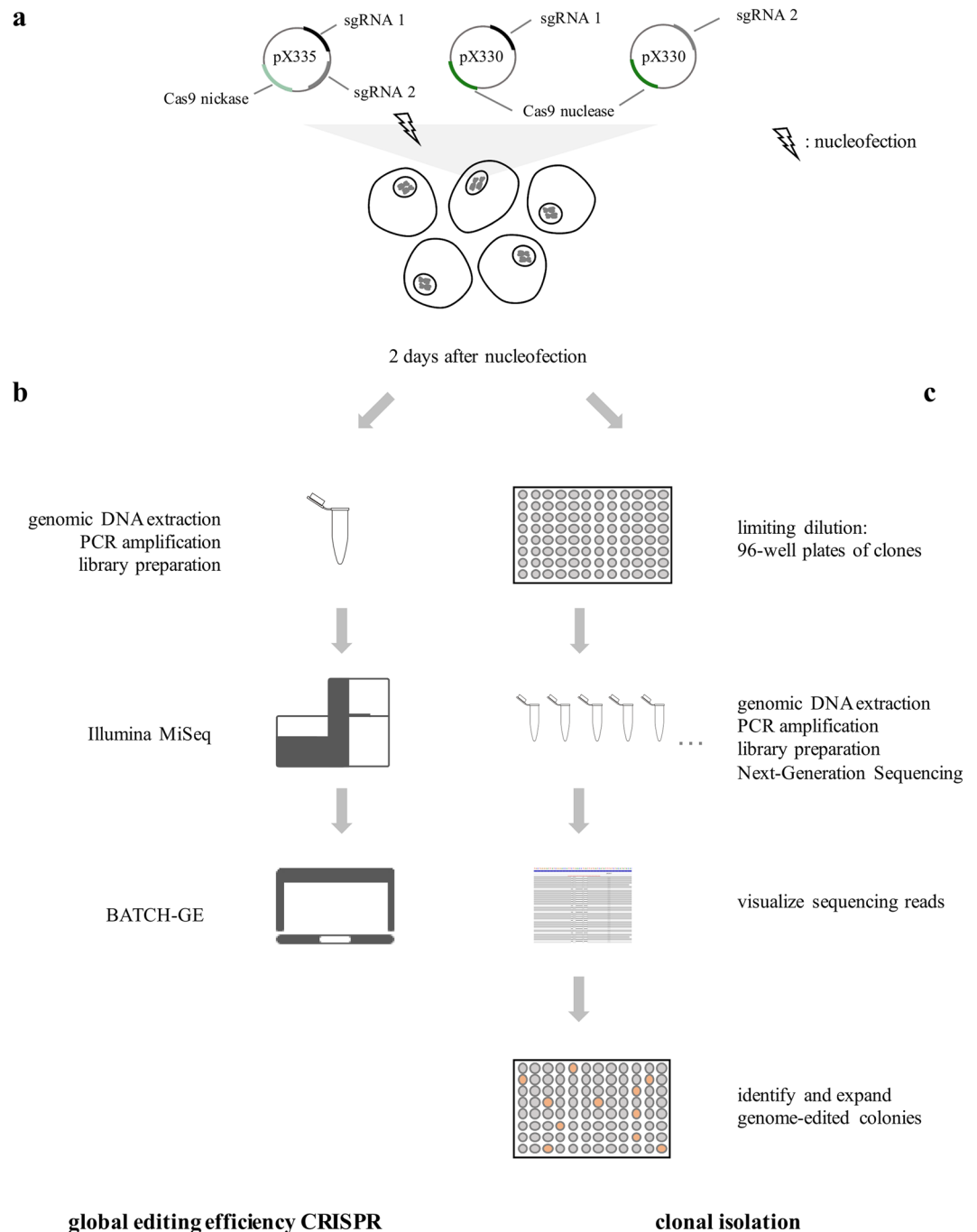
Mali *et al.*<sup>29</sup> described in HEK-293T cells that most indels resulting from NHEJ after Cas9 nuclease activity are present near the 3' end of the target site, particularly ~3–4 bp upstream of the PAM, consistent with the theoretical cleavage site of Cas9. Also in the naïve hESCs we see that most indels are located around this position after editing with Cas9 nuclease. For *TUNA*, *EMX1* and *MEG3* predominant indels of 1–6 bp are present 3–5 bp upstream of the PAM sequence (Fig. 3). Remarkably, the same predominant indels are also present in HEK-293T cells with similar editing efficiencies as in the naïve hESCs (Supplemental Figs 1 and 2).

For *MEG3*-sgRNA1 however, only a few variants are located near this expected cleavage site (Fig. 3), suggestive of a none to very low active sgRNA in the naïve hESCs. Nonetheless in HEK-293T cells, the same sgRNA also shows low editing efficiency (2.2%), but a predominant indel is present at 3 bp upstream of the PAM sequence (Supplemental Figs 1 and 2).

After editing with the Cas9 nickase, no predominant indels are seen. Here, approximately 60% of the indels for *TUNA*, *EMX1* and *MEG3* are present completely between or spanning both the cleavage sites and 70–90% of indels span at least one of the two sites (Fig. 4). This is also the case in HEK-293T cells (Supplemental Fig. 3). These data demonstrate that our CRISPRs, both the Cas9 nuclease as nickase, cut at the target site.

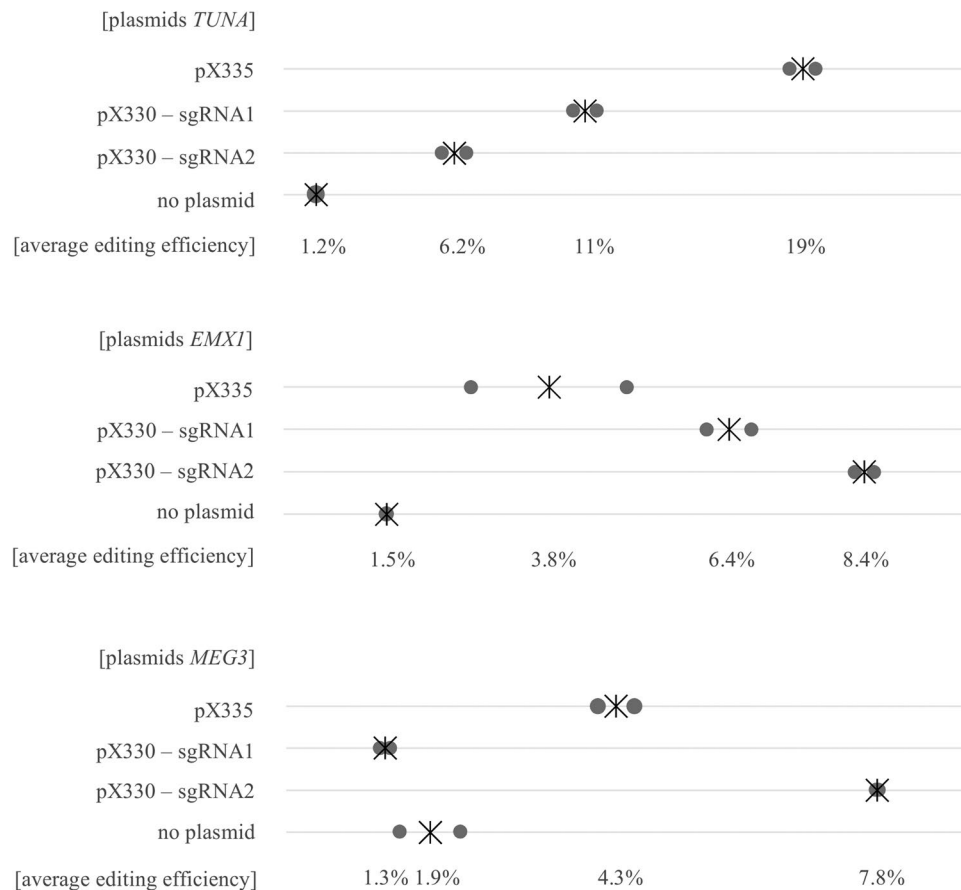
After screening the top 4 highest ranked off-target sites for each sgRNA, no off-targets were detected for both sgRNAs for *TUNA* and *MEG3*. However, indels were present in one off-target region of *EMX1*-sgRNA2 in the cells cleaved with the Cas9 nuclease (12% editing efficiency), but not with the double nicking system (Supplemental Fig. 4). Since these undesired indels are completely intronic (first intron of *EXOC6*), this off-target is unlikely to have a functional effect.

**Cas9 nickase generates larger indels compared to wildtype Cas9 nuclease.** Based on median deletion and insertion sizes, indels generated with the Cas9 nickase are larger compared to the wildtype Cas9 nuclease (Supplemental Table 1). This is due to the creation of two nicks in each other's neighbourhood compared to only one break point with the Cas9 nuclease. Since both sgRNAs for *EMX1* overlap six base pairs, making the two cleavage sites closer to each other compared to *TUNA* and *MEG3*, there is a smaller difference in deletion size between the Cas9 nuclease and nickase for this gene. For the nickase, a median size of 3 and 7 bp (rep 1 and 2) is seen and for both sgRNAs combined with Cas9 nuclease a median size of 6–8 and 3 bp (rep 1 and 2). The largest deletions were generated for *TUNA* (median size nickase 17 and 21 bp (rep 1 and rep2)) and the largest insertions for *MEG3* (median size nickase 5 and 25 bp (rep 1 and rep2)). Interestingly, NHEJ generates more deletions than insertions both after editing with the Cas9 nuclease as well as the nickase (Supplemental Table 1). Only for the Cas9 nuclease plasmids for *TUNA*, more insertions occur due to the predominant one base pair insertion at the cleavage site. The same indel characteristics are seen in HEK-293T cells (Supplemental Table 2).



**Figure 1.** Optimized workflow for efficient genome-editing of naïve human embryonic stem cells. **(a)** Two sgRNAs were designed for each gene to allow double nicking and cloned in the pX335 plasmid. Both sgRNAs were also individually paired with the Cas9 nuclease (pX330 plasmid). The CRISPR plasmids were transfected in the naïve hESCs using nucleofection at day two after splitting. **(b)** To determine the overall editing efficiency of the CRISPR, DNA was extracted from the pool of transfected cells two days after nucleofection. After subsequent PCR and massive parallel sequencing (Illumina MiSeq), the overall editing efficiency and the detected indel variants were quantified with BATCH-GE<sup>51</sup>. **(c)** To obtain monoclonal genome-edited colonies, clonal isolation was performed two days after nucleofection, using limiting dilution (0.5 cells/100  $\mu$ l). After 6–8 days, surviving clones were observed and DNA was extracted. After amplification and massive parallel sequencing of the target region, the sequencing reads were visualized in the Integrative Genomics Viewer (IGV, Broad Institute) to identify the genome-edited colonies.

**Successful generation of monoclonal genome edited naïve hESCs.** Lin *et al.*<sup>28</sup> reported *Tuna*, a lncRNA, to play an important role in pluripotency and neural differentiation of mESCs. As a proof-of-principle for our workflow, consisting of genome-editing in naïve hESCs followed by neural differentiation, we generated

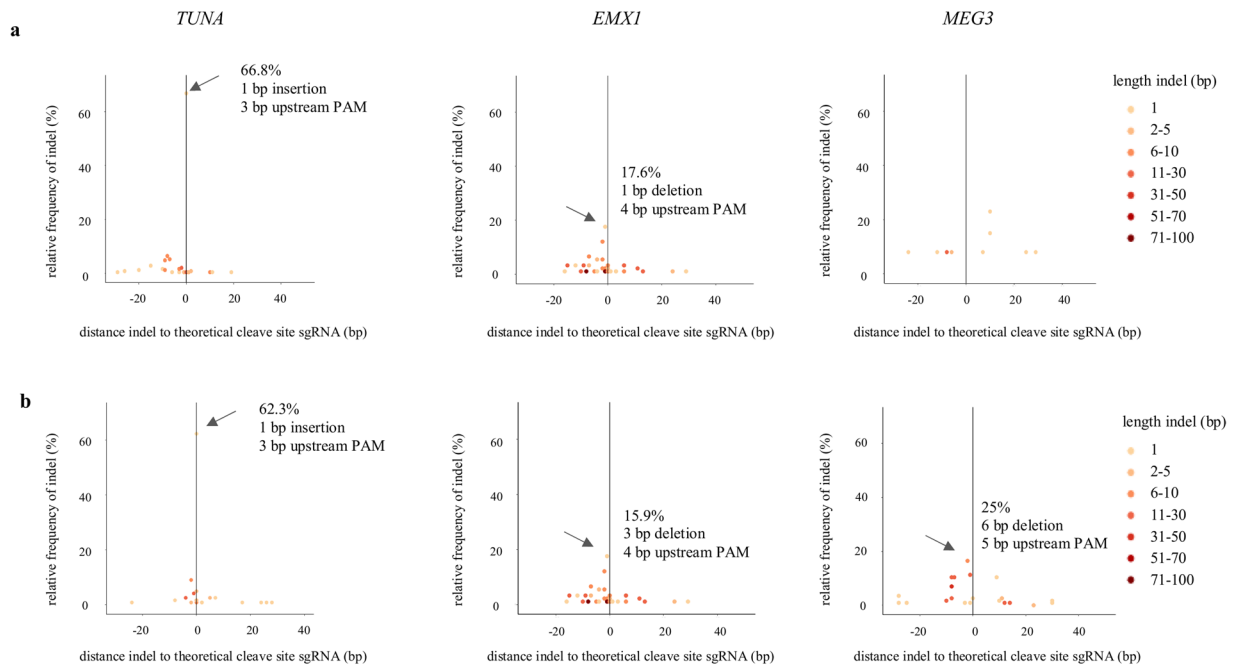


**Figure 2.** CRISPR/Cas9 works efficiently in naïve human embryonic stem cells. Editing efficiencies for *TUNA*, *EMX1* and *MEG3* after transfecting pX335, pX330-sgRNA1 or pX330-sgRNA2 plasmids in the naïve hESCs. As a control, mock-treated naïve hESCs (no plasmid) were used. Grey dots represent two biological replicates, the asterisk indicates the average of both replicates.

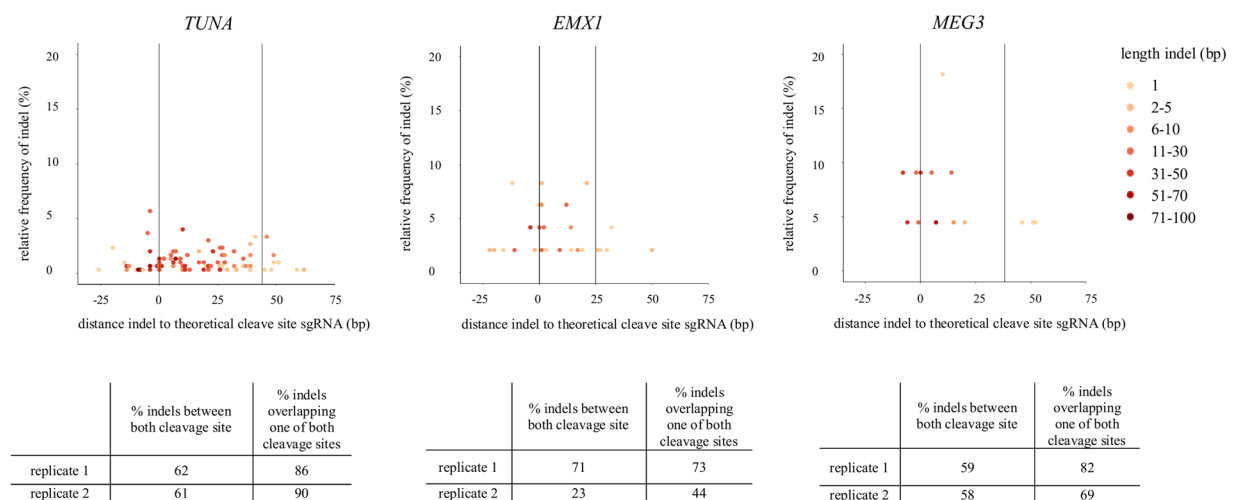
genome-edited hESC lines for *TUNA* using the double nicking system, as this system generated the highest editing efficiency combined with larger indels. The latter can be important as in contrast to coding genes, the functional effect of a certain indel is unpredictable for non-coding genes.

After transfection, the naïve hESCs were monoclonal isolated through limiting dilution in 96-well plates and a survival rate of 7% (7 out of 102 clones) was observed. The surviving clones were expanded and in the end, we obtained two monoclonal genome-edited *TUNA* colonies, one with a homozygous deletion and one with compound heterozygous deletions (Supplemental Fig. 5). To exclude that the latter is the result of a polyclonal colony, the clonal isolation was repeated and all resulting monoclonal colonies carry the same compound heterozygous deletions.

To confirm the naïve state of pluripotency after genome-editing, expression of transcription factors associated with either naïve or primed pluripotency was examined through qPCR in both genome-edited hESCs and compared to the wildtype naïve and primed hESCs (Supplemental Fig. 6a). We analyzed the expression of the naïve-specific pluripotency genes *KLF2* and *TCL1B*<sup>30</sup> in our primed and naïve stem cell lines and noticed a similar expression of both genes in the wildtype and genome-edited naïve hESCs and a trend of increased expression compared to primed hESCs. As expression analysis was performed on bulk cells, it is possible that certain naïve cells were already in a more differentiated state than others. This could explain the lower expression of *KLF2* and *TCL1B* in some biological replicates. No differences in expression of *NANOG* was seen between wildtype and genome-edited naïve hESCs, as well as between these naïve hESCs and their primed counterparts. This was however to be expected, as *NANOG* is a core-pluripotency factor with an important role in both the naïve and the primed state<sup>31</sup>. Next to this, in mESCs a variable expression of *Nanog* can be seen due to some heterogeneity in the mESC population<sup>32</sup>. This supports the assumption that at the time of RNA collection, a subpopulation of already differentiated cells was present in the naïve hESCs. De novo methyltransferase *DNMT3B*, a marker for the primed state, was clearly higher expressed in primed cells compared to all naïve cells<sup>30</sup>. Therefore, we can conclude that no differences in expression of naïve markers between the wildtype and genome-edited naïve hESCs were observed in our culture conditions. Moreover, there was no difference in morphology observed between the wildtype and genome-edited naïve hESC colonies (Supplemental Fig. 6b). These results indicate that the hESCs were still naïve after genome-editing.

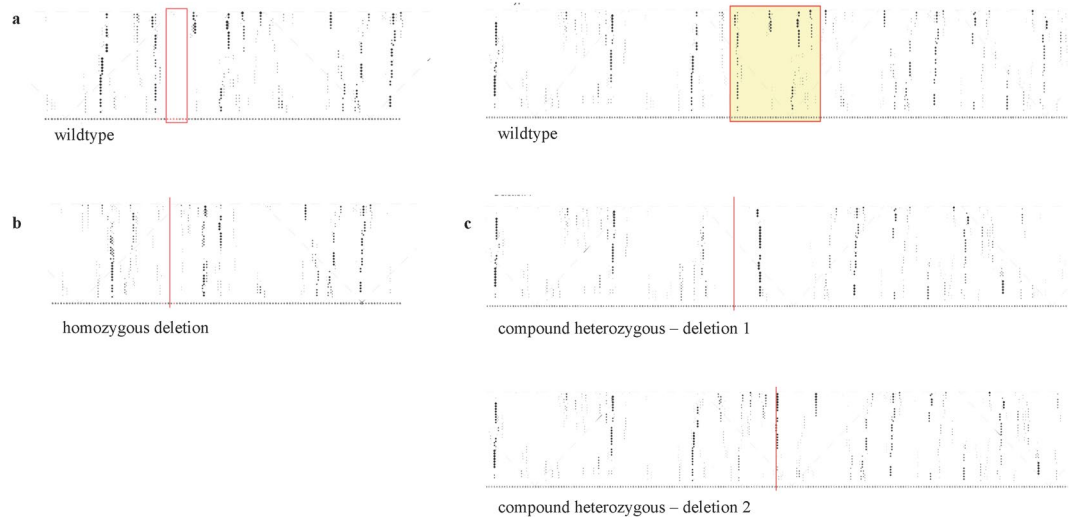


**Figure 3.** Cas9 nuclease generates predominant indels near the cleavage site. Distance of the start position of each insertion or deletion to the theoretical cleavage site (=3 bp upstream of the PAM sequence) is presented on the x-axis, while the relative frequency of the indel is presented on the y-axis for (a) nuclease – sgRNA1 and (b) nuclease – sgRNA2. Length of indels are represented by the different colors. Predominant indels are indicated with an arrow. Only 1 replicate is shown.

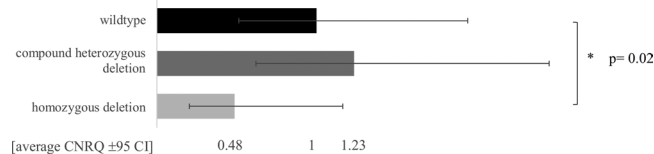


**Figure 4.** Majority of indels is present between both cleavage sites after editing with Cas9 nickase. Distance of the start position of each insertion or deletion to the theoretical cleavage site (=3 bp upstream of the PAM sequence) is presented on the x-axis, while the relative frequency of the indel is presented on the y-axis after editing with the Cas9 nickase. Length of indels are represented by the different colors. Only 1 replicate is shown. The majority of indels is present between the two theoretical cleavage sites (indicated by the black lines at position 0 (sgRNA1) and at positions 44 bp (*TUNA*), 25 bp (*EMX1*) and 38 bp (*MEG3*). Percentages of indels 1) within or spanning both the two cleavage sites and 2) indels spanning at least one of the two sites are shown below the graph. Replicate 2 of *EMX1* has an editing efficiency of 2.7% (compared to 5% for replicate 1), which can explain the lower percentage of indels located between the two cleavage sites.

**The generated deletions have an effect on both the secondary structure as well as the expression of *TUNA* in naïve hESCs.** In comparison with the consensus RNA structure, both the homozygous as well as the compound heterozygous deletions showed a modification of the secondary structure of *TUNA*, which was assessed in silico using RNAplfold<sup>33</sup>. Fig. 5a–c respectively show the pair probabilities for the secondary



**Figure 5.** Local structures in the targeted region of *TUNA* are altered after genome-editing. Using RNAplfold, local averaged pair probabilities are computed and represented in a dot plot<sup>33</sup>. (a) Wildtype local RNA structure of *TUNA* (left red box: position of homozygous deletion; right red box: position of the compound heterozygous deletions). (b) and (c) respectively show the local RNA structure of *TUNA* in the colony with the homozygous deletion and in the colony with the compound heterozygous deletions after CRISPR/Cas9 editing. The red lines indicate the position of the deletions.



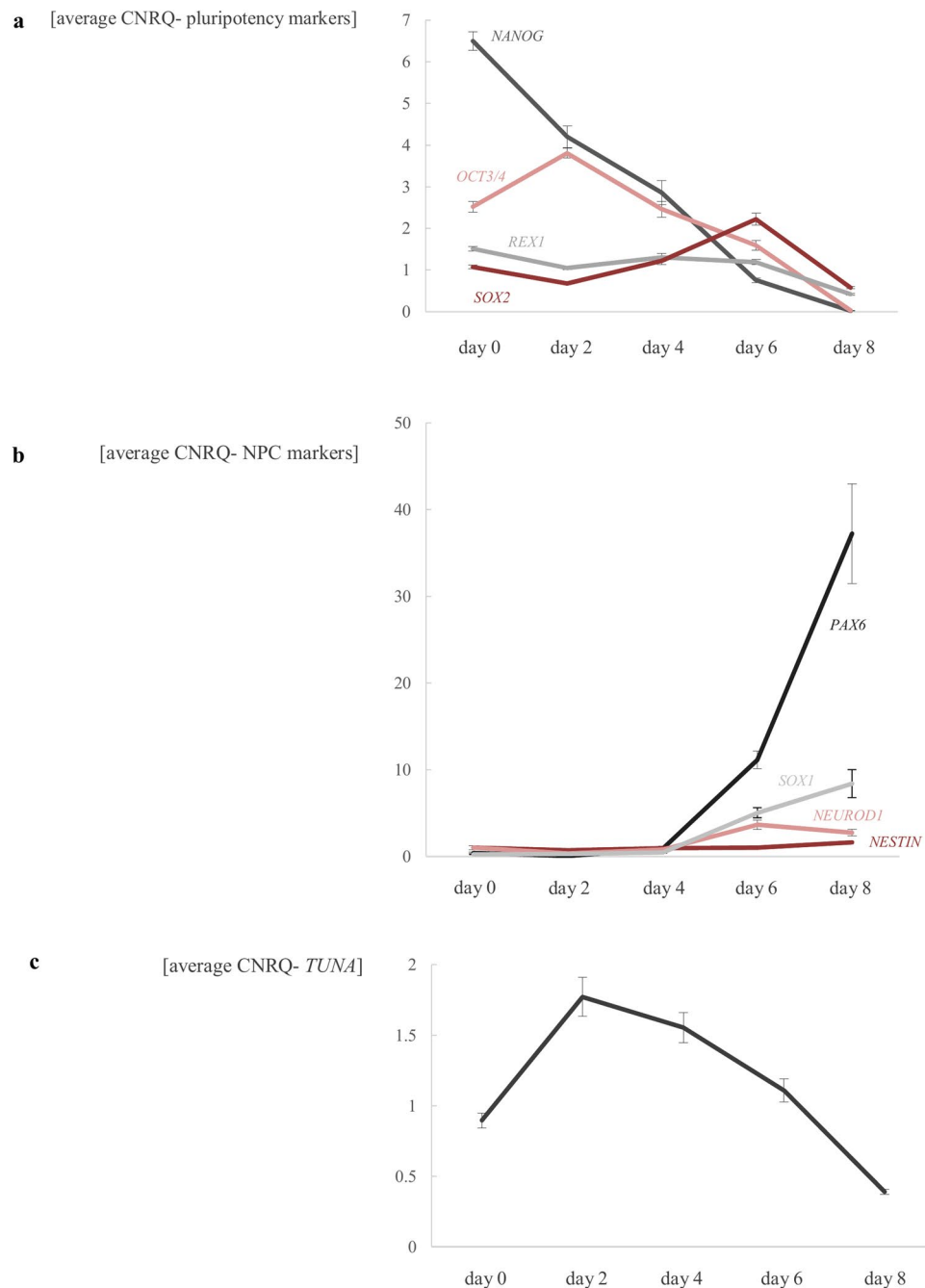
**Figure 6.** Altered expression of *TUNA* in genome-edited hESCs. Expression of *TUNA* in the wildtype, compound heterozygous and homozygous edited hESCs. Average Calibrated Normalized Relative Quantity (CNRQ) values  $\pm$  95% confidence interval (CI) is shown for 10 biological replicates of each genotype. The asterisk indicates significant difference in expression for *TUNA* (Wilcoxon rank sum test,  $\alpha = 0.05$ ).

structures of the wildtype RNA structure, the structure of the RNA with the homozygous deletion and the structure of the RNA with both deletions in the compound heterozygous edited cells. All three deletions show a local alteration of the secondary structure.

Also on expression level an effect was observed for the naïve hESCs with the homozygous *TUNA* deletion, namely a significant decrease in expression of *TUNA* ( $p$ -value = 0.02; Wilcoxon rank sum test,  $\alpha = 0.05$ ). This is not the case for the compound heterozygous deletion where no significant difference in expression was seen compared to the naïve wildtype hESCs ( $p$ -value = 0.19, Wilcoxon rank sum test,  $\alpha = 0.05$ ) (Fig. 6).

**Efficient differentiation of naïve hESCs to neural progenitor cells.** We set up a neural differentiation with STEMdiff Neural Induction Medium starting from naïve wildtype hESCs. RNA was isolated at different time points (day 0, 2, 4, 6 and 8 of neural induction) from three independent biological replicates. As shown in Fig. 7a, the expression of core-pluripotency genes *OCT3/4*, *REX1*, *SOX2* and *NANOG* decreases in time. At the same moment, the expression of the neural progenitor markers *PAX6*, *NESTIN*, *NEUROD1* and *SOX1* increases (Fig. 7b), indicating that the neural differentiation was successful.

***TUNA*, a lncRNA involved in neural differentiation?** We see a similar expression pattern of *TUNA* during neural differentiation as described by Lin *et al.*<sup>28</sup>, namely, an increase during the first two days of neural induction (Fig. 7c). To test whether knockdown of *TUNA* would interfere with the neural differentiation, as seen by Lin *et al.* after knockdown with shRNAs, the expression of the pluripotency and neural markers was monitored during differentiation of both wildtype as well as the compound heterozygous and homozygous edited stem cell lines of *TUNA* (Fig. 8, Supplemental Fig. 7). From ten independent biological replicates RNA was extracted from the naïve hESCs (day zero) and at day six after neural induction. Because ESCs are often subject to high variability in expression between biological replicates, we standardized our expression data as described by Willems *et al.*<sup>34</sup>. During neural differentiation, a decrease in expression of *TUNA* was observed in both genome-edited stem cell lines, which was significant for the homozygous deletion ( $p$ -value =  $4.10 \cdot 10^{-6}$ , false discovery rate (FDR) = 0.05). In mESCs, *Nanog*, *Sox2* and *Fgf4* are described as direct target genes of *Tuna* and decreased expression of these genes was observed after knockdown of *Tuna* with shRNAs<sup>28</sup>. This is not the case for *SOX2*, *FGF4* and *NANOG* in both

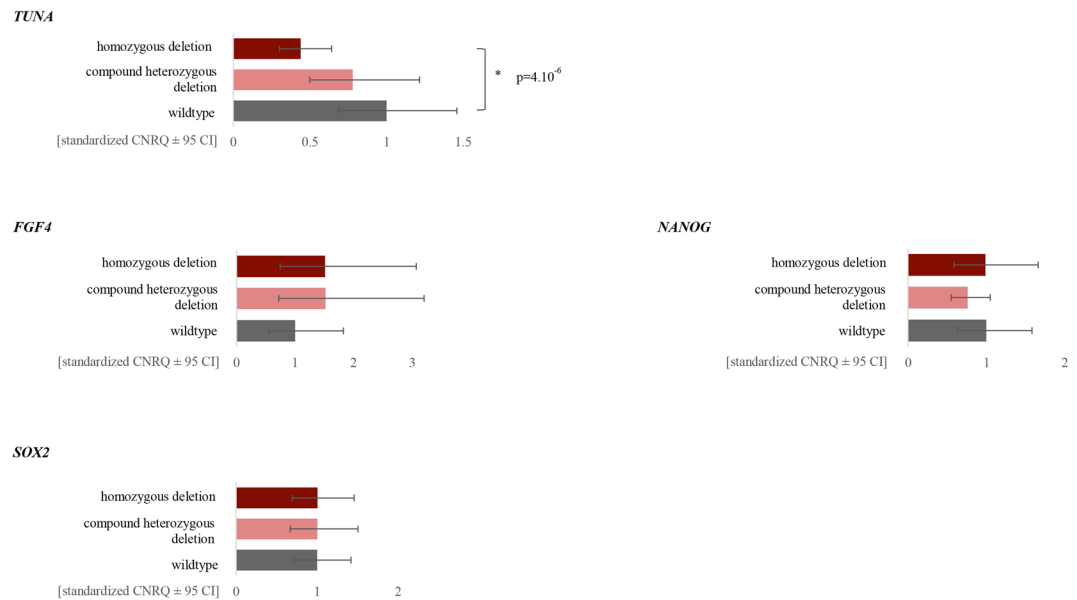


**Figure 7.** Successful differentiation of wildtype naïve hESCs to neural progenitor cells (NPCs). **(a)** Pluripotency markers (*OCT3/4*, *REX1*, *SOX2* and *NANOG*) decrease during neural differentiation of wildtype naïve hESCs, while the expression of neural progenitor cell markers (*PAX6*, *NESTIN*, *NEUROD1*, *SOX1*) increases **(b)**. Expression of *TUNA* increases during first two days **(c)**. RNA was sampled at day 0, 2, 4, 6 and 8 of neural induction. Average CNRQ  $\pm$  standard error of three independent biological replicates is shown.

*TUNA* genome-edited hESCs. We did not observe a significant decrease in expression of the core-pluripotency genes *REX1* and *OCT3/4*. In addition, no decrease in expression of neural progenitor markers *PAX6*, *NESTIN*, *NEUROD1* and *SOX1* was noticed during differentiation in the *TUNA* genome-edited hESCs. Remarkably, a significant increased expression of *NESTIN* is observed in the compound heterozygous stem cell line ( $p = 1,99.10^{-5}$ ,  $FDR = 0.05$ ) (Fig. 8, Supplemental Fig. 7).

## Discussion

The naïve state of hESCs can offer a better alternative for efficient genome-editing as there is no labor-intensive maintenance of the stem cell culture and exhibit increased single-cell cloning efficiency, compared to primed hESCs<sup>7</sup>. Based on the three tested loci, the CRISPR/Cas9 system works efficiently in naïve hESCs both with the



**Figure 8.** Generated indels in conserved region of *TUNA* don't cause a significant effect on pluripotency and neural differentiation. Expression of *TUNA* and its downstream targets (*FGF4*, *NANOG*, *SOX2*) in the wildtype, compound heterozygous and homozygous edited hESCs during neural differentiation. Average standardized (according Willems *et al.*<sup>34</sup>) CNRQ values  $\pm$  95% confidence interval (CI) are shown for 10 biological replicates of each genotype. RNA was collected of the naïve hESCs (day 0) and at day 6 of neural induction. Significant p-values are indicated with an asterisk (Wilcoxon rank sum test with BH correction, FDR = 0.05).

Cas9 nuclease as well as with the Cas9 nickase (up to 19% editing efficiency). The use of a plasmid based CRISPR approach already proved its effectiveness in pluripotent stem cells and the observed editing efficiencies are similar as described in literature for iPSCs (1–25% NHEJ in iPSCs, also without any selection steps<sup>29,35,36</sup>). Higher efficiencies are noticed when using a selection marker (up to 79% in iPSCs<sup>37–39</sup>). Next to this strategy, delivery of ribonucleoproteins (RNPs) is reported as more efficient and with reduced off-target effects compared to plasmids expressing the Cas9 nuclease<sup>40,41</sup>. For example, Kim *et al.* reported an editing efficiency of 66% when using Cas9 RNP electroporation in H9 hESCs<sup>42</sup>. Therefore, it is of great interest to optimize this system in naïve hESCs and compare the editing efficiencies with the Cas9 plasmids. When only considering the off-target effects, these will also be minimized when using the double nicking system with the plasmid approach.

Although we only screened four predicted off-target sites, for one sgRNA of *EMX1* one of the screened sites showed cleavage with the wildtype Cas9 nuclease and not with the Cas9 nickase. This confirms the higher specificity of the double nicking system, as described by Ran *et al.*<sup>24</sup>. It also stresses the importance of screening for off-target sites when using CRISPR-Cas9, especially for certain applications (e.g. gene or cell therapy). Therefore, one can screen all predicted off-targets sites or identify the off-targets empirically using methods like GUIDE-Seq<sup>43</sup>, Digenome-Seq<sup>44</sup>, SITE-Seq<sup>45</sup> and CIRCLE-Seq<sup>46</sup>.

In contrast to Miyaoka *et al.*<sup>36</sup>, for our sgRNAs we don't see an effect of the chosen cell type on the efficiency of gene-editing (naïve hESCs and HEK-293T), however we do observe an influence of the used sgRNA and chosen target locus.

For the three observed loci, the Cas9 nickase doesn't generate predominant indels like the Cas9 nuclease does, but the indels are larger. The latter can be beneficial to perturb the function of a lncRNA, which is not as straightforward as for protein coding genes. It's described that small deletions within lncRNAs are unlikely to disrupt their biological activity and hence larger deletions are preferred<sup>47</sup>.

Lin *et al.*<sup>28</sup> described a highly-conserved region of ~200 bp in the largest exon of *TUNA* that contains a functional motif. Therefore, we chose to target this motif with CRISPR/Cas9 to disturb its function. Because of the favorable characteristics of the naïve hESCs, no problems were experienced with the monoclonal isolation and we rapidly obtained two genome-edited hESCs colonies with (a) homozygous/compound heterozygous deletion(s) in this motif.

The function of lncRNAs is probably related to their higher-order structures through interactions with partner proteins<sup>48</sup>. After in silico assessment with RNAplfold we see that both deletions cause a change in the local secondary structure of *TUNA*, compared to the wildtype structure. Since it's shown by Lin *et al.*<sup>28</sup> that *Tuna* modulates its function through its conserved functional domain by interacting with the RNA-binding proteins PTBP1 and hnRNP-K (but not NCL), it can be anticipated that the obtained alterations in secondary structure can have a functional effect.

Recently, Lee *et al.*<sup>49</sup> showed that naïve hESCs are intrinsically predisposed to differentiate towards the neural fate and have impaired mesoderm and endoderm differentiation capacity. They compared the neural differentiation potential of naïve and primed hESCs and noticed significantly more NPCs generated from the naïve hESCs. Our group also successfully differentiated wildtype and genome-edited naïve hESCs to NPCs, which has great potential for studying neurodevelopmental disorders. After differentiating the two genome-edited *TUNA*



hESCs to NPCs, both showed a decrease in expression of *TUNA* during neural differentiation (significant for the homozygous deletion). Large differences between biological replicates can be noticed and hence the method by Willems *et al.* to standardize our expression data was implemented.

A possible explanation why we don't observe the same phenotype as Lin *et al.* might be the use of different mechanisms to generate the gene knockdown: shRNA interferes at the (post)transcript-level and CRISPR/Cas9 at the DNA-level. It might be that the generated indels with CRISPR/Cas9 are too small to generate a sufficient knockdown and subsequent a functional effect. Lin *et al.* achieves a knockdown of almost 90% in the mESCs, while only the hESCs with a homozygous deletion shows a knockdown of ~50%. It might be that only these hESC clones survived because these indels have limited effect on their pluripotency. It would be interesting to investigate whether genome-edited colonies with larger indels in the functional domain or even a deletion of the entire lncRNA will have a functional effect on pluripotency and neural differentiation and will repeat the phenotype as described in Lin *et al.*

Combining pluripotent stem cells and genome-editing, in particular the CRISPR/Cas9 technology with its broad applications, will be a powerful tool for future research. The use of naïve hESCs in comparison to primed hESCs has several advantages, therefore, we optimized a general workflow for the CRISPR/Cas9 system in naïve human embryonic stem cells. Here we showed that the system works efficiently. Next to this, we also established neural differentiation starting from naïve hESCs. In future research, this workflow can be used to study candidate genes, both coding as well as noncoding, involved in neural differentiation.

## Material and Methods

**Ethical permission.** The derivation of naïve human embryonic stem cells for this study was approved by Ghent University Hospital Ethical Committee (2013/822). Genome-editing of naïve human embryonic stem cells with CRISPR/Cas9 was approved by Ghent University Hospital Ethical Committee (2017/0923). All experiments were performed in accordance with relevant guidelines and regulations.

**Generation and culture of hESC.** In-house derived primed and naïve hESCs were cultured as described<sup>6</sup>. The naïve hESCs were maintained in 20% O<sub>2</sub> and 5% CO<sub>2</sub> conditions instead of hypoxic conditions (5% O<sub>2</sub>) and 6% CO<sub>2</sub>. In brief, the naïve hESCs were cultured on inactivated Mouse Embryonic Fibroblasts (MEFs) (GSC-6201G, Tebu-Bio) in naïve hESC medium consisting of KO-DMEM (10829018, Invitrogen) containing 20% knockout serum replacement (10828028, Invitrogen), 1% non-essential amino acids (11140050, Invitrogen), 1% Penicillin/Streptomycin (15140122, Invitrogen), 0.1 mM L-glutamine (25030024, Invitrogen) and 0.1 mM β-mercaptoethanol (31350010, Invitrogen). This naïve hESC medium is supplemented with 12 ng/ml bFGF (100-18B, Peprotech), 1000U recombinant human LIF (L5283, Sigma) and the small molecules 1 μM PD0325903 (13034, Sanbio), 3 μM CHIR99021 (1386, Axon Medchem), 10 μM Forskolin (F6886, Sigma) and 50 ng/ml ascorbic acid (A8960, Sigma). The naïve colonies were passaged as single cells using 0.05% trypsin/EDTA (25300054, Invitrogen) every three days and were re-plated on inactivated MEFs.

**Design and construction CRISPR plasmids.** CRISPR plasmids were constructed for three target genes: *TUNA*, *MEG3* and *EMX1*. sgRNAs were designed in the largest exon of *TUNA*, which is highly conserved<sup>28</sup>, and in the promoter region of *MEG3*. For *EMX1*, sgRNAs 9 and 10 from Ran *et al.*<sup>24</sup> were used (Supplemental Table 3). For each gene, two guide RNAs were designed to allow double nicking using the CRISPR/Cas9 design tool developed by the Zhang lab at MIT (crispr.mit.edu). For the selected sgRNA target sequence, the recognition site for BbsI was added to the sequence and the single stranded DNA molecules were ordered at Integrated DNA Technologies (IDT, Belgium): 5'-CACC-(G)-N<sub>20</sub>-3' (forward) and 3'-(C)-N<sub>20</sub>-CAAA-5' (reverse). If the target sequence does not start with a G, add a G and a C to respectively the forward and reverse oligonucleotide. Both oligonucleotides were resuspended to a concentration of 100 μM, mixed in equimolar concentration, heated for 2 minutes at 95 °C, cooled down to room temperature and diluted to 0.04 μM.

**Cas9 nuclease.** The two sgRNAs used with the Cas9 nickase were individually cloned into the pX330 plasmid (Addgene Plasmid #42230). The vector was digested with BbsI (R0539S, Bioké) and annealed oligonucleotides were ligated into the digested plasmid with the DNA ligation kit (TB6023, Westburg) according manufacturer's instructions.

**Double nicking D10A Cas mutant.** Both sgRNAs were first individually cloned into the pX335 plasmid (Addgene Plasmid #42335). To enhance transfection efficiency, both guide RNAs were then cloned in the same pX335 plasmid. Therefore, the U6-sgRNA2 cassette was amplified using primers containing the XbaI and KpnI sites and cloned into the pX335 plasmid containing the U6-sgRNA1 cassette.

**Determine off-target sites.** The CRISPR/Cas9 design tool from the Zhang lab gives a ranked list of possible off-target sites for each sgRNA, based on the number and position of mismatches. For each sgRNA the top four ranked off-targets were screened for presence of undesired indels by next-generation-sequencing. Supplemental Table 4 contains the primer sequences and the genomic positions (hg19).

**Transfection of CRISPR plasmids into naïve human embryonic stem cells.** Naïve hESCs cultured on MEFs were dissociated into single cells using 0.05% Trypsin/EDTA. 2.10<sup>5</sup> cells were nucleofected with 0.4 μg plasmid in 20 μl P3 Primary Cell Nucleofector Solution with program DN100 (4D Nucleofector, Lonza), according to the manufacturer's instructions. For every transfection, a mock-treated control was included (cells were exposed to the Nucleofector solution and subjected to the Nucleofection procedure, but in the absence of DNA). The transfected cells were subsequently plated onto fresh inactivated MEFs and cultured in naïve human embryonic stem cells medium, supplemented with 10 μM Rock inhibitor (S1049, Selleckchem). To test overall editing efficiency of the CRISPR, DNA was isolated from a pool of cells two days after nucleofection.

**Transfection of CRISPR plasmids into HEK-293T cells.**  $2.10^5$  cells HEK-293T cells were nucleofected with 0.4  $\mu$ g plasmid in 20  $\mu$ l SF Primary Cell Nucleofector Solution with program CM-130 (4D Nucleofector, Lonza), according to the manufacturer's instruction. For every transfection, a mock-treated control was included (cells were exposed to the Nucleofector solution and subjected to the Nucleofection procedure, but in the absence of DNA). To test overall editing efficiency of the CRISPR, DNA was isolated from a pool of cells two days after nucleofection.

**Isolation of clonal cell lines by limiting dilution.** Two days after nucleofection, the transfected naïve hESCs were dissociated into single cells using 0.05% Trypsin/EDTA. To prevent clumping of cells, a cell strainer was used. The cells were diluted to a final concentration of 0.5 cells per 100  $\mu$ l of naïve medium, supplemented with 10  $\mu$ M Rock inhibitor (S1049, Selleckchem), to reduce the likelihood of having multiple cells per well. 100  $\mu$ l of diluted cells were added to each well of a 96-well plate, coated with inactivated MEFs. After six to eight days, individual colonies were observed and these were picked and further expanded.

**Genomic DNA extraction.** Genomic DNA was extracted with the QuickExtract DNA isolation kit (QE09050, Epicentre) according to the manufacturer's instructions. Depending on the number of cells, 50–100  $\mu$ l of the solution was used. Genomic DNA was diluted to 20–25 ng/ $\mu$ l and 2  $\mu$ l was used as input for PCR amplification of the desired target site.

**Next-generation-sequencing (NGS) and data analysis.** Extracted DNA was singleplex PCR-amplified with the Kapa2G Robust HotStart Ready Mix (KK5702, Kapa Biosystems). PCR products were subsequently sequenced on an Illumina MiSeq platform according to De Leener *et al.*<sup>50</sup>. For amplicons from DNA extracted from pools of cells, BATCH-GE generated a detailed overview on genome editing efficiency and the detected indel variants after NGS<sup>51</sup>. The region of interest for Cas9 nuclease and nickase was respectively defined as [theoretical cleavage site  $\pm$  30 bp] and [theoretical cleavage site sgRNA1–30 bp; theoretical cleavage site sgRNA2 + 30 bp]. See Supplemental Table 4 for primers.

**RNA isolation, cDNA synthesis and quantitative PCR.** Total RNA was extracted from cells using TRIzol reagent (15596026, Invitrogen) and the miRNeasy Mini RNA extraction kit (217004, Qiagen) according to the manufacturer's instructions. Quality assessment of RNA was performed with the High Sensitivity RNA Analysis Kit on a Fragment Analyser (DNF-472-0500, Advanced Analytical Technologies). Complementary DNA synthesis was performed with the iScript cDNA synthesis kit (170-8891, Bio-Rad Laboratories) in accordance with the manufacturer's instructions. Quantitative reverse transcription-polymerase chain reaction was performed using 2x SYBR Green SsoAdvanced Supermix (172-5274, BioRad Laboratories), 5  $\mu$ M forward and reverse primers, and 5 ng complementary DNA as input. For all experiments, expression was normalised using three reference genes (*ACTB*, *TBP* and *YWHAZ*) and evaluated with qBasePlus software (Biogazelle). See Supplemental Table 5 for primers used for qPCR.

**Analysis expression data and statistical testing.** Statistical differences between groups were evaluated by Wilcoxon rank sum test (alpha level = 0.05). For the neural differentiation experiment, in every group (wildtype, homozygous deletion, heterozygous deletion) 10 independent biological replicates were included. We standardized our expression data from the neural differentiation experiment as described by Willems *et al.*<sup>34</sup> (performing series of sequential corrections including log transformation, mean centering and autoscaling) to reduce the variability between biological replicates. Benjamini-Hochberg (BH) correction was used to account for multiple comparison testing (false discovery rate = 0.05).

**Neural differentiation of naïve human embryonic stem cells.** Naïve human embryonic stem cells were differentiated to neural progenitor stem cells with STEMdiff Neural Induction Medium (05835, STEMCell Technologies) according to manufacturer's instructions with some adaptations. Naïve hESCs, cultured on MEFs, were dissociated to single cells with trypsin/EDTA. To reduce the concentration of MEFs in the cell suspension, the cell suspension was seeded on 0.1% gelatin coated flasks (G-1890, Sigma-Aldrich) and incubated for 1 hour at 37 °C. Since the MEFs will attach to the gelatin first, the remaining suspension of unattached cells was collected.  $7.10^4$ – $8.10^4$  cells were resuspended in STEMdiff Neural Induction Medium + 10  $\mu$ M Rock-Inhibitor (S1049, Selleckchem) and plated on Poly-L-Ornithine/Laminin coated 12-well plates. Medium was daily changed until the desired time point. 10  $\mu$ M Rock-Inhibitor was added at day 0 and day 1.

**Data availability.** All data generated or analyzed during this study are included in this published article (and its Supplementary Information files). The raw datasets generated and/or analyzed are available from the corresponding author on reasonable request.

## References

1. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
2. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
3. Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199, <https://doi.org/10.1038/nature05972> (2007).
4. Brons, I. G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195, <https://doi.org/10.1038/nature05950> (2007).
5. Nichols, J. & Smith, A. The origin and identity of embryonic stem cells. *Development* **138**, 3–8, <https://doi.org/10.1242/dev.050831> (2011).

6. Duggal, G. *et al.* Alternative Routes to Induce Naive Pluripotency in Human Embryonic Stem Cells. *Stem Cells* **33**, 2686–2698, <https://doi.org/10.1002/stem.2071> (2015).
7. Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286, <https://doi.org/10.1038/nature12745> (2013).
8. Hanna, J. *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad Sci USA* **107**, 9222–9227, <https://doi.org/10.1073/pnas.1004584107> (2010).
9. Chan, Y. S. *et al.* Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663–675, <https://doi.org/10.1016/j.stem.2013.11.015> (2013).
10. Ware, C. B. *et al.* Derivation of naive human embryonic stem cells. *Proc Natl Acad Sci USA* **111**, 4484–4489, <https://doi.org/10.1073/pnas.1319738111> (2014).
11. Theunissen, T. W. *et al.* Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487, <https://doi.org/10.1016/j.stem.2014.07.002> (2014).
12. Gu, Q. *et al.* Rapid conversion of human ESCs into mouse ESC-like pluripotent state by optimizing culture conditions. *Protein Cell* **3**, 71–79, <https://doi.org/10.1007/s13238-012-2007-8> (2012).
13. Avior, Y., Sagi, I. & Benvenisty, N. Pluripotent stem cells in disease modelling and drug discovery. *Nat Rev Mol Cell Biol* **17**, 170–182, <https://doi.org/10.1038/nrm.2015.27> (2016).
14. Chen, Y. W. *et al.* A three-dimensional model of human lung development and disease from pluripotent stem cells. *Nat Cell Biol* **19**, 542–549, <https://doi.org/10.1038/ncb3510> (2017).
15. Fonoudi, H. & Bosman, A. Turning Potential Into Action: Using Pluripotent Stem Cells to Understand Heart Development and Function in Health and Disease. *Stem Cells Transl Med* **6**, 1452–1457, <https://doi.org/10.1002/sctm.16-0476> (2017).
16. Choudhary, P. *et al.* Directing Differentiation of Pluripotent Stem Cells Toward Retinal Pigment Epithelium Lineage. *Stem Cells Transl Med* **6**, 490–501, <https://doi.org/10.5966/sctm.2016-0088> (2017).
17. Saurat, N. G., Livesey, F. J. & Moore, S. Cortical Differentiation of Human Pluripotent Cells for *In Vitro* Modeling of Alzheimer's Disease. *Methods Mol Biol* **1303**, 267–278, [https://doi.org/10.1007/978-1-4939-2627-5\\_16](https://doi.org/10.1007/978-1-4939-2627-5_16) (2016).
18. Wang, P. *et al.* CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in cerebral organoids derived from iPS cells. *Mol Autism* **8**, 11, <https://doi.org/10.1186/s13229-017-0124-1> (2017).
19. Kotini, A. G. *et al.* Functional analysis of a chromosomal deletion associated with myelodysplastic syndromes using isogenic human induced pluripotent stem cells. *Nat Biotechnol* **33**, 646–655, <https://doi.org/10.1038/nbt.3178> (2015).
20. Liao, J. *et al.* Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet* **47**, 469–478, <https://doi.org/10.1038/ng.3258> (2015).
21. Brouns, S. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964, <https://doi.org/10.1126/science.1159689> (2008).
22. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821, <https://doi.org/10.1126/science.1225829> (2012).
23. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096, <https://doi.org/10.1126/science.1258096> (2014).
24. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389, <https://doi.org/10.1016/j.cell.2013.08.021> (2013).
25. Buecker, C. *et al.* A murine ESC-like state facilitates transgenesis and homologous recombination in human pluripotent stem cells. *Cell Stem Cell* **6**, 535–546, <https://doi.org/10.1016/j.stem.2010.05.003> (2010).
26. Dodsworth, B. T., Flynn, R. & Cowley, S. A. The Current State of Naive Human Pluripotency. *Stem Cells* **33**, 3181–3186, <https://doi.org/10.1002/stem.2085> (2015).
27. Yang, Y. *et al.* Naive Induced Pluripotent Stem Cells Generated From beta-Thalassemia Fibroblasts Allow Efficient Gene Correction With CRISPR/Cas9. *Stem Cells Transl Med* **5**, 267, <https://doi.org/10.5966/sctm.2015-0157erratum> (2016).
28. Lin, N. *et al.* An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol Cell* **53**, 1005–1019, <https://doi.org/10.1016/j.molcel.2014.01.021> (2014).
29. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826, <https://doi.org/10.1126/science.1232033> (2013).
30. Warriar, S. *et al.* Direct comparison of distinct naive pluripotent states in human embryonic stem cells. *Nat Commun* **8**, 15055, <https://doi.org/10.1038/ncomms15055> (2017).
31. Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492, <https://doi.org/10.1016/j.stem.2009.05.015> (2009).
32. Nair, G., Abranches, E., Guedes, A. M., Henrique, D. & Raj, A. Heterogeneous lineage marker expression in naive embryonic stem cells is mostly due to spontaneous differentiation. *Sci Rep* **5**, 13339, <https://doi.org/10.1038/srep13339> (2015).
33. Bernhart, S. H., Hofacker, I. L. & Stadler, P. F. Local RNA base pairing probabilities in large sequences. *Bioinformatics* **22**, 614–615, <https://doi.org/10.1093/bioinformatics/btk014> (2006).
34. Willems, E., Leyns, L. & Vandesompele, J. Standardization of real-time PCR gene expression data from independent biological replicates. *Anal Biochem* **379**, 127–129, <https://doi.org/10.1016/j.ab.2008.04.036> (2008).
35. Byrne, S. M., Mali, P. & Church, G. M. Genome editing in human stem cells. *Methods Enzymol* **546**, 119–138, <https://doi.org/10.1016/B978-0-12-801185-0.00006-4> (2014).
36. Miyaoka, Y. *et al.* Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci Rep* **6**, 23549, <https://doi.org/10.1038/srep23549> (2016).
37. Horii, T., Tamura, D., Morita, S., Kimura, M. & Hatada, I. Generation of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system. *Int J Mol Sci* **14**, 19774–19781, <https://doi.org/10.3390/ijms141019774> (2013).
38. Horii, T. *et al.* Genome engineering of mammalian haploid embryonic stem cells using the Cas9/RNA system. *PeerJ* **1**, e230, <https://doi.org/10.7717/peerj.230> (2013).
39. Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393–394, <https://doi.org/10.1016/j.stem.2013.03.006> (2013).
40. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J. S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* **24**, 1012–1019, <https://doi.org/10.1101/gr.171322.113> (2014).
41. Liang, X. *et al.* Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* **208**, 44–53, <https://doi.org/10.1016/j.jbiotec.2015.04.024> (2015).
42. Kim, S. J. *et al.* A homozygous Keap1-knockout human embryonic stem cell line generated using CRISPR/Cas9 mediates gene targeting. *Stem Cell Res* **19**, 52–54, <https://doi.org/10.1016/j.scr.2016.12.028> (2017).
43. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187–197, <https://doi.org/10.1038/nbt.3117> (2015).
44. Kim, D. *et al.* Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* **12**, 237, <https://doi.org/10.1038/Nmeth.3284> (2015).
45. Cameron, P. *et al.* Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat Methods* **14**, 600, <https://doi.org/10.1038/Nmeth.4284> (2017).

46. Tsai, S. Q. *et al.* CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR Cas9 nuclease off-targets. *Nat Methods* **14**, 607–+, <https://doi.org/10.1038/Nmeth.4278> (2017).
47. Goyal, A. *et al.* Challenges of CRISPR/Cas9 applications for long non-coding RNA genes. *Nucleic Acids Res* **45**, e12, <https://doi.org/10.1093/nar/gkw883> (2017).
48. Li, R., Zhu, H. & Luo, Y. Understanding the Functions of Long Non-Coding RNAs through Their Higher-Order Structures. *Int J Mol Sci* **17**, <https://doi.org/10.3390/ijms17050702> (2016).
49. Lee, J. H. *et al.* Lineage-Specific Differentiation Is Influenced by State of Human Pluripotency. *Cell Rep* **19**, 20–35, <https://doi.org/10.1016/j.celrep.2017.03.036> (2017).
50. De Leeneer, K. *et al.* Flexible, scalable, and efficient targeted resequencing on a benchtop sequencer for variant detection in clinical practice. *Hum Mutat* **36**, 379–387, <https://doi.org/10.1002/humu.22739> (2015).
51. Boel, A. *et al.* BATCH-GE: Batch analysis of Next-Generation Sequencing data for genome editing assessment. *Sci Rep* **6**, 30330, <https://doi.org/10.1038/srep30330> (2016).

## Acknowledgements

E.Z.J. and E.D. are supported by respectively doctoral grants 11B0117N and 11A9817N, from the Fund for Scientific Research Flanders (FWO). S.V. is supported by postdoctoral grant 12Q7817N from the FWO. We thank Ferring Pharmaceuticals (Aalst, Belgium) for an unrestricted educational grant. B.H. and this research are supported by a Ghent University grant, Faculty Health and Medical Sciences; a research grant from FWO (FWO/KAN/1507816N) and the Agency for Innovation by Science and Technology (IWT, Project Number 150042). We thank Mina Popovic and Jasin Taelman for the fruitful discussions regarding the primed and naïve pluripotency markers.

## Author Contributions

E.Z.J., B.M. and S.V. conceptualized and designed the study. E.Z.J., E.D., E.V.L. and L.V. collected and assembled data. S.W., M.V.d.J. and B.H. provided the naïve stem cell cultures. E.Z.J. and S.V. analyzed and interpreted the results. P.-J.V. performed RNAplfold analysis. E.Z.J. and S.V. wrote the manuscript. B.M. and S.V. supervised the study. All authors critically reviewed the manuscript and approved the final version.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-16932-y>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017