Precise *in vivo* functional analysis of DNA variants with base editing using ACEofBASEs target prediction

Authors

Alex Cornean^{1,2*}, Jakob Gierden^{1,3,4*}, Bettina Welz^{1,2,4*}, Juan L. Mateo⁵, Thomas Thumberger¹, Joachim Wittbrodt^{1,4#}

Affiliations

¹Centre for Organismal Studies, Heidelberg University, Heidelberg, Germany

²Heidelberg Biosciences International Graduate School (HBIGS), Heidelberg, Germany

³Department of Pediatric Cardiology, University Hospital Heidelberg, Heidelberg, Germany

⁴DZHK (German Centre for Cardiovascular Research), partner site Heidelberg/Mannheim

⁵Department of Computer Science, University of Oviedo, Oviedo 33005, Spain

*Equal contribution

#Corresponding Author

Contact: jochen.wittbrodt@cos.uni-heidelberg.de

Abstract

Single nucleotide variants (SNVs) are prevalent genetic factors shaping individual trait profiles and disease susceptibility. The recent development and optimizations of base editors, rubber and pencil genome editing tools now promise to enable direct functional assessment of SNVs in model organisms. However, the lack of bioinformatic tools aiding target prediction limits the application of base editing *in vivo*. Here, we provide a framework for adenine and cytosine base editing in medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*), ideal for scalable validation studies. We developed an online base editing tool ACEofBASEs (a careful evaluation of base-edits), to facilitate decision-making by streamlining sgRNA design and performing off-target evaluation. We used state-of-the-art adenine (ABE) and cytosine base editors (CBE) in medaka and zebrafish to edit eye pigmentation genes and transgenic GFP function with high efficiencies. Base editing in the genes encoding troponin T and the potassium channel ERG faithfully recreated known cardiac phenotypes. We finally validated missense mutations in novel candidate genes of congenital heart disease (CHD) *dapk3*, *ube2b*, *usp44*, and *ptpn11* and present the resulting heart specific phenotypes. This base editing framework applies to a wide range of SNV-susceptible traits accessible in fish, facilitating straight-forward candidate validation and prioritization for detailed mechanistic downstream studies.

Keywords: base editing, SNV, medaka, zebrafish, disease variant validation

Introduction

Single nucleotide variants (SNVs) are the most prevalent disease-causing alterations in human genes (Claussnitzer et al., 2020). While genomic studies continue to unravel SNVs rapidly, functional validation to rank their phenotypic impact *in vivo* remains the central bottleneck, hindering causal assessment. Precise genome editing at the single nucleotide level is a prerequisite to determine the functional relevance and pathogenic contribution of SNVs in cell-based or animal disease models. Although CRISPR-Cas9 gene deletion is powerful in generating valuable insight into null phenotypes, a means to go beyond gene-level analysis to address the functional impact of single nucleotide changes directly is highly desirable.

Base editing has recently emerged as a single-nucleotide-level rubber and pencil tool (Ravindran, 2019), utilizing targeted hydrolytic deamination while largely omitting double-strand breaks (DSBs) (Gaudelli et al., 2017; Komor et al., 2016). Engineered proteins combining the PAM-specific action of a Cas9-D10A nickase or a dead Cas9 nuclease with cytidine or deoxyadenosine deaminases constitute the base editor complex. The deaminases thereby enable cytosine to thymine (C-to-T) or adenine to guanine (A-to-G) conversion within a predefined window of activity on the DNA segment, with limited indel formation (Figure 1-supplement 1). Established in human tissue culture, cytosine base editors (CBEs) and adenine base editors (ABEs) combined enable all transition mutations (Gaudelli et al., 2017; Komor et al., 2016). External development, rapid growth, and transparency of medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*) embryos make these established vertebrate human disease models ideal for studying genetic variants' consequences *in vivo* (Gut et al., 2017; Hammouda et al., 2021; Meyer et al., 2020). Both ABEs and CBEs have been employed in zebrafish (Carrington et al., 2020; Qin et al., 2018; Rosello et al., 2021b; Zhang et al., 2017; Zhao et al., 2020), reaching up to 40% and 91% editing efficiencies, respectively. Notably, Rosello et al. (2021b) have recently established a proof-of-concept for cytosine base editing to study developmental and human disease variants in fish. CRISPR/Cas9-mediated gene knockout can be readily used to discern gene function in fish in F0 enabled by high Cas9 nuclease efficiencies and robust workflows that include versatile sgRNA design tools (Hoshijima et al., 2019; Kroll et al., 2021; Stemmer et al., 2015). Various iterations of improvements yielding highly efficient base

editing tools *in vitro*, also led to high C-to-T editing efficiencies in some of the loci tested (Rosello et al., 2021b). However, for both CBE and ABE systems tested in fish, reliable information for base editor selection to consistently achieve high efficiencies is absent, impeding the validation of specific DNA variants by linking these to specific traits in F0. Moreover, the lack of user-friendly web-based software for sgRNA design and selection for target DNA variants has hampered routine applications of base editing in fish.

Here, we present a comprehensive workflow for adenine and cytosine base editing in medaka and zebrafish. To design and evaluate nucleotide and codon changes and predict potential off-target sites, we built on our CRISPR-design tool CCTop (Stemmer et al., 2015), now including a careful evaluation of base-edits (ACEofBASEs). We employ this simple, efficient, and easy-to-use software to design sgRNAs for the state-of-the-art cytosine base editors BE4-Gam (Komor et al., 2017), ancBE4max (Koblan et al., 2018), and evoBE4max (evoAPOBEC1-BE4max; Thuronyi et al., 2019), as well as the adenine base editor ABE8e (Richter et al., 2020). Editing efficiencies close to homozygosity in injected (F0) embryos allowed us to directly link specific missense mutations or premature termination codons (PTC) to phenotypic outcomes, with a recapitulation of expected phenotypes. Notably, introducing a single missense mutation in an ultra-conserved pore-domain of *kcnh6a* (*O1* ERG), an essential potassium channel gene, resulted in a non-contractile ventricle with striking secondary morphological defects. We further used base editing in medaka to address SNVs associated with human congenital heart disease (CHD). Efficient introduction of conserved missense mutations in four novel candidate genes, *dapk3*, *ube2b*, *usp44*, and *ptpn11*, enabled validating phenotypic consequences and uncovering a functional role of these genes in cardiovascular development.

The conservation of a significant proportion of coding SNVs in fish opens the door to address a plethora of disease variants robustly in F0. The presented software and workflow for base editing in medaka and zebrafish coupled to our experimental data allow rational editor choice for *in vivo* SNV validation experiments. The efficiencies demonstrated are compatible with F0 phenotype quantification, a reliable and robust shortcut to endpoints of detailed developmental studies, and rapid candidate validation in human genetics.

Results

ACEofBASEs for base editor sgRNA design

The lack of comprehensive software tools for base editor sgRNA design and on- and off-target editing predictions significantly limits widespread base editing applications. We developed the online software tool ACEofBASEs as an extension of our CRISPR-Cas prediction tool CCTop (Stemmer et al., 2015) for straightforward sgRNA design for CBEs and ABEs. ACEofBASEs identifies all sgRNA target sites of a given query nucleotide sequence with adenine or cytosine residues present in the respective base-editing windows. A dropdown menu enables the selection of one of four recent state-of-the-art base editors: BE4-Gam, ancBE4max, evoBE4max, and ABE8e (Figure 1a). The standard base editing window parameters are set according to *in vitro* data, which differentiates between a window with high (blue) and observed activity (light blue) (Arbab et al., 2020; Huang et al., 2021; Koblan et al., 2018; Komor et al., 2017; Richter et al., 2020; Thuronyi et al., 2019, Figure 1-supplement 2). All detected sgRNAs for a particular target site are ranked considering the dinucleotide context and the editing window (displayed on a yellow-orange-brown scale; Figure 1-supplement 2). Alternatively, the limits of the base editing window can be freely adjusted (without sgRNA ranking, Figure 1a). The specifications of the oligonucleotides for sgRNA cloning are customizable in the user interface. Off-target prediction in the selected target genome follows the same routine used in CCTop (Stemmer et al., 2015) but only presents sgRNA off-target sites with adenines or cytosines present in the respective base editing window.

ACEofBASEs expects the query sequence to be coding and assumes that the ORF starts at the first nucleotide provided. The reading frame can be adjusted in the dropdown menu and the codons, translation and predicted sequence alterations are highlighted on the results page. The sgRNA display includes base edited amino acid sequence changes specified as missense (purple) or nonsense (black) mutations for the selected ORF (Figure 1b). Avoiding off-target editing is essential for organismal F0 screens as well as in cell-based assays, and consequently, the precise off-target prediction represents a crucial feature of any base editing software tool. The ACEofBASEs overview of sgRNA targets provides this information at a glance and details each sgRNA with potential off-target coordinates and the type of genomic location (intergenic, intronic, or exonic) (Figure 1-supplement 2).

Biallelic PTC and missense mutations through cytosine and adenine base editing in fish

Single nucleotide changes through PTCs resulting in nonsense-mediated decay or truncated proteins, and missense mutations substituting individual amino acids, can result in a considerable number of functionally relevant phenotypes. We investigated editing efficiencies for cytosine editors by their potential of introducing PTCs, and in contrast, for adenine base editors through the installation of missense mutations resulting in loss-of-function phenotypes. We compared the original BE4-Gam (Komor et al., 2017) to two next-generation CBEs, ancBE4max (Koblan et al., 2018) and evoBE4max (Thuronyi et al., 2019), the latter of which had not been previously tested in fish (Figure 2a). In addition, we also employed the highly processive adenine base editor ABE8e (Richter et al., 2020), potentially overcoming target constraints reported for ABE7.10 in zebrafish (Qin et al., 2018).

An acute and early onset of base editing activity and high base conversion rates are prerequisite for efficient biallelic editing causing reliable phenotypes in F0 experiments. We examined the *in vivo* efficiencies of state-of-the-art base editors by editing the *oculocutaneous albinism II* gene (*oca2*) required for eye pigmentation, a gene that had been successfully targeted with BE4-Gam in medaka (Thumberger et al., 2021)). The loss of retinal pigmentation provides a scorable readout to address the efficacy of biallelic base editing or Cas9 based knock-out experiments (Lischik et al., 2019; Zhang et al., 2017), since a single allele of *oca2* is sufficient to generate normal retinal pigmentation and only cells affected in both alleles will exhibit fail to get pigmented. Within conserved regions of *oca2* (Figure 2-supplement 1a), ACEofBASEs proposed a sgRNA to introduce a PTC (*p.[T332I; Q333X]*) or two missense mutations (*p.[T332A; Q333R]*) in medaka with CBEs or ABEs, respectively.

We used microinjection of mRNA to deliver base editors and sgRNAs into 1-cell stage medaka embryos. Phenotyping of base edited embryos, henceforth referred to as “editants”, at 4 dpf revealed a spectrum from wild-type to complete loss of ocular pigmentation divided into five arbitrary categories (Figure 2b-c). While BE4-Gam produced at least one pigment-free clone in almost all editants, overall, pigmentation was still prevalent. In sharp contrast, both ancBE4max and evoBE4max in a large

fraction of editants led to the almost complete eye pigmentation loss (Figure 2d). Sanger-sequencing of randomly pooled editants reflected the same trend, i.e., substantially remaining wild-type C peaks (protospacer position 5-7) for BE4-Gam, but complete C-to-T conversion achieved at protospacer position 6-7 using ancBE4max and evoBE4max (Figure 2e). Importantly, no other DNA sequence changes such as indels or unwanted editing were observed around the locus (Figure 2-supplement 1b). Quantification of Sanger sequencing results by editR (Kluesner et al., 2018) revealed editing efficiencies at C7 (c.997C>T, i.e., p.Q333X) of 29.3±7.4% (n=3), 93.8±7.9% (n=5), and 93.3±9.8% (n=3) for BE4-Gam, ancBE4max and evoBE4max, respectively (Figure 2f, Table 1). Notably, due to inherent limitations of Sanger sequencing, both editing <5% and monoallelic sequences are not reliably resolved (Kluesner et al., 2018). Interestingly, evoBE4max showed reduced C-to-T conversion at C5 within an AC dinucleotide context (Figure 2e-f), in line with its context-dependent substrate C editing efficiency detailed in-depth in cell culture (Arbab et al., 2020). Accordingly, we observed a reduced average conversion rate of 20.7±0.6% at target C5 in an AC context for evoBE4max, compared to 29.3±6.7% and 86.4±11.5%, for BE4-Gam and ancBE4max, respectively (Figure 2f). In addition to ancBE4max and evoBE4max displaying prominently increased editing activity, a slight increase in aberrant F0 phenotypes was observed with both editors (Figure 2-supplement 1c).

In summary, ancBE4max and evoBE4max are highly efficient in converting targeted cytosines *in vivo*, demonstrated by introducing a PTC in *oca2*, resulting in near-complete loss of eye pigmentation in medaka embryos compatible with accelerated F0 genotype-phenotype studies.

We tested the effects of single (*oca2-Q333*) and pooled sgRNA (*oca2-Q256*, *T306*, *Q333*) editing events, aided by ACEofBASEs, with ABE8e to compare the impact of multiple vs a single amino acid exchange. Following our described workflow and categorization with ABE8e, we noted that pooled *oca2*-sgRNA injections led to a higher fraction (32 of 32) of embryos with reduced eye pigmentation, in contrast to the *oca2*-Q333 injection alone (14 of 48) (Figure 2d). The surge in expected phenotypes, however, coincided with an increase in aberrant phenotypes (Figure 2-supplement 1b). We observed close to homozygous editing resulting in missense mutations p.T332A (99.3±0.6%, 93.7±11.0%) and p.Q333R (89.0±6.6%, 90.3±2.1%) by single and pooled

oca2 experiments, respectively (Figure 2e-g, Table 2). For pooled sgRNA injections, we also observed efficient editing at *oca2*-Q256 (100%) and -T306 (52.3±2.3%) (Figure 2-supplement 2a-b), which may implicate a causative role of *oca2*-Q256 in the increased phenotypic fraction. Closer analysis of *oca2*-Q333 editants with evoBE4max and ABE8e revealed no off-target (OT) editing at three partially matching sites indicated by ACEofBASEs (Figure1-supplement 3, Figure 2-supplement 1d).

Next, we compared the performance of ABE8e in injected zebrafish embryos and determined the state of eye pigmentation of successfully gastrulating embryos at 56 hpf. Here, seven out of 54 editants (13%) displayed reduced eye pigmentation (Figure 2h-i'). Genotyping of individual embryos, grouped according to phenotype (normal vs. pigment-loss), revealed a direct correlation between phenotype and A-to-G transition efficiency with 29.0±29.8% and 61.8±29.6% at A5 (Figure 2j), with no aberrant sequence changes in the locus (Figure 2-supplement 2c). Moreover, we observed editing of up to 95% introducing a close to homozygous missense p.L293P mutation, indicating that, similarly to medaka *oca2*-Q333 editing, these missense mutations do not fully impair protein function.

Taken together, we showed that ABE8e efficiently introduced missense mutations with all four sgRNAs tested in medaka and zebrafish. To examine whether missense mutations can lead to fully penetrant phenotypes in F0, we next addressed the consequences of a specific missense loss-of-function mutation on transgenic GFP.

Figure 2

Figure 2. Somatic cytosine and adenine base editing at the *oculocutaneous albinism II* (*oca2*) locus in medaka and zebrafish allows direct functional assessment (continued on next page).

Figure 2. (a) Schematic diagram of the cytosine base editors BE4-Gam, ancBE4max and evoBE4max (evoAPOBEC1-BE4max) and the adenine base editor ABE8e. Cas9n-D10A nickase (light grey) with N-terminally linked cytidine or deoxyadenosine deaminase (pink) and C-terminal SV40 or bipartite (bp) nuclear localisation sequence (NLS, red). All except BE4-Gam also contain the bpNLS N-terminally. CBEs contain variations of the rat APOBEC-1 cytidine deaminase, whereas ABE8e contains the TadA* domain (tRNA adenine deaminase), CBEs further contain C-terminally linked Uracil glycosylase inhibitors (UGI, green). Gam protein from bacteriophage *Mu* (purple) and linkers of varying lengths (dark grey). (b) Scheme of the experimental workflow. Cytosine or adenine base editor (CBE/ABE) mRNA and *oca2-Q333* or a pool of three *oca2-sgRNAs* (–Q256, –T306, –Q333) were injected into the cell of a medaka zygote. Control injections only contained *oca2-Q333* or *ABE8e* mRNA. (c) Phenotypic inspection of eye pigmentation was performed at 4 dpf (dorsal view). (d) Grouped and quantified pigmentation phenotypes shown for BE4-Gam, ancBE4max, evoBE4max, and ABE8e experiments. Control only contains *oca2-Q333* sgRNA. n shown excludes embryos that are otherwise abnormal or dead, with abnormality rate given in supplement 1b. (e) Exemplary Sanger sequencing reads for each experimental condition, obtained from a pool of five randomly selected embryos at the *oca2-Q333* locus. (f–g) Quantification of Sanger sequencing reads (by EditR, Kluesner et al., 2018) for BE4-Gam (n=3), ancBE4max (n=5) and evoBE4max (n=3) (f), and ABE8e for single (n=3) and pooled *oca2-sgRNA* experiments (n=3) (g). Pools of five embryos per data point summarizes editing efficiencies. Mean data points are summarized in Tables 1 and 2. To highlight the dinucleotide context, the nucleotide preceding the target C or A is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective C or A. (h) Microinjections into the yolk of 1-cell stage zebrafish were performed with *ABE8e* mRNA and *oca2-L293* sgRNA. Zebrafish larvae were phenotypically analysed at 56 hpf and individual larvae were subsequently genotyped. (i–i') Larvae were scored as without ("normal", black) or with loss of eye pigment (grey). (j) Sanger sequencing on individually scored larvae was analysed by EditR and plotted according to phenotype. Scale bars = 400 μ m (c) or 100 μ m (i). dpf / hpf=days/hours post fertilization.

The highly processive ABE8e efficiently introduces missense mutations inactivating GFP fluorescence

To expand our initial analysis on ABE8e base editing efficiency, we took advantage of the deep functional annotation of individual amino acids in GFP (Fu et al., 2015; Li et al., 1997; Patterson et al., 1997) as target protein.

We assayed the loss of GFP fluorescence in a medaka EGFP and mCherry double transgenic line with heart-specific, nuclear fluorophore expression (*myl7::EGFP*, *myl7::H2A-mCherry*) (Hammouda et al., 2021). Co-injection of ABE8e mRNA with the *GFP-C71* sgRNA into 1-cell stage embryos resulted in the complete loss of GFP fluorescence in all hearts assayed at 4 dpf (n=41), indicating quantitative editing as also reflected by the absence of mosaicism.

Cells in the medaka 4-cell embryo form a syncytium allowing limited diffusion of the base editing machinery provided to one of the blastomeres. Only when forced by injection into a single blastomere at the 4-cell stage, mosaicism, apparent by speckled

GFP-positive hearts, could be forced in all embryos analyzed (100%, n=23) (Figure 3a-b), further underpinning the high editing efficiency.

Sanger sequencing and EditR quantification confirmed the strikingly high efficiencies in introducing the desired p.C71R missense mutation ($97.0 \pm 4.4\%$ and $90.0 \pm 7.8\%$), as well as a notable bystander edit causing p.V69A missense mutation (61.2 ± 9.4 and $12.6 \pm 8.1\%$) for 1-cell and 4-cell stage injections, respectively (Figure 3f-g; Table 2).

The highly efficient editing, even when only provided to one of the four cells, underscored the high *in vivo* potency of ABE8e even at reduced levels. The high, almost quantitative activity of both CBEs and ABE8e, even at low concentrations *in vivo*, led us to address phenotypic consequences of editing individual sites in established and so far, unknown targets.

Figure 3

Figure 3. ABE8e efficiently introduces missense mutation and completely abolishes GFP fluorescence in F0. (a) Co-injection of ABE8e mRNA with GFP-C71 sgRNA into a single cell of 1-cell or 4-cell stage medaka embryos (*myl7::EGFP*, *myl7::H2A-mCherry*). Control injections with ABE8e mRNA. Scoring of fluorescence was performed at 7 dpf followed by genotyping of each individual embryo (continued on next page).

Figure 3. (b-d') Confocal microscopy of chemically arrested hearts (representative images) at 7 dpf (lateral view with V=ventricle, A=atrium). Overview images, overlaid with transmitted light, show maximum z-projections of optical slices acquired with a z-step size of 5 μ m. Scale bar = 200 μ m (**b-d**). Close-up images show maximum z-projections of optical slices acquired with a z-step size of 1 μ m. Note the display of A-to-G conversion rates for A4 causing the p.C71R missense mutation (see g) and Table 2). Scale bar = 50 μ m (**b-d'**). (**e-f**) Quantification of Sanger sequencing reads shows close to homozygosity rates of A-to-G transversions installing the C71R missense mutation (Table 2). Note: sgRNA GFP-C71 targets the complementary strand (arrow in f). To highlight the dinucleotide context, the nucleotide preceding the target A is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective A. dpf=days post fertilization.

Functional intervention in F0 by CBE induced stop-gain mutations in cardiovascular troponin T

The high efficiency of base editing already in the injected generation opens the door for functional evaluation of candidates rapidly uncovered by extensive population-scale sequencing studies of cardiovascular disease (CVD) (Ingles et al., 2020; Kathiresan and Srivastava, 2012; Pierpont et al., 2018). The understanding of the mechanistic contribution of those variants holds a great promise for prevention and precision medicine.

We first validated the relevance of *in vivo* base editing in the injected generation by phenocopying reference heart phenotypes associated with fully penetrant, recessive lethal genes from fish (Chen et al., 1996; Meyer et al., 2020; Stainier et al., 1996). We targeted *troponin T type 2a* (*tnnt2a*) that, when mutated, results in silent, non-contractile hearts in zebrafish *sih* mutants (Sehnert et al., 2002) as well as in medaka "crispants" of *tnnt2a* (Meyer et al., 2020). We addressed the ACEofBASEs predicted conversion of Q114 into a PTC mediated by the previously employed sgRNA (Figure 4a) combined with a cytosine base editor by co-injection of this sgRNA with different CBEs. Those resulted in the impairment of cardiac contractility at various degrees (Figure 4b, Video 1).

Notably, the penetrance of the phenotype drastically increased from BE4-Gam to ancBE4max and evoBE4max (Figure 4c), reflecting the editing efficiency (c.340C>T resulting in p.Q114X of 0% (BE4-Gam), 27.8±16.8% (ancBE4max), and 85.9±23.5% (evoBE4max) (Table 1)). Analysis of editing efficiencies in the silent heart group indicated prevalent homozygous C8-to-T edits (p.Q114X) by evoBE4max (Figure 4d-d'). Highly efficient base editing of *tnnt2a* in injected F0 embryos therefore allowed us to faithfully reproduce reference mutant phenotypes.

Expanding our proof of concept analyses, we determined editing efficiencies using a sgRNA for *tnnt2c* (paralog of silent heart-associated *tnnt2a*) and two sgRNAs for *s1pr2*, a sphingolipid receptor known to direct the midline migration of the bilateral heart precursors disrupted in the *miles apart* (*mil*) zebrafish mutant (Kupperman et al., 2000). We observed average C-to-T editing efficiencies of 38-88% across these loci in F0 (Figure 4-supplement 1, Table 1), further substantiating the high editing power in F0, allowing us to address the role of individual amino acids in complex proteins.

Figure 4

Figure 4. Stop-gain mutation in medaka troponin T gene by cytosine base editing accurately recreates mutant phenotypes in F0. (a) Schematic diagram of protocol after co-injection of cytosine editor mRNA (comparing BE4-Gam, ancBE4max, evoBE4max) with sgRNA *tnnt2a*-Q114 (PTC) into 1-cell stage medaka embryos; control injections only contained the sgRNA. (b) Editing of *tnnt2a* resulted in a range of phenotypes classified into five categories, including general morphogenic (global), dysmorphic but still functional heart chambers (cardiac), and non-contractile hearts (silent), where cardiac and silent phenotype groups displayed homogeneously additional developmental retardation. Scale bar = 400 μ m (overview) and 100 μ m (zoom-in). Ventricle (V), atrium (A), diastole (Dia) and systole (Sys) are indicated. (b') Representative scheme of fractional shortening of the heart chambers in specified phenotype groups highlighting significant morphological consequences (small ventricle) in the silent heart group. (c) Fraction of phenotype scores as a function of cytosine base editor injections (continued on next page).

Figure 4. (d) Summary of editor type-specific C-to-T conversion efficiencies relative to the target C protospacer position grouped by phenotype class for BE4-Gam (n=6), ancBE4max (n=6) and evoBE4max (n=12). To highlight the dinucleotide context, the nucleotide preceding the target C is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective C. **(d')** Example Sanger sequencing reads of single edited embryos with resulting missense and stop-gain mutations through editing at C5, C7, and C8, respectively. dpf=days post fertilization.

***In vivo* modeling of heritable LQTS mutations using cytosine and adenine base editing**

We next specifically targeted crucial amino acids in the cardiac-specific potassium channel ERG, essential for cardiac repolarization (Arnaout et al., 2007; Hassel et al., 2008). The ERG channel is a homotetramer with each subunit consisting of six alpha-helical transmembrane domains S1-S4 forming a voltage sensor and the pore-forming domains S5-S6 (Figure 5-supplement 1) (Wang and MacKinnon, 2017; Zhang et al., 2004). Functional impairment of the human potassium channel *Hs* ERG (*KCNH2*) causes inherited or acquired long QT syndrome (LQTS), respectively, associated with sudden cardiac death (Abriel and Zaklyazminskaya, 2013).

Using CBEs, we revealed the loss-of-function phenotype in *kcnh6a* editants by PTC introduction in the F0 and F1 generation (Figure 5-supplement 1-3, Movie 2). To specifically alter the voltage-sensing and gating behavior of ERG channels *in vivo*, we altered critical amino acids in the conserved transmembrane domain by introducing precise point mutations (Figure 5-supplement 1a-b), modeled on alleles described in zebrafish (Arnaout et al., 2007).

We used ACEofBASEs to select sgRNAs that facilitate base edits targeting and putatively inactivating the S4 and S4-S5 linker domain of the medaka ERG channel, resulting in the sgRNAs *kcnh6a-R512*, *kcnh6a-T507*, and *kcnh6a-D521* (Figure 5-supplement 1a-b). We addressed the consequences of editing crucial amino acids in combination and individually.

Co-injection of all three sgRNAs with ABE8e editor mRNA into medaka embryos revealed a high frequency of specific heart phenotypes at 4 dpf (82% and 88% respectively, two independent biological replicates; Figure 5c, Figure 5-supplement 4), including significantly reduced or complete loss of ventricular contraction (Figure 5b-c, Video 3).

The introduction of a missense mutation to exchange a single arginine at *kcnh6a-R512* *in organismo* was particularly interesting, since the edit of a single nucleotide removes a positive charge at a crucial position. Injection of the sgRNA *kcnh6a-R512* together with ABE8e editor mRNA into medaka embryos demonstrated that this S4 sensor charge is functionally essential *in vivo*. Its exchange resulted in partially or entirely silent ventricles in the majority of scored F0 embryos (Figure 5c, Figure 5-supplement 4) as suggested by electrophysiological studies on heterologously expressed *Hs* ERG (Zhang et al., 2004).

Analysis of the genomic *kcnh6a* locus around R512 of pooled (*kcnh6a-T507*, *-R512*, *-D521*) and *kcnh6a-R512* single editants revealed reproducibly high conversion rates of AGG (R512) to GGG (G) ($92.1 \pm 6.6\%$ and $91.7 \pm 9.5\%$), respectively (Figure 5d). In pooled editants the two other sgRNA sites T507 and D521 also reached high editing efficiencies of $81.6 \pm 7.2\%$ (A6-to-G) and $61.0 \pm 12.5\%$ (A5-to-G), respectively (Figure 5d'-d''; Table 2). Notably, ABE8e-mediated adenine base editing achieved homozygosity in three out of four imaged R512G embryos in F0 (Figure 5e).

Investigating the ABE8e-mediated R512G missense editants in a *myl7::GFP* reporter (Gierten et al., 2020) revealed a striking morphological impact of the primary repolarization phenotype in specimens with silent hearts. While atrioventricular differentiation is complete, the ventricular muscle showed major growth and differentiation deficiencies. High-resolution imaging of silent ventricles showed ventricular collapse and multiple vesicle-shaped, aneurysm-like structures (Figure 5f). Whether this structural phenotype directly relates to *kcnh6a* effects or, more likely, is a secondary consequence of the lack of forces generated by chamber contractions and directed blood flow is an exciting starting point for future investigations.

Taken together, applying ABE8e in medaka embryos enabled modifying a conserved voltage sensor domain of the LQTS/SQTS-related potassium channel ERG *in vivo* at the level of a single amino acid in F0.

Figure 5

Figure 5. *In vivo* modeling of human LQTS-associated mutations using adenine base editing of the medaka ERG channel gene *kcnh6a*. (a) Regime of ABE8e mRNA injections with a single (*kcnh6a-R512*) or pooled sgRNAs (*kcnh6a-T507*, -*R512*, -*D521*) targeting different amino acid codons in the voltage sensor S4 domain/S4-S5 linker of the medaka potassium channel ERG in *myl7::EGFP* (*Cab* strain) transgenic embryos; control injection included ABE8e mRNA only. (b) Phenotypes in F0 comprised primary cardiac malformation (dysmorphic ventricle with impaired contractility) and more severe global phenotypes with general retarded development and prominently dysmorphic hearts the proportions of which are given in (c). (d-d'') Genotyping summaries of the three sgRNA loci with phenotype class annotations for each genotyped specimen with a comparison of single sgRNA-R512 injection to a pool with two additional sgRNAs (T507 and D521) targeting the medaka ERG S4 voltage sensor; inlets display Sanger reads with the editing of A8 (d), A6 and A8 (d') and A5 (d'') contained in the core editing windows; sgRNA pool (n=8) and sgRNA-R512 (n=6). To highlight the dinucleotide context, the nucleotide preceding the target A is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective A (continued on next page).

Figure 5. (e-f) Confocal microscopy imaging of the heart in a *myl7::EGFP* reporter line injected with ABE8e mRNA and sgRNA-R512 at 7 dpf revealing significant chamber wall defects of non-contractile/spastic ventricles with A-to-G editing of 100% in 3/4 of the specimen (e). Overview images, overlaid with transmitted light, show maximum z-projections of optical slices acquired with a z-step size of 5 μm ; scale bar = 200 μm . Close-up images show maximum z-projections of optical slices acquired with a z-step size of 1 μm . Note the display of A-to-G conversion rates. Scale bar = 50 μm . e=eye, b=brain, y=yolk, V=ventricle, A=atrium, dpf=days post fertilization.

Functional validation of congenital heart disease-associated missense mutations with SNV resolution using cytosine editing

We have demonstrated the functionality and potential of CBEs and ABEs by introducing PTCs or missense mutations into well-studied genes with reference phenotypes. We next used base editing to validate specific mutations in novel candidate genes associated with congenital heart disease. Extensive sequencing studies have revealed a polygenic origin of CHD, uncovering a plethora of candidate genes with enrichment of SNVs with predicted pathogenic relevance acting as protein-truncating or missense variants (Homsy et al., 2015; Sifrim et al., 2016; Zaidi et al., 2013). To validate novel SNVs using base editing, we chose four highly ranked candidate genes, *DAPK3*, *UBE2B*, *USP44*, and *PTPN11*, each with a specific CHD-associated missense mutation (Homsy et al., 2015; Zaidi et al., 2013).

Using the missense mutation loci, ACEofBASEs identified sgRNAs compatible with cytosine editing to introduce missense mutations in the highly conserved medaka orthologues resulting in identical or alternative amino acid changes (Figure 6a, Figure 6-supplement 1).

Each of these sgRNAs was co-injected with evoBE4max into medaka zygotes using wild-type strains Cab or HdrR with dual-color cytosolic and nuclear heart-specific labels (*myl7::EGFP*, *myl7::H2A-mCherry*). Individual embryos were subjected to Sanger sequencing and EditR analysis, quantifying C-to-T transition. For each of the four genes we obtained the expected edits with 86.9 $\pm$ 8.7% (*dapk3*-p.P204L), 54.2 $\pm$ 24.8% (*ube2b*-p.R8Q), 99.7 $\pm$ 0.9% (*usp44*-p.E68K), and 88.0 $\pm$ 7.3% (*ptpn11*-p.G504E/K) average efficiency respectively (Figure 6, Figure 6-supplement 2, Table 3). In addition, missense bystander edits were observed in *dapk3* (p.L205F, 61.7 $\pm$ 14.1%), *ube2b* (p.R7K, 67.4 $\pm$ 16.8%), and *usp44* (p.M67I, 97.0 $\pm$ 5.0%).

The first time point of phenotypic analysis at 4 dpf revealed cardiovascular phenotypes for all four CBE-driven missense mutations (Figure 6d). While atrium and ventricle were specified, we observed a lack of normal dextra heart looping (R-loop) with a displacement of the ventricle to the left (L-loop, left-right asymmetry defect), vessel malformation and blood clotting (*dapk3*), defective looping, and aberrant heart morphology (*ube2b*, *usp44*), and mesocardia, i.e., lack of looping, and smaller ventricles (*ptpn11*) (Figure 6d, Figure 6-supplement 3). Notably, missense mutations in each of the four candidate genes specifically impacted the cardiovascular system (Figure 6-supplement 3), validating their suspected impact.

Representative embryos were further characterized by confocal microscopy of the heart at 7 dpf. High-resolution imaging revealed abnormal AV channel formation, dysmorphic atrium and ventricle in the *usp44*-p.E68K editant (100% edited allele, Figure 6f). The *ptpn11*-p.G503E editant (88% edited allele) showed a small ventricle and linear atrium (Figure 6g). Additionally, *ube2b* and *dapk3* editants displayed consistent cardiac phenotypes (Figure 6-supplement 4).

In summary, we have established and validated a comprehensive toolbox for the context-dependent editing of single nucleotides in model systems. We transferred human mutations to medaka to translate an association into causality. We applied base editing in medaka and precisely modeled single SNVs *in vivo*, validating single missense mutations initially associated with human congenital heart defects and highlighting the potential of targeted base editing for disease modeling.

Figure 6

Figure 6. Cytosine base editing enables human CVD-associated SNV validation. (a) Candidate human CVD gene SNV validation workflow. (b) To target the SNVs *evoBE4max* mRNA was co-injected into the 1-cell stage of the medaka wild-type or *myl7::EGFP*, *myl7::H2A-mCherry* reporter strain together with the corresponding target or *oca2-Q333* (control) sgRNAs. Individual, imaged, embryos were then further analysed to determine the rate of C-to-T transversions. (c) Cytosine editing efficiencies are substantial for all candidate genes tested. Data shown in Figure 6-supplement 2 was replotted, including all data points from a-d across all target cytosine along the protospacer. Sample numbers: *dapk3-P204* (n=7), *ube2b-R8* (n=5), *usp44-E68* (n=11), and *ptpn11-G504* (n=11). (d) Representative phenotypes of 4 dpf base edited embryos are shown for all four tested candidate CVD genes including *oca2-Q333* controls. Top, ventral view, with V=ventricle, A=atrium. (e-g) Confocal microscopy of selected candidate validations in the reporter background. Hearts were imaged in 7 dpf hatched double fluorescent embryos. Images show maximum projections of the entire detectable cardiac volume with a step size of 1 μ m. Cartoons (left) highlight the looping defects observed in *usp44* and *ptpn11* base edited embryos with ventricle-atrium inversion (f) or tubular heart (g) (continued on next page).

Figure 6. (e'-g') Imaged embryos were subsequently genotyped and quantified C-to-T transversions for the target codon are shown. Note: due to the inverted nature of the confocal microscope used, raw images display a mirroring of observed structures, which we corrected here for simpler appreciation. Scale bar = 100 μ m (**d, e-g**). dpf=days post fertilization.

Discussion

Determining the *in vivo* functional consequences of SNVs associated with physiological trait variation and disease in humans is critical to assess causative genetic variants. We provide a framework for *in vivo* base editing to close the gap between SNV discovery and validation using small fish model systems, medaka and zebrafish, allowing accurate phenotype assessment of gene function interference. Our online base editing tool ACEofBASEs allowed us to probe the efficiency of current CBEs and one ABE at various test loci in fish and to uncover the developmental impact of point mutations in four novel CHD candidate genes *dapk3*, *ube2b*, *ups44* and *ptpn11*.

The synergy between CBEs and ABEs and our open-access web tool ACEofBASEs facilitates immediate access to base editing experiments in cell and organism-based assays. The modular design of our software provides entry- as well as expert-level for base editing applications. In contrast to existing tools like the BE-designer (Hwang et al., 2018) or PnB Designer (Siegnier et al., 2021) or custom scripts to generate nonsense mutations (Rosello et al., 2021b), ACEofBASEs presents the altered amino acid composition, direct off-target assessment with linked sequence information, and details for sgRNA cloning. In addition, comprehensive editing annotations, a selection of base editor PAM variants, and genome selection available for an extensive (and constantly growing) collection of species expand the applicability of base editing to laboratories working on a wide range of different model systems.

ACEofBASEs guided immediate and efficient base editing at multiple relevant loci. We used it to assess the *in vivo* performance of three recent state-of-the-art CBEs, BE4-Gam, ancBE4max, evoBE4max, and ABE8e, in comparison to the respective *in vitro* specifications. Our results show comparable overall editing efficiencies for ancBE4max, evoBE4max, and ABE8e, with the highest level of edits at efficiencies ranging from 94 to 100%, indicating almost quantitative bi-allelic editing in the injected generation (Tables 1-3). We have determined the *in vivo* editing windows and context-

dependent efficiencies of all four editors (Figure 7a-a'; Table 4). Our *in vivo* findings confirm the *in vitro* characteristics of the APOBEC-1 deaminase-inherent canonical dinucleotide context sequence preference (TC>CC>AC>GC), which was ameliorated in ancBE4max and evoBE4max, both with improved editing at AC or GC-dinucleotides, respectively (Arbab et al., 2020; Huang et al., 2021; Komor et al., 2016). Similarly and consistent with the *in vitro* observations (Richter et al., 2020) a dinucleotide preference was not observed for ABE8e (Figure 7b-b').

Our analyses show enhanced efficiencies compared to previous reports for cytosine base editing in fish (Rosello et al., 2021b; Zhang et al., 2017; Zhao et al., 2020). In particular, the adenine base editor variant ABE8e in both medaka and zebrafish remarkably exceeded efficiencies previously reported for ABE7.10 in zebrafish (Qin et al., 2018). Overall, high efficiencies and precision *in vivo* with limited bystander edits and undetectable off-target DNA editing allowed to overcome the bottleneck of current genome engineering strategies for the reliable, somatic interrogation of DNA variants. We demonstrate the power of *in vivo* analysis and validation of human genetic variants in fish by generating PTCs and specific missense mutations in essential cardiac genes resulting in phenocopies of the respective reference heart phenotypes. Exchanging a single positively charged acid to a neutral one in the ultra-conserved voltage sensor domain (S4) of the potassium channel ERG (*kcnh6a*) resulted in the formation of a non-contractile ventricle equal to previously reported phenotypes of null mutants established in zebrafish (Hoshijima et al., 2016). Given the high editing efficiency, the ventricle collapse with secondary morphological aberrations revealed by high-resolution imaging highlights a previously unreported ERG functionality consistently observed in the injected generation. In contrast to the knockin-knockout strategy (Hoshijima et al., 2016), the base editing approach allows to efficiently interrogate distinct human point mutations in potentially disease-relevant genes as well as addressing structure function relationships by editing individual amino acids. This is particularly appealing to address clinically relevant mutations in ERG to investigate acquired LQTS by drug exposure or, vice versa, the pharmacological control of heritable LQTS.

Precise base editing in medaka pinpointed the developmental impact of novel missense mutations in four genes, *DAPK3*, *UBE2B*, *USP44*, and *PTPN11*. Those were

associated with structural CHD in parent-offspring trio exome sequencing studies (Homsy et al., 2015; Zaidi et al., 2013). We accurately modeled these human *de novo* mutations guided by ACEofBASEs and applied cytosine base editing to the conserved medaka orthologs.

While missense mutations in the kinase domain of *DAPK3* had been observed in various carcinomas (Brognard et al., 2011), there was only circumstantial evidence linking it to the heart. In biochemical analyses with isolated proteins, *DAPK3* could phosphorylate the regulatory light chain associated with cardiac myosin (*MYL2*) (Chang et al., 2010). Our base editing experiment provides the first functional evidence linking the missense mutation p.P193L, associated with conotruncal defects in humans (Zaidi et al., 2013), to CHD. Introduction of the p.P204L mutation into the corresponding position of the *dapk3* medaka ortholog resulted in looping and morphological heart defects during embryonic development.

Our base editing experiments modeling patient-based mutations to *UBE2B* and *USP44*, both involved in the post-translational modifications of histones (*H2Bub1*), show their essential role in heart looping. While the knockdown of *ube2b* in *Xenopus* did not result in an apparent phenotype (Robson et al., 2019), mutations in the gene were correlated in CHD patients. Strikingly, the introduction of the single p.R8Q missense mutation into the medaka *ube2b* ortholog resulted in aberrant L-loop phenotypes in medaka embryos.

USP44 regulates the cell cycle as deubiquitinase acting on *H2Bub1* during embryonic stem cell differentiation (Fuchs et al., 2012; Stegmeier et al., 2007). Respective mouse knockout models hint at a tumor suppressor role (Zhang et al., 2012). Interestingly, a specific glutamate residue (O1 E68, Hs E71) was associated with CHD in GWAS (Zaidi et al., 2013), demanding a precise editing for functional validation that we applied in the medaka ortholog at the corresponding position. This allowed to disentangle the pleiotropic effects observed in *USP44* null mutants and emphasize the physiological relevance of a single amino acid for the proper development of the heart.

Previous experiments in zebrafish (mRNA overexpression of pathogenic *PTPN11* variants, Bonetti et al., 2014) had argued for a potential role of *PTPN11* in cardiac development. Engineering the *PTPN11* missense mutation p.G503E residing in the protein tyrosine phosphatase (PTP) domain (Homsy et al., 2015) into the corresponding position of the medaka ortholog resulted in impaired cardiac looping,

demonstrating that this amino acid is essential for proper PTPN11 function in cardiac development.

Those four examples highlight the power of the combination of ACEofBASEs as a prediction tool and the respective base editors to instantly validate associations from human datasets. This allowed the immediate establishment of highly informative small animal models for human diseases that could be queried already in the F0 generation, underpinning the power of the approach in addressing the pathophysiology of human disease-associated DNA variants.

Although the list of successfully *in vivo* validated SNVs remains limited (Claussnitzer et al., 2020), the combined CBE-ABE (standard editing window, NGG PAM) action provides access to almost 30% of human variants in ClinVar (Gaudelli et al., 2017; Komor et al., 2016; Landrum et al., 2016) for functional interrogation. We analyzed the pathogenic ClinVar SNVs' conservation revealing that CBEs and ABEs can directly model 5896 human CDS SNVs in the orthologous medaka and zebrafish genes combined, consistent with the high fraction of conserved orthologs in zebrafish (Howe et al., 2013) and medaka (Kasahara et al., 2007). Notably, 1951 pathogenic CDS SNVs can be modeled without bystander mutations (Figure 7c). The practical use of CBE and ABE Cas variants with broadened PAM compatibilities, such as xCas9 or SpRY, is estimated to collectively expand variant validation possibilities, covering up to 95% of pathogenic ClinVar transition mutations (Anzalone et al., 2020; Hu et al., 2018; Walton et al., 2020). A recent report of NR PAM cytosine base editing in fish (Rosello et al., 2021a) opens the door for future validation experiments using expanded PAM editors. With the recent advances in the evolution of deaminases (Chen et al., 2021; Koblan et al., 2021; Kurt et al., 2021; Zhao et al., 2021), the substitution flexibility is constantly extended.

Given the high precision and efficiency, the parallel introduction of multiple edits is within reach in the near future. This will ultimately allow tackling polygenic traits in the organismal context. It will likewise provide experimental access to the regulatory genome, as recently demonstrated by combining epigenomics with adenine base editing *in vitro* that mapped and validated sickle cell disease-associated cis-regulatory elements of the fetal haemoglobin (Cheng et al., 2021).

Figure 7

Figure 7. Recapitulation of *in vitro* base editing characteristics combined with a plethora of conserved variants make fish excellent models to validate human pathogenic SNVs.

(a) Transition efficiencies for CBEs and ABE8e tested in this study in medaka across the entire protospacer are shown as mean with 95% confidence interval (CI). Each data point represents the mean efficiency for the locus (sgRNA) tested. N=represents the number of loci tested with the respective editor. n=total number of genomes used for quantification of editing efficiencies (a pool of five embryos was counted as five genomes). (a') Simplified scheme only showing mean and CI. (b) Summary of dinucleotide preference for the tested base editors, calculated for the standard editing window (4-8). Each data point represents the mean editing efficiency of the corresponding editor for a particular protospacer position. (b') Simplified scheme of overlaid dinucleotide logic. (c) Analysis of human pathogenic CDS SNVs annotated in ClinVar reveals that a remarkable portion of these SNVs have orthologous sequences in medaka or zebrafish that can be mimicked by CBEs or ABEs following editing window (4-8) and NGG PAM restrictions. Modeling SNVs mutations may be achieved with stringent criteria (no bystander mutations accepted, n=1951) or less stringent selection (allowing bystander mutations "all", n=5896). Note: the number of SNVs shown for medaka + zebrafish, corresponds to a set-up in which these species are complementing each other.

In conclusion, we demonstrate that *in vivo* base editing, critically guided by our online tool ACEofBASEs, is instantly applicable to a wide range of developmental and genetic disease studies in native genomic contexts. Validation of the plethora of genetic variants surfaced by large-scale sequencing in humans, which require effective functional testing strategies, can now be immediately addressed by ACEofBASEs guided *in vivo* base editing in fish. Future studies in model organisms focusing on the pre-clinical examination of genetic disease variants will help direct the translational discovery process to improve personalized clinical care of patients with individual genetic variant profiles.

Acknowledgements

This research was supported by the German Science funding agency (DFG, FOR2509 project 10 (WI 1824/9-1)), the European Research Council Synergy Grant IndiGene (Number 810172) and the DZHK (German Centre for Cardiovascular Research) to J.W.. J.G. was supported by the Deutsche Herzstiftung e.V. (S/02/17) and by an Add-On Fellowship for Interdisciplinary Science of the Joachim Herz Stiftung. A.C., B.W. and J.G. are members/alumni of the Heidelberg Biosciences International Graduate School (HBIGS). We thank S. Lemke, O. Hammouda, T. Tavhelidse, L. Doering, F. Farkas and all members of the Wittbrodt lab for their critical, constructive feedback on the manuscript. We thank T. Kellner for excellent technical support. We further thank M. Majewsky, E. Leist, S. Erny and A. Saraceno for expert fish husbandry.

Author contributions

A.C., J.G. and J.W. conceived the study and implemented the methodology. J.L.M. and T.T. conceptualised the ACEofBASEs tool, which J.L.M. implemented. A.C. and B.W. performed the experiments with initial contributions from J.G.. A.C. analysed the data with contributions from B.W. and J.L.M.. A.C. visualised the data. A.C., J.G. and J.W. wrote the manuscript with contributions from B.W., J.L.M. and T.T.. J.W. provided resources and supervised this work.

Competing Interests statement

All authors declare no competing interests.

References

- Abriel H, Zaklyazminskaya E V. 2013. Cardiac channelopathies: Genetic and molecular mechanisms. *Gene* **517**:1–11. doi:<https://doi.org/10.1016/j.gene.2012.12.061>
- Anzalone A V., Koblan LW, Liu DR. 2020. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* **38**:824–844. doi:10.1038/s41587-020-0561-9
- Arbab M, Shen MW, Mok B, Wilson C, Matuszek Ż, Cassa CA, Liu DR. 2020. Determinants of Base Editing Outcomes from Target Library Analysis and Machine Learning. *Cell* **182**:463-480.e30. doi:10.1016/j.cell.2020.05.037
- Arnaout R, Ferrer T, Huiskens J, Spitzer K, Stainier DYR, Tristani-Firouzi M, Chi NC. 2007. Zebrafish model for human long QT syndrome. *Proc Natl Acad Sci* **104**:11316 LP – 11321. doi:10.1073/pnas.0702724104
- Becker C, Lust K, Wittbrodt J. 2021. Igf signaling couples retina growth with body growth by modulating progenitor cell division. *Development* **148**. doi:10.1242/dev.199133
- Bert B, Chmielewska J, Bergmann S, Busch M, Driever W, Finger-Baier K, Höbner J, Köhler A, Leich N, Misgeld T, Nöldner T, Reiher A, Scharl M, Seebach-Sproedt A, Thumberger T, Schönfelder G, Grune B. 2016. Considerations for a European animal welfare standard to evaluate adverse phenotypes in teleost fish. *EMBO J* **35**:1151–1154. doi:<https://doi.org/10.15252/embj.201694448>
- Bonetti M, Overman JP, Tessadori F, Noël E, Bakkers J, den Hertog J. 2014. Noonan and LEOPARD syndrome Shp2 variants induce heart displacement defects in zebrafish. *Dev* **141**:1961–1970. doi:10.1242/dev.106310
- Brinkman EK, Chen T, Amendola M, van Steensel B. 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**:e168–e168. doi:10.1093/nar/gku936
- Brognaud J, Zhang Y-W, Puto LA, Hunter T. 2011. Cancer-Associated Loss-of-Function Mutations Implicate DAPK3 as a Tumor-Suppressing Kinase. *Cancer Res* **71**:3152–3161. doi:10.1158/0008-5472.CAN-10-3543
- Carrington B, Weinstein RN, Sood R. 2020. BE4max and AncBE4max Are Efficient in Germline Conversion of C:G to T:A Base Pairs in Zebrafish. *Cells* **9**:1690. doi:10.3390/cells9071690
- Chang AN, Chen G, Gerard RD, Kamm KE, Stull JT. 2010. Cardiac Myosin Is a Substrate for Zipper-interacting Protein Kinase (ZIPK). *J Biol Chem* **285**:5122–5126. doi:10.1074/jbc.C109.076489
- Chen JN, Haffter P, Odenthal J, Vogelsang E, Brand M, Van Eeden FJM, Furutani-Seiki M,

Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Kelsh RN, Mullins MC, Nüsslein-Volhard C. 1996. Mutations affecting the cardiovascular system and other internal organs in zebrafish. *Development* **123**:293–302. doi:10.1242/dev.123.1.293

Chen L, Park JE, Paa P, Rajakumar PD, Prekop HT, Chew YT, Manivannan SN, Chew WL. 2021. Programmable C:G to G:C genome editing with CRISPR-Cas9-directed base excision repair proteins. *Nat Commun* **12**. doi:10.1038/s41467-021-21559-9

Cheng L, Li Y, Qi Q, Xu P, Feng R, Palmer L, Chen J, Wu R, Yee T, Zhang J, Yao Y, Sharma A, Hardison RC, Weiss MJ, Cheng Y. 2021. Single-nucleotide-level mapping of DNA regulatory elements that control fetal hemoglobin expression. *Nat Genet*. doi:10.1038/s41588-021-00861-8

Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurles ME, Kathiresan S, Kenny EE, Lindgren CM, MacArthur DG, North KN, Plon SE, Rehm HL, Risch N, Rotimi CN, Shendure J, Soranzo N, McCarthy MI. 2020. A brief history of human disease genetics. *Nature* **577**:179–189. doi:10.1038/s41586-019-1879-7

Fu JL, Kanno T, Liang S-C, Matzke AJM, Matzke M. 2015. GFP Loss-of-Function Mutations in *Arabidopsis thaliana*. *G3 Genes/Genomes/Genetics* **5**:1849–1855. doi:10.1534/g3.115.019604

Fuchs G, Shema E, Vesterman R, Kotler E, Wolchinsky Z, Wilder S, Golomb L, Pribluda A, Zhang F, Haj-Yahya M, Feldmesser E, Brik A, Yu X, Hanna J, Aberdam D, Domany E, Oren M. 2012. RNF20 and USP44 Regulate Stem Cell Differentiation by Modulating H2B Monoubiquitylation. *Mol Cell* **46**:662–673. doi:10.1016/j.molcel.2012.05.023

Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR. 2017. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**:464–471. doi:10.1038/nature24644

Gierten J, Pylatiuk C, Hammouda OT, Schock C, Stegmaier J, Wittbrodt J, Gehrig J, Loosli F. 2020. Automated high-throughput heartbeat quantification in medaka and zebrafish embryos under physiological conditions. *Sci Rep* **10**:1–12. doi:10.1038/s41598-020-58563-w

Gut P, Reischauer S, Stainier DYR, Arnaout R. 2017. Little fish, big data: Zebrafish as a model for cardiovascular and metabolic disease. *Physiol Rev* **97**:889–938. doi:10.1152/physrev.00038.2016

Gutierrez-Triana JA, Tavhelidse T, Thumberger T, Thomas I, Wittbrodt B, Kellner T, Anlas K, Tsingos E, Wittbrodt J. 2018. Efficient single-copy HDR by 5' modified long dsDNA donors. *Elife* **7**:1–15. doi:10.7554/eLife.39468

Hammouda OT, Böttger F, Wittbrodt J, Thumberger T. 2019. Swift large-scale examination of directed genome editing. *PLoS One* **14**:1–11. doi:10.1371/journal.pone.0213317

Hammouda OT, Wu MY, Kaul V, Thumberger T, Wittbrodt J. 2021. *In vivo* identification and validation of novel potential predictors for human cardiovascular diseases. *bioRxiv* 2021.02.03.429563. doi:10.1101/2021.02.03.429563

Hassel D, Scholz EP, Trano N, Friedrich O, Just S, Meder B, Weiss DL, Zitron E, Marquart S, Vogel B, Karle CA, Seemann G, Fishman MC, Katus HA, Rottbauer W. 2008. Deficient zebrafish ether-à-go-go-related gene channel gating causes short-QT syndrome in zebrafish reggae mutants. *Circulation* **117**:866–875. doi:10.1161/CIRCULATIONAHA.107.752220

Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, DePalma SR, McKean D, Wakimoto H, Gorham J, Jin SC, Deanfield J, Giardini A, Porter GA, Kim R, Bilguvar K, López-Giráldez F, Tikhonova I, Mane S, Romano-Adesman A, Qi H, Vardarajan B, Ma L, Daly M, Roberts AE, Russell MW, Mital S, Newburger JW, Gaynor JW, Breitbart RE, Iossifov I, Ronemus M, Sanders SJ, Kaltman JR, Seidman JG, Brueckner M, Gelb BD, Goldmuntz E, Lifton RP, Seidman CE, Chung WK. 2015. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science (80-)* **350**:1262 LP – 1266. doi:10.1126/science.aac9396

Hoshijima K, Juryneć MJ, Grunwald DJ. 2016. Precise Editing of the Zebrafish Genome Made Simple and Efficient. *Dev Cell* **36**:654–667. doi:10.1016/j.devcel.2016.02.015

Hoshijima K, Juryneć MJ, Klatt Shaw D, Jacobi AM, Behlke MA, Grunwald DJ. 2019. Highly Efficient CRISPR-Cas9-Based Methods for Generating Deletion Mutations and F0 Embryos that Lack Gene Function in Zebrafish. *Dev Cell* **51**:645-657.e4. doi:<https://doi.org/10.1016/j.devcel.2019.10.004>

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, McLaren S, Sealy I, Caccamo M, Churcher C, Scott C, Barrett JC, Koch R, Rauch G-J, White S, Chow W, Kilian B, Quintais LT, Guerra-Assunção JA, Zhou Y, Gu Y, Yen J, Vogel J-H, Eyre T, Redmond S, Banerjee R, Chi J, Fu B, Langley E, Maguire SF, Laird GK, Lloyd D, Kenyon E, Donaldson S, Sehra H, Almeida-King J, Loveland J, Trevanion S, Jones M, Quail M, Willey D, Hunt A, Burton J, Sims S, McLay K, Plumb B, Davis J, Clee C, Oliver K, Clark R, Riddle C, Elliott D, Threadgold G, Harden G, Ware D, Begum S, Mortimore B, Kerry G, Heath P, Phillimore B, Tracey A, Corby N, Dunn M, Johnson C, Wood J, Clark S, Pelan S, Griffiths G, Smith M, Glithero R, Howden P, Barker N, Lloyd C, Stevens C, Harley J, Holt K, Panagiotidis G, Lovell J, Beasley H, Henderson C, Gordon D, Auger K, Wright D, Collins J, Raisen C, Dyer L, Leung K, Robertson L, Ambridge K, Leongamornlert D, McGuire S, Gilderthorp R, Griffiths C, Manthavadi D, Nichol S, Barker G, Whitehead S, Kay M, Brown J, Murnane C, Gray E, Humphries M, Sycamore N, Barker D, Saunders D, Wallis J, Babbage A,

Hammond S, Mashreghi-Mohammadi M, Barr L, Martin S, Wray P, Ellington A, Matthews N, Ellwood M, Woodmansey R, Clark G, Cooper JD, Tromans A, Grafham D, Skuce C, Pandian R, Andrews R, Harrison E, Kimberley A, Garnett J, Fosker N, Hall R, Garner P, Kelly D, Bird C, Palmer S, Gehring I, Berger A, Dooley CM, Ersan-Ürün Z, Eser C, Geiger H, Geisler M, Karotki L, Kirn A, Konantz J, Konantz M, Oberländer M, Rudolph-Geiger S, Teucke M, Lanz C, Raddatz G, Osoegawa K, Zhu B, Rapp A, Widaa S, Langford C, Yang F, Schuster SC, Carter NP, Harrow J, Ning Z, Herrero J, Searle SMJ, Enright A, Geisler R, Plasterk RHA, Lee C, Westerfield M, de Jong PJ, Zon LI, Postlethwait JH, Nüsslein-Volhard C, Hubbard TJP, Crollius HR, Rogers J, Stemple DL. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**:498–503. doi:10.1038/nature12111

Howe KL, Achuthan P, Allen James, Allen Jamie, Alvarez-Jarreta J, Amodé MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Charkhchi M, Cummins C, Da Rin Fioretto L, Davidson C, Dodiya K, El Houdaigui B, Fatima R, Gall A, Garcia Giron C, Grego T, Guijarro-Clarke C, Haggerty L, Hemrom A, Hourlier T, Izuogu OG, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Gonzalez Martinez J, Marugán JC, Maurel T, McMahon AC, Mohanan S, Moore B, Muffato M, Oheh DN, Paraschas D, Parker A, Parton A, Prosovetskaia I, Sakthivel MP, Salam AIA, Schmitt BM, Schuilenburg H, Sheppard D, Steed E, Szpak M, Szuba M, Taylor K, Thormann A, Threadgold G, Walts B, Winterbottom A, Chakiachvili M, Chaubal A, De Silva N, Flint B, Frankish A, Hunt SE, Ilsley GR, Langridge N, Loveland JE, Martin FJ, Mudge JM, Morales J, Perry E, Ruffier M, Tate J, Thybert D, Trevanion SJ, Cunningham F, Yates AD, Zerbino DR, Flicek P. 2021. Ensembl 2021. *Nucleic Acids Res* **49**:D884–D891. doi:10.1093/nar/gkaa942

Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X, Rees HA, Lin Z, Liu DR. 2018. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**:57–63. doi:10.1038/nature26155

Huang TP, Newby GA, Liu DR. 2021. Precision genome editing using cytosine and adenine base editors in mammalian cells. *Nat Protoc* **16**:1089–1128. doi:10.1038/s41596-020-00450-9

Hwang GH, Park J, Lim K, Kim S, Yu J, Yu E, Kim ST, Eils R, Kim JS, Bae S. 2018. Web-based design and analysis tools for CRISPR base editing. *BMC Bioinformatics* **19**. doi:10.1186/s12859-018-2585-4

Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JRJ, Joung JK. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* **31**:227–229. doi:10.1038/nbt.2501

Ingles J, Macciocca I, Morales A, Thomson K. 2020. Genetic Testing in Inherited Heart Diseases. *Hear Lung Circ* **29**:505–511. doi:<https://doi.org/10.1016/j.hlc.2019.10.014>

Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, Ahsan B, Yamada T, Nagayasu Y, Doi K, Kasai Y, Jindo T, Kobayashi D, Shimada A, Toyoda A, Kuroki Y, Fujiyama A, Sasaki T, Shimizu A, Asakawa S, Shimizu N, Hashimoto S, Yang J, Lee Y, Matsushima K, Sugano S, Sakaizumi M, Narita T, Ohishi K, Haga S, Ohta F, Nomoto H, Nogata K, Morishita T, Endo T, Shin-I T, Takeda H, Morishita S, Kohara Y. 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature* **447**:714–719. doi:10.1038/nature05846

Kassambara A. 2020. ggpubr: “ggplot2” Based Publication Ready Plots.

Kathiresan S, Srivastava D. 2012. Genetics of human cardiovascular disease. *Cell* **148**:1242–1257. doi:10.1016/j.cell.2012.03.001

Kerr R, Stevens G, Manga P, Salm S, John P, Haw T, Ramsay M. 2000. Identification of P gene mutations in individuals with oculocutaneous albinism in sub-Saharan Africa. *Hum Mutat* **15**:166–172. doi:[https://doi.org/10.1002/\(SICI\)1098-1004\(200002\)15:2<166::AID-HUMU5>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1098-1004(200002)15:2<166::AID-HUMU5>3.0.CO;2-Z)

Kluesner MG, Nedveck DA, Lahr WS, Garbe JR, Abrahante JE, Webber BR, Moriarity BS. 2018. EditR: A Method to Quantify Base Editing from Sanger Sequencing. *Cris J* **1**:239–250. doi:10.1089/crispr.2018.0014

Koblan LW, Arbab M, Shen MW, Hussmann JA, Anzalone A V, Doman JL, Newby GA, Yang D, Mok B, Replogle JM, Xu A, Sisley TA, Weissman JS, Adamson B, Liu DR. 2021. Efficient C•G-to-G•C base editors developed using CRISPRi screens, target-library analysis, and machine learning. *Nat Biotechnol*. doi:10.1038/s41587-021-00938-z

Koblan LW, Doman JL, Wilson C, Levy JM, Tay T, Newby GA, Maianti JP, Raguram A, Liu DR. 2018. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol* **36**:843–848. doi:10.1038/nbt.4172

Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**:420–424. doi:10.1038/nature17946

Komor AC, Zhao KT, Packer MS, Gaudelli NM, Waterbury AL, Koblan LW, Kim YB, Badran AH, Liu DR. 2017. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv* **3**:1–10. doi:10.1126/sciadv.aao4774

Kraeussling M, Wagner TU, Scharl M. 2011. Highly Asynchronous and Asymmetric Cleavage Divisions Accompany Early Transcriptional Activity in Pre-Blastula Medaka Embryos. *PLoS One* **6**:e21741.

Kroll F, Powell GT, Ghosh M, Gestri G, Antinucci P, Hearn TJ, Tunbak H, Lim S, Dennis HW, Fernandez JM, Whitmore D, Dreosti E, Wilson SW, Hoffman EJ, Rihel J. 2021. A simple and effective f0 knockout method for rapid screening of behaviour and other complex phenotypes, *eLife*. doi:10.7554/eLife.59683

Kupperman E, An S, Osborne N, Waldron S, Stainier DYR. 2000. A sphingosine-1-phosphate receptor regulates cell migration during vertebrate heart development. *Nature* **406**:192–195. doi:10.1038/35018092

Kurt IC, Zhou R, Iyer S, Garcia SP, Miller BR, Langner LM, Grünewald J, Joung JK. 2021. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat Biotechnol* **39**:41–46. doi:10.1038/s41587-020-0609-x

Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W, Maglott DR. 2016. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* **44**:D862–D868. doi:10.1093/nar/gkv1222

Li X, Zhang G, Ngo N, Zhao X, Kain SR, Huang C-C. 1997. Deletions of the Aequorea victoria Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence. *J Biol Chem* **272**:28545–28549. doi:https://doi.org/10.1074/jbc.272.45.28545

Li Y, Liu Y, Yang H, Zhang T, Naruse K, Tu Q. 2020. Dynamic transcriptional and chromatin accessibility landscape of medaka embryogenesis. *Genome Res* **30**:924–937. doi:10.1101/gr.258871.119

Lischik CQ, Adelmann L, Wittbrodt J. 2019. Enhanced in vivo-imaging in medaka by optimized anaesthesia, fluorescent protein selection and removal of pigmentation. *PLoS One* **14**:1–19. doi:10.1371/journal.pone.0212956

Loosli F, Köster RW, Carl M, Kühnlein R, Henrich T, Mücke M, Krone A, Wittbrodt J. 2000. A genetic screen for mutations affecting embryonic development in medaka fish (*Oryzias latipes*). *Mech Dev* **97**:133–139. doi:https://doi.org/10.1016/S0925-4773(00)00406-8

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. 2016. The Ensembl Variant Effect Predictor. *Genome Biol* **17**:122. doi:10.1186/s13059-016-0974-4

Meyer H V, Dawes TJW, Serrani M, Bai W, Tokarczuk P, Cai J, de Marvao A, Henry A, Lumbers RT, Gierten J, Thumberger T, Wittbrodt J, Ware JS, Rueckert D, Matthews PM, Prasad SK, Costantino ML, Cook SA, Birney E, O'Regan DP. 2020. Genetic and functional insights into the fractal structure of the heart. *Nature* **584**:589–594. doi:10.1038/s41586-020-2635-8

Nakajima T, Furukawa T, Hirano Y, Tanaka T, Sakurada H, Takahashi T, Nagai R, Itoh T,

Katayama Y, Nakamura Y, Hiraoka M. 1999. Voltage-shift of the current activation in HERG S4 mutation (R534C) in LQT2. *Cardiovasc Res* **44**:283–293. doi:10.1016/S0008-6363(99)00195-9

Ng CA, Perry MD, Tan PS, Hill AP, Kuchel PW, Vandenberg JI. 2012. The S4-S5 linker acts as a signal integrator for hERG K⁺ channel activation and deactivation gating. *PLoS One* **7**. doi:10.1371/journal.pone.0031640

Patterson GH, Knobel SM, Sharif WD, Kain SR, Piston DW. 1997. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J* **73**:2782–2790. doi:https://doi.org/10.1016/S0006-3495(97)78307-3

Pierpont ME, Brueckner M, Chung WK, Garg V, Lacro R V, McGuire AL, Mital S, Priest JR, Pu WT, Roberts A, Ware SM, Gelb BD, Russell MW, null null. 2018. Genetic Basis for Congenital Heart Disease: Revisited: A Scientific Statement From the American Heart Association. *Circulation* **138**:e653–e711. doi:10.1161/CIR.0000000000000606

Qin W, Lu X, Liu Y, Bai H, Li S, Lin S. 2018. Precise A•T to G•C base editing in the zebrafish genome. *BMC Biol* **16**:139. doi:10.1186/s12915-018-0609-1

R Core Team. 2020. R: A Language and Environment for Statistical Computing.

Ravindran S. 2019. Got mutation? “Base editors” fix genomes one nucleotide at a time. *Nature* **575**:553–555. doi:10.1038/d41586-019-03536-x

Rebbeck TR, Kanetsky PA, Walker AH, Holmes R, Halpern AC, Schuchter LM, Elder DE, Guerry D. 2002. P Gene as an Inherited Biomarker of Human Eye Color. *Cancer Epidemiol Biomarkers & Prev* **11**:782 LP – 784.

Rembold M, Lahiri K, Foulkes NS, Wittbrodt J. 2006. Transgenesis in fish: efficient selection of transgenic fish by co-injection with a fluorescent reporter construct. *Nat Protoc* **1**:1133–1139. doi:10.1038/nprot.2006.165

Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, Thuronyi BW, Wilson C, Koblan LW, Zeng J, Bauer DE, Doudna JA, Liu DR. 2020. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat Biotechnol* **38**:883–891. doi:10.1038/s41587-020-0453-z

Robson A, Makova SZ, Barish S, Zaidi S, Mehta S, Drozd J, Jin SC, Gelb BD, Seidman CE, Chung WK, Lifton RP, Khokha MK, Brueckner M. 2019. Histone H2B monoubiquitination regulates heart development via epigenetic control of cilia motility. *Proc Natl Acad Sci U S A* **116**:14049–14054. doi:10.1073/pnas.1808341116

Rosello M, Serafini M, Mione MC, Concordet J-P, Del Bene F. 2021a. Disease modeling by efficient genome editing using a near PAM-less base editor *in vivo*. *bioRxiv* 2021.06.28.450169. doi:10.1101/2021.06.28.450169

Rosello M, Vougny J, Czarny F, Mione MC, Concordet J-P, Albadri S, Del Bene F. 2021b.

Precise base editing for the in vivo study of developmental signaling and human pathologies in zebrafish. *Elife* **10**. doi:10.7554/eLife.65552

Rupp RA, Snider L, Weintraub H. 1994. Xenopus embryos regulate the nuclear localization of XMyoD. *Genes Dev* **8**:1311–1323. doi:10.1101/gad.8.11.1311

Sanguinetti MC, Xu QP. 1999. Mutations of the S4-S5 linker alter activation properties of HERG potassium channels expressed in Xenopus oocytes. *J Physiol* **514**:667–675. doi:<https://doi.org/10.1111/j.1469-7793.1999.667ad.x>

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**:676–682. doi:10.1038/nmeth.2019

Sehnert AJ, Huq A, Weinstein BM, Walker C, Fishman M, Stainier DYR. 2002. Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. *Nat Genet* **31**:106–110. doi:10.1038/ng875

Siegner SM, Karasu ME, Schröder MS, Kontarakis Z, Corn JE. 2021. PnB Designer: a web application to design prime and base editor guide RNAs for animals and plants. *BMC Bioinformatics* **22**:1–12. doi:10.1186/s12859-021-04034-6

Sifrim A, Hitz MP, Wilsdon A, Breckpot J, Turki SHA, Thienpont B, McRae J, Fitzgerald TW, Singh T, Swaminathan GJ, Prigmore E, Rajan D, Abdul-Khaliq H, Banka S, Bauer UMM, Bentham J, Berger F, Bhattacharya S, Bu'Lock F, Canham N, Colgiu IG, Cosgrove C, Cox H, Daehnert I, Daly A, Danesh J, Fryer A, Gewillig M, Hobson E, Hoff K, Homfray T, Kahlert AK, Ketley A, Kramer HH, Lachlan K, Lampe AK, Louw JJ, Manickara AK, Manase D, McCarthy KP, Metcalfe K, Moore C, Newbury-Ecob R, Omer SO, Ouwehand WH, Park SM, Parker MJ, Pickardt T, Pollard MO, Robert L, Roberts DJ, Sambrook J, Setchfield K, Stiller B, Thornborough C, Toka O, Watkins H, Williams D, Wright M, Mital S, Daubeney PEF, Keavney B, Goodship J, Abu-Sulaiman RM, Klaassen S, Wright CF, Firth H V., Barrett JC, Devriendt K, Fitzpatrick DR, Brook JD, Hurles ME. 2016. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nat Genet* **48**:1060–1065. doi:10.1038/ng.3627

Stainier DY, Fouquet B, Chen JN, Warren KS, Weinstein BM, Meiler SE, Mohideen MA, Neuhauss SC, Solnica-Krezel L, Schier AF, Zwartkruis F, Stemple DL, Malicki J, Driever W, Fishman MC. 1996. Mutations affecting the formation and function of the cardiovascular system in the zebrafish embryo. *Development* **123**:285–292. doi:10.1242/dev.123.1.285

Stegmeier F, Rape M, Draviam VM, Nalepa G, Sowa ME, Ang XL, McDonald III ER, Li MZ,

Hannon GJ, Sorger PK, Kirschner MW, Harper JW, Elledge SJ. 2007. Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. *Nature* **446**:876–881. doi:10.1038/nature05694

Stemmer M, Thumberger T, del Sol Keyer M, Wittbrodt J, Mateo JL. 2015. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* **10**:e0124633.

Thumberger T, Tavhelidse T, Gutierrez-Triana JA, Medert R, Cornean A, Welz B, Freichel M, Wittbrodt J. 2021. hei-tag: a highly efficient tag to boost targeted genome editing. *bioRxiv* 2021.05.27.445956. doi:10.1101/2021.05.27.445956

Thuronyi BW, Koblan LW, Levy JM, Yeh WH, Zheng C, Newby GA, Wilson C, Bhaumik M, Shubina-Oleinik O, Holt JR, Liu DR. 2019. Continuous evolution of base editors with expanded target compatibility and improved activity. *Nat Biotechnol* **37**:1070–1079. doi:10.1038/s41587-019-0193-0

Walton RT, Christie KA, Whittaker MN, Kleinstiver BP. 2020. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science (80-)* **368**:290–296. doi:10.1126/science.aba8853

Wang W, MacKinnon R. 2017. Cryo-EM Structure of the Open Human Ether-à-go-go-Related K⁺ Channel hERG. *Cell* **169**:422-430.e10. doi:10.1016/j.cell.2017.03.048

Westerfield M. 2000. The zebrafish book: a guide for the laboratory use of zebrafish, 4th ed. ed. Univ. of Oregon Press, Eugene.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham H. 2011. The Split-Apply-Combine Strategy for Data Analysis. *J Stat Softw* **40**:1–29.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the tidyverse. *J Open Source Softw* **4**:1686. doi:10.21105/joss.01686

Wickham H, François R, Henry L, Müller K. 2020. dplyr: A Grammar of Data Manipulation.

Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, Romano-Adesman A, Bjornson RD, Breitbart RE, Brown KK, Carrierio NJ, Cheung YH, Deanfield J, Depalma S, Fakhro KA, Glessner J, Hakonarson H, Italia MJ, Kaltman JR, Kaski J, Kim R, Kline JK, Lee T, Leipzig J, Lopez A, Mane SM, Mitchell LE, Newburger JW, Parfenov M, Pe’Er I, Porter G, Roberts AE, Sachidanandam R, Sanders SJ, Seiden HS, State MW, Subramanian S, Tikhonova IR, Wang W, Warburton D, White PS, Williams IA, Zhao H, Seidman JG, Brueckner M, Chung WK, Gelb BD, Goldmuntz E, Seidman CE, Lifton RP. 2013. De novo mutations in histone-modifying genes in congenital heart disease. *Nature*

498:220–223. doi:10.1038/nature12141

Zhang M, Liu J, Tseng G-N. 2004. Gating Charges in the Activation and Inactivation Processes of the hERG Channel . *J Gen Physiol* **124**:703–718. doi:10.1085/jgp.200409119

Zhang Y, Foreman O, Wigle DA, Kosari F, Vasmatzis G, Salisbury JL, van Deursen J, Galardy PJ. 2012. USP44 regulates centrosome positioning to prevent aneuploidy and suppress tumorigenesis. *J Clin Invest* **122**:4362–4374. doi:10.1172/JCI63084

Zhang Y, Qin W, Lu X, Xu J, Huang H, Bai H, Li S, Lin S. 2017. Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat Commun* **8**:118. doi:10.1038/s41467-017-00175-6

Zhang YH, Colenso CK, Sessions RB, Dempsey CE, Hancox JC. 2011. The hERG K⁺ channel S4 domain L532P mutation: Characterization at 37 °C. *Biochim Biophys Acta - Biomembr* **1808**:2477–2487. doi:10.1016/j.bbamem.2011.07.001

Zhao D, Li J, Li S, Xin X, Hu M, Price MA, Rosser SJ, Bi C, Zhang X. 2021. Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat Biotechnol* **39**:35–40. doi:10.1038/s41587-020-0592-2

Zhao Y, Shang D, Ying R, Cheng H, Zhou R. 2020. An optimized base editor with efficient C-to-T base editing in zebrafish. *BMC Biol* **18**:190. doi:10.1186/s12915-020-00923-z

Materials and Methods

Key Resources Table

Reagent type, (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain (Oryzias latipes)	Cab	(Loosli et al., 2000)		medaka Southern wild type population
Strain (Oryzias latipes)	Cab (myl7::EGFP)	(Gierten et al., 2020)		See Materials and methods
Strain (Oryzias latipes)	HdrR (myl7::EGFP myl7::H2A-mCherry)	(Hammouda et al., 2021)		See Materials and methods
Strain (Danio rerio)	AB	ZIRC	ZFIN: ZBD-GENO-960809-7)	Wildtype zebrafish strain

Recombinant DNA reagent	pCMV_BE4-Gam (plasmid)	Addgene	Addgene plasmid #100806; http://n2t.net/addgene:100806 ; RRID:Addgene_100806	
Recombinant DNA reagent	pGGEV_4_BE4-Gam (plasmid)	(Thumberger et al., 2021)		
Recombinant DNA reagent	pCMV_AncBE4max, (plasmid)	Addgene	Addgene plasmid #112094; http://n2t.net/addgene:112094 ; RRID:Addgene_112094	
Recombinant DNA reagent	pBT281(evoAPOBE C1-BE4max) (plasmid)	Addgene	Addgene plasmid #122611 ; http://n2t.net/addgene:122611 ; RRID:Addgene_122611	
Recombinant DNA reagent	pCS2+ (plasmid)	(Rupp et al., 1994)		
Recombinant DNA reagent	pCS2+_evoBE4max	This paper		See Materials and methods
Recombinant DNA reagent	pBABE8e (plasmid)	Addgene	Addgene plasmid #138489; http://n2t.net/addgene:138489 ; RRID:Addgene_138489	
Recombinant DNA reagent	DR274 (plasmid)	Addgene	Addgene plasmid #42250; http://n2t.net/addgene:42250 ; RRID:Addgene_42250	
Gene (<i>Oryzias latipes</i>)	<i>oca2</i>	Ensemble genome browser	Ensemble (release 96): ENSORLG00000015893 CDS	
Gene (<i>Danio rerio</i>)	<i>oca2</i>	Ensemble genome browser	Ensemble (release 103): ENSDARG00000061303.8 CDS	
Gene (<i>Homo sapiens</i>)	<i>oca2</i>	Ensemble genome browser	Ensemble (release 103): ENSG00000104044.16 CDS	
Gene (<i>Oryzias latipes</i>)	<i>tnnt2a</i>	Ensemble genome browser	Ensemble (release 95): ENSORLG00000024544.1 CDS	
Gene (<i>Oryzias latipes</i>)	<i>kcnh6a</i>	Ensemble genome browser	Ensemble (release 93): ENSORLG00000002317.1 CDS	Encodes <i>O/</i> ERG
Gene (<i>Danio rerio</i>)	<i>kcnh6a</i>	Ensemble genome browser	Ensemble (release 103): ENSDARG00000001803.12 CDS	Encodes <i>Dr</i> ERG
Gene (<i>Homo sapiens</i>)	<i>kcnh2</i>	Ensemble genome browser	Ensemble (release 103): ENSG00000055118.16 CDS	Encodes <i>Hs</i> ERG
Gene (<i>Oryzias latipes</i>)	<i>tnnt2c</i>	Ensemble genome browser	Ensemble (release 93): ENSORLG00000016386.1 CDS	

Gene (<i>Oryzias latipes</i>)	<i>s1pr2</i>	Ensemble genome browser	Ensemble (release 93): ENSORLG00000005560.1 CDS	
Gene (<i>Oryzias latipes</i>)	<i>dapk3</i>	Ensemble genome browser	Ensemble (release 103): ENSORLG00000017965.2 CDS	
Gene (<i>Homo sapiens</i>)	<i>dapk3</i>	Ensemble genome browser	Ensemble (release 103): ENSG00000167657.14 CDS	
Gene (<i>Oryzias latipes</i>)	<i>ube2b</i>	Ensemble genome browser	Ensemble (release 103): ENSORLG00000000951.2 CDS	
Gene (<i>Homo sapiens</i>)	<i>ube2b</i>	Ensemble genome browser	Ensemble (release 103): ENSG00000119048, <u>CCDS4174</u>	
Gene (<i>Oryzias latipes</i>)	<i>usp44</i>	Ensemble genome browser	Ensemble (release 103): ENSORLG00000016627.3 CDS	
Gene (<i>Homo sapiens</i>)	<i>usp44</i>	Ensemble genome browser	Ensemble (release 103): ENSG00000136014.12 CDS	
Gene (<i>Oryzias latipes</i>)	<i>ptpn11</i>	Ensemble genome browser	Ensemble (release 103): ENSORLG00000000470.2 CDS	<i>ptpn11a</i>
Gene (<i>Homo sapiens</i>)	<i>ptpn11</i>	Ensemble genome browser	Ensemble (release 103): ENSG00000179295.18 CDS	
Recombinant DNA reagent	oligonucleotides	Eurofins Genomics		PCR, sgRNA cloning, Sanger sequencing
Commercial kit	NEBuilder HiFi DNA Assembly Cloning kit	New England Biolabs	Catalog #E5520S	
Commercial kit	mMessage mMachine Sp6 Transcription Kit	Thermo Fisher Scientific	Catalog #AM1340	
Commercial kit	mMessage mMachine T7 Transcription Kit	Thermo Fisher Scientific	Catalog #AM1344	
Commercial kit	T7 MEGAscript Kit	Thermo Fisher Scientific	Catalog #AM1334	
Commercial kit	InnuPREP Gel Extraction Kit	Analytik Jena	Catalog #845-KS-5030250	
Commercial kit	Monarch DNA Gel Extraction Kit	New England Biolabs	Catalog #T1020	
Commercial kit	RNeasy Mini Kit	Qiagen	Catalog #74106	

Peptide, recombinant protein	Q5 High-Fidelity DNA Polymerase	New England Biolabs	Catalog #M0491	
Chemical compound, drug	2,3-Butanedione 2-monoxime (BDM)	Abcam	Catalog #ab120616	
Chemical compound, drug	N-Phenylthiourea (PTU)	Sigma-Aldrich	Catalog #P7629	
Chemical compound, drug	Ethyl 3-aminobenzoate methanesulfonate salt (Tricaine)	Sigma-Aldrich	Catalog #A5040	
Software, algorithm	ACEofBASEs	This paper	https://aceofbases.cos.uni-heidelberg.de	
Software, algorithm	EditR	(Kluesner et al., 2018)	https://moriaritylab.shinyapps.io/editr_v10/	
Software, algorithm	TIDE	(Brinkman et al., 2014)	http://shinyapps.datacurators.nl/tide/	
Software, algorithm	Geneious	Biomatters	Version 8.1.9	
Software, algorithm	Fiji distribution of ImageJ	(Schindelin et al., 2012)	Version 2.0.0	
Software, algorithm	Adobe Illustrator	Adobe	Version 23.2.1	
Software, algorithm	R, R studio	(R Core Team, 2020)	https://www.R-project.org/	
Software, algorithm	R package	(Wickham et al., 2019)	Tidyverse	
Software, algorithm	R package	(Wickham, 2016)	ggplot2	
Software, algorithm	R package	(Kassambara, 2020)	ggpubr	
Software, algorithm	R package	(Wickham, 2011)	plyr	
Software, algorithm	R package	(Wickham et al., 2020)	dplyr	

1052

1053 Fish lines and husbandry

1054 Medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*) stocks were maintained (fish husbandry,
1055 permit number 35–9185.64/BH Wittbrodt) and experiments (permit numbers 35–9185.81/G-
1056 145/15 and 35-9185.81/G-271/20 Wittbrodt) were performed in accordance with local animal
1057 welfare standards (Tierschutzgesetz §11, Abs. 1, Nr. 1) and with European Union animal
1058 welfare guidelines (Bert et al., 2016). Fish were maintained in closed stocks and constant

recirculating systems at 28°C on a 14 h light/10 h dark cycle. The fish facility is under the supervision of the local representative of the animal welfare agency. The following medaka lines: Cab as wild-type (Loosli et al., 2000), Cab (*myl7::EGFP*) (Gierten et al., 2020), HdrR (*myl7::EGFP myl7::H2A-mCherry*) (Hammouda et al., 2021). The zebrafish line AB was used in this study as wild-type.

Plasmids

To generate the pCS2+_evoBE4max plasmid, pBT281(evoAPOBEC1-BE4max; Addgene plasmid #122611, a gift from David Liu; Thuronyi et al., 2019) was assembled into pCS2+ (Rupp et al., 1994) by NEBuilder HiFi DNA Assembly (NEB) using Q5 polymerase PCR products (NEB). pCMV_AncBE4max (Addgene plasmid #112094; Koblan et al., 2018) and pBABE8e (Addgene plasmid #138489; Richter et al., 2020) are gifts from David Liu. pGGEV_4_BE4-Gam was used as previously published (Thumberger et al., 2021).

Primer name	primer sequence 5'-3'
forward	
pCS2+ backbone	GCCTCTAGAACTATAGTGAGTCG
evoBE4max fragment 1	GTTCTTTTTCGAGGATCCCATTACCATGAAACGGACAGCCGAC
evoBE4max fragment 2	CAAGGACAAGGACTTCCTG
reverse	
pCS2+ backbone	ATGGGATCCTGCAAAAAGAACAAG
evoBE4max fragment 1	CAGGAAGTCCTTGTCTCTG
evoBE4max fragment 2	CTCACTATAGTTCTAGAGGCTTAGACTTTCCTCTTCTTCTTGG

sgRNA design and synthesis

sgRNAs were designed with ACEofBASEs (<https://aceofbases.cos.uni-heidelberg.de>), applying the following parameters: presence of at least one A or C nucleotide in the respective base editing window of the respective BaseEditor variant, limitation of maximal 2 mismatches in the core region (set to 12 nucleotides adjacent to PAM) or maximal 4 mismatches in the entire spacer sequence. All sgRNA target sites in the query sequence were evaluated for potential off-targets against the respective genome (Japanese medaka HdrR (*Oryzias latipes*) Ensembl V 103; Zebrafish (*Danio rerio*) Ensemble V 103) and sgRNAs were selected for the desired codon change. Oligos were designed and selected with substitution (i.e. replacement of the two most 5' nucleotides with Gs if necessary, to foster T7 *in vitro* transcription) for transcription from DR274 (DR274 was a gift from Keith Joung, Addgene plasmid #42250; Hwang et al., 2013).

sgRNA description	target site [PAM]	Reference
<i>oca2-Q333</i>	GAAACCCAGGTGGCCATTGC[AGG]	(Lischik et al., 2019)
<i>oca2-Q256</i>	GATCCAAGTGGAGCAGACTG[AGG]	(Lischik et al., 2019)
<i>oca2-T306</i>	CACAATCCAGGCCTTCCTGC[AGG]	
<i>dr_oca2_L293</i>	GTACAGCGACTGGTTAGTCA[TGG]	(Hammouda et al., 2019)
<i>tnnt2a_Q114</i>	AGAGCGCCAAAAACGTCTTG[AGG]	(Meyer et al., 2020)
<i>kcnh6a_Q11</i>	GGCGCTCCAGAACACCTATT[TGG]	
<i>kcnh6a_T507</i>	TGAAGACAGCCCGACTGCTC[AGG]	
<i>kcnh6a_R509</i>	GACAGCCCGACTGCTCAGGT[TGG]	
<i>kcnh6a_L511/R512</i>	ACTGCTCAGGTTGGTGCGAG[TGG]	
<i>kcnh6a_D521/R522</i>	CTGGACCGTTACTCGGAGTA[CGG]	
<i>tnnt2c_R112</i>	AAGCACGAGTGGCTGAGGAG[AGG]	
<i>s1pr2_R150</i>	AACATAGCGCTCTATAGCTA[TGG]	
<i>s1pr2_R167</i>	TGCCGCATGTTTCTGCTGAT[AGG]	
<i>GFP_C71</i>	AGCACTGCACGCCGTAGGTC[AGG]	(Hammouda et al., 2021)
<i>dapk3_P204</i>	CGAGCCGCTCGGCCTGGAGG[CGG]	
<i>ube2b_R8</i>	TAGCCGTCTTCTTGCTGGTG[TGG]	
<i>usp44_E68</i>	TAACCTCCATAGCTAACGGG[TGG]	
<i>ptpn11_G504</i>	CCGCTCCAAGCGCTCGGGGA[TGG]	

Base editing sgRNA rank score

To evaluate sgRNAs and rank them we used three characteristics, efficiency, dinucleotide preference and editing window. The following describes how we, based on empirical data published previously and confirmed in this study, attributed scores.

In total we considered these, for three CBEs: BE4-Gam, ancBE4max and evoBE4max and ABE8e, the ABE we tested.

Our data suggests that BE4-Gam demonstrates efficiencies much lower than the three second generation editors tested, ancBE4max, evoBE4max and ABE8e. We therefore ascribed 5 points (BE4-Gam) and 10 points (second gen. editors) – the starting value points, when selecting a specific editor in the drop-down menu. To include dinucleotide preferences in our score, we decided to apply penalties for such dinucleotide contexts that are disfavoured when used with the selected editor. Moreover, we considered base editing windows. We considered two types of windows, such which observed editing (usually broader) and those with optimal efficiency. Below, a summary of penalty scores is provided.

Editor	Efficiency (E)	Dinucleotide context penalty, DP (penalty factor, PF)				Editing window on protospacer: peak (observed) activity	Editing window penalty (for outside of peak window editing) WP (PF = 0.8)
		T	C	A	G		
BE4-Gam	5	0 (0)	2 (0.4)	3 (0.6)	4 (0.8)	3-10 (4-8)	4
ancBE4max	10	0 (0)	2 (0.2)	2 (0.2)	8 (0.8)	3-9 (4-7)	8
evoBE4max	10	0 (0)	2 (0.2)	8 (0.8)	2 (0.2)	1-11 (4-8)	8
ABE8e	10	0 (0)	0 (0)	0 (0)	0 (0)	3-11 (4-8)	8

We identified the cytosine along the protospacer that received the highest score value and ranked this sgRNA accordingly (Figure 1-supplement 2).

$$S_{efficiency} = E - DP - WP$$

S = score.

E = efficiency

DP = dinucleotide context penalty = E x penalty factor (PF)

WP = Editing window penalty = E x penalty factor (PF)

***In vitro* transcription of mRNA**

pGGEV_4_BE4-Gam was linearised with SpeI and pCS2+_evoBE4max was linearised with NotI, then mRNA was transcribed *in vitro* for both with the mMessage mMachine Sp6 Transcription Kit. pCMV_AncBE4max and pBABE8e were both linearized with SapI and mRNA was transcribed *in vitro* with mMessage mMachine T7 Transcription Kit.

Microinjection

Medaka 1-cell stage embryos were injected into the cytoplasm as previously described (Rembold et al., 2006). Zebrafish 1-cell stage embryos were injected into the yolk. For medaka, injection solutions contained 150 ng/μl base editor mRNA, 30 ng/μl of the respective sgRNAs and 20 ng/μl GFP mRNA as injection tracer. Control siblings were injected with 20 ng/μl GFP mRNA and 30 ng/μl sgRNA only or rather 20 ng/μl GFP mRNA and 150 ng/μl base editor mRNA only. For zebrafish injections, injection mixes contained 360 pg ABE8e mRNA, 40 pg *oca2-L293* sgRNA and 12 pg GFP mRNA as injection tracer. Injected embryos were incubated at 28°C in zebrafish medium (Westerfield, 2000) or medaka embryo rearing medium (ERM) (Becker et al., 2021) and selected for GFP expression 7 hours post injection.

Genotyping

Single or pools of five randomly selected embryos from respective injection experiments and adult fin clips were lysed in DNA extraction buffer (0.4 M Tris/HCl pH 8.0, 0.15 M NaCl, 0.1%

SDS, 5 mM EDTA pH 8.0, 1 mg/ml proteinase K) at 60°C overnight. Proteinase K was inactivated at 95°C for 20 min and the solution was diluted 1:2 with nuclease-free water. Genomic DNA was precipitated in 300 mM sodium acetate and 3x vol. absolute ethanol at 20,000 x g at 4°C. Precipitated genomic DNA was resuspended in TE buffer (10 mM Tris pH 8.0, 1 mM EDTA in RNase-free water) and stored at 4°C. Genotyping-PCR was performed in 1x Q5 reaction buffer, 200 µM dNTPs, 200 µM forward and reverse primer, 1 µl precipitated DNA sample and 0.3 U Q5 polymerase (NEB).

oligo pair description	Fwd (5'-3')	Rev (5'-3')	Reference
oca2_T306&Q33_seq	GTAAAAACAGTTTCTT AAAAAGAACAGGA	AGCAGAAGAAATGACT CAACATTTTG	(Hammouda et al., 2019)
oca2_CBE_ABE_OT1_seq	CTCTGGTTACACAAT GCGCG	ACAGTAGCATGCAGGC TCTC	
oca2_CBE_ABE_OT2_seq	TTGTGATGCTGCTGTT GCAC	CCCTTAATGGACGAGC AGCA	
oca2_CBE_OT3_seq	AGTGTCTGGATTGGA TCAGTAGATG	GTGCCTGACCACTCTG ACAT	
oca2_ABE_OT3_seq	AGAGTGGGACTTTAA AGATGCACA	ACTTGTGCAGCACTTT GGATG	
dr_oca2_L293_seq	ACAGGTGCTGTATAA TTGGACCAT	AAAGAGTGGTCATAAA CGGCTACT	(Hammouda et al., 2019)
tnnt2a_Q114_seq	TGGAGAAAGACCTGA TGGAGC	TTCCCGCTCCTCTTCT CTGT	
kcnh6a_Q11_seq	ACATCCTGCATCTGC CATCG	GCAGGTGCAGTGAACC AAAA	
kcnh6a_S4domain_seq	GCTTTGCAAGGTATA GAGCACAG	AACGTTGCCAAAACCC ACAC	
tnnt2c_R112_seq	GTGCCTAACATGGTC CCTCC	ACCTCTGGTGGTCACT GACT	
s1pr2_R150_seq	CCTGGTTCTGATGGC TGTGT	CCCAGCACTATTGTGA CCGT	
GFP_C71_seq	GTGAGCAAGGGCGA GGAGCT	CTTGACAGCTCGTCC ATGC	(Gutierrez-Triana et al., 2018)
dapk3_P204_seq	CCTTAAGGAGGCAGC GAGTC	ACAGACATGAGTGTGG GCTG	
ube2b_R8_seq	AGGCGTTTTAATTGAC ATTTTGACG	CCTGTCTGGCTTCATA GACTGT	
USP44_E68_seq	TCTAGCTTTTTGGCTC CCCG	CTGTAGCTCCTGTGCT CCAC	
ptpn11_G504_seq	ACCTTTTCTCTGAACT GTCGTGT	AGGTCGGACAGCGAG TACT	

The conditions were: 98°C 2 min, 30 cycles of 98°C 30 s, annealing for 20 s and 72°C 30 s per kb, and a final extension time of 5 min at 72°C. Following agarose gel electrophoresis, amplicons were gel purified with the InnuPREP Gel Extraction Kit (Analytik Jena) or the Monarch DNA Gel Extraction Kit (NEB), and submitted for Sanger sequencing (Eurofins Genomics). Sequence analysis was performed in Geneious (R8) and the transition rates of base editing experiments were estimated using EditR (Kluesner et al., 2018). For sequence reads with observed background signal we additionally used TIDE (Brinkman et al., 2014) to estimate putative indel frequencies. For one editing experiment we then estimated overall indel frequencies, as follows: $\text{sum}(\text{individual indel rates}) / \text{total samples sequenced}$ (e.g. if one out of five sequencing reads showed 20% indel rates, overall indel frequency was 4%) (Tables 1-3).

Imaging

Gross morphology and heart dynamics of embryos were assessed with a Nikon SMZ18 Stereomicroscope equipped with a Nikon DS-Ri1 camera. For *in vivo* imaging of hearts medaka embryos were kept in 5x PTU in 1x ERM from 5-7 dpf to block thoracic and abdominal pigmentation. To capture morphological phenotypes, embryos were anesthetized in 1x Tricaine diluted in 1x ERM with subsequent induction of cardiac arrest by treatment with 50 mM BDM 40-60 min. For confocal imaging embryos were mounted laterally or ventrally (indicated in respective Figure legends) on glass-bottomed Petri dishes (MatTek Corporation, Ashland, MA) in 1% low melting agarose supplemented with 30 mM BDM and 1x Tricaine. Hearts of live embryos were imaged with a Sp8 confocal microscope (Leica) with 10x (air) or 20x (glycerol) objectives.

Analysis of SNVs in ClinVar

The ClinVar vcf file was downloaded from https://ftp.ncbi.nlm.nih.gov/pub/clinvar/vcf_GRCh38/ (date 2021-04-04). From this file, only the variants annotated as pathogenic (CLNSIG=Pathogenic) and single nucleotide (CLNVC=single_nucleotide_variant) were considered. Furthermore, as SNVs affecting the coding sequence were counted the variants annotated with any of these Sequence Ontology IDs: SO:0001578, SO:0001582, SO:0001583, SO:0001587. To identify which SNVs can be modelled in medaka or zebrafish, the REST API from Ensembl (version 103, Howe et al., 2021) to retrieve the orthologous position defined by LastZ alignments was used. Only SNV whose coordinates were aligned uniquely and in which the aligned base is the identical in medaka or zebrafish to human were kept. To obtain the molecular consequence of the inferred SNVs in fish, Variant Effect Predictor from the Ensembl web site was used (McLaren et al., 2016).

Analysis and data visualisation

Images were processed with the Fiji distribution of ImageJ (Schindelin et al., 2012). Analysis and graphical data visualisation were performed in R with the Tidyverse, ggplot2, ggpubr, plyr, and dplyr packages (Kassambara, 2020; R Core Team, 2020; Wickham, 2016, 2011; Wickham et al., 2020, 2019). Figures were assembled in Adobe Illustrator. Sample size (n) and number of independent experiments are mentioned in every figure/figure legend or the main text. No statistical methods were used to predetermine sample sizes. The experimental groups were allocated randomly, and no blinding was done during allocation.

Supporting Figures

Figure 1 - Supplement 1

Figure 1-supplement 1. Basic mechanism of cytosine base editing exemplifying the base editing principle. (a) The cytosine base editor (CBE) complex will be guided to the genomic locus of interest by the selected sgRNA. While the protospacer adjacent motif (PAM) is essential for target recognition the editing window on the protospacer is, however, more volatile depending on the editor used and resides between positions 4-8 for standard CBEs and ABEs. (b-d) Mechanism of CBE-DNA binding, DNA opening and hydrolytic deamination. Following cellular DNA repair or replication processes target cytosines are converted to thymines. Note that within the hypothetical editing window any cytosine may be deaminated.

Figure 1 - Supplement 2

Figure 1-supplement 2. sgRNA score explanation and example. (a) Exemplified by three potential sgRNAs, the scoring system penalizes cytosines or adenines to be edited (1) outside the editing window and (2) where the dinucleotide context is unfavorable. See Methods section for details. Note that only the best scoring cytosine or adenine is considered for the ranking. Here, both #2 and #3 score equally well. #2 or #3: favorable TC (C₄ or C₇) context within optimal editing window; #1: unfavorable AC and CC within optimal editing window (**continued on next page**).

Figure 1-supplement 2. (b) Shows sgRNA ranking exemplified by using the cytosine base editor BE4-Gam and the input sequence as in Figure 1a. **(b')** Note the different order of sgRNAs compared to Figure 1b. **(b'')** The TC context in the optimal editing window (green arrow) strongly favors T1, whereas the CC context in the observed editing window (magenta arrow) predicts reduced efficiencies for T26.

Figure 1 - Supplement 3

Figure 1-supplement 3. Details on selected sgRNA in ACEofBASEs. Overview of sgRNA *oca2-Q333* with target sequence, potential codon changes, and oligonucleotides required for cloning into the DR274 *in vitro* transcription vector for use with cytosine (a) and adenine base editors (b). All sgRNAs designed and tested were cloned with 5' substitution. A list of genomic targets is shown below, including a list of potential DNA off-target sequences (OTs) provided with sequence coordinates, mismatches (shown in red), gene name and gene ID.

Figure 2 - Supplement 1

Figure 2-supplement 1. Somatic cytosine and adenine base editing at the medaka *oculocutaneous albinism II* (*oca2*) gene occurs in the absence of detectable indels and DNA off-target editing. (a) Translation of sections of exon 6 and 8 of human (Hs), medaka (Ol) and zebrafish (Dr) OCA2 shows high degree of conservation. Amino acids edited or targeted in this study are highlighted in yellow. Missense mutations R305W (green) and A334V (pink) highlight albinism-associated human OCA2 mutations (Kerr et al., 2000; Rebbeck et al., 2002). (b) Broader window of the Sanger sequencing result displayed in Figure 2e, indicates lack of indels or unwanted editing (Kluesner et al., 2018) (continued on next page).

Figure 2-supplement 1. (c) Microinjection of *oca2* sgRNAs with respective editors at the 1-cell stage causes low levels of aberrant phenotypes in medaka. Abnormal embryos either showed developmental delay or sublethal phenotype, at rates of 6.5%, 6.1%, 17.9%, 10.1%, 3.7%, 9.4%, and 15.8% (left to right), respectively. **(d)** The top three DNA off-targets (OTs) as predicted by ACEofBASEs for evoBE4max and ABE8e were evaluated for off-target base editing.

Figure 2-supplement 2. ABE8e efficiently introduces A-to-G mediated missense mutations in medaka and zebrafish. (a-b) EditR quantification of Sanger sequencing reads for pooled *oca2* ABE8e experiments (*Q256*, *T306*, *Q333*) from pools of five medaka embryos per data point, summarizes editing efficiencies for *Q256* (a) and *T306* (b) loci. For mean values see Table 2. To highlight the dinucleotide context, the nucleotide preceding the target A is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective A. **(c)** Wider window of exemplary Sanger sequencing reads for individual, base edited embryos, from each experimental condition following ABE8e experiments at the zebrafish *oca2*-L293 locus.

Figure 4 - Supplement 1

Figure 4-supplement 1. Cytosine base editing enables efficient installation of PTCs or missense mutations in two additional cardiac genes. Summary of C-to-T conversion efficiencies relative to the target C protospacer position for (Sample number used for evoBE4max: n=5, 6, 6, respectively for (b-d); and BE4-Gam: n=2 (b). To highlight the dinucleotide context, the nucleotide preceding the target C is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective C.

Figure 5 - Supplement 1

Figure 5-supplement 1. Robustness of interrogating gene function by introducing stop-gain mutations in medaka demonstrated at the N-terminal *kcnh6a*-Q11 locus in F0 through CBEs. (a) Cartoon of ERG structure according to Wang and MacKinnon (2017), highlighting the S4-6 domains. (b) Summary of mutations in the S4 domain in model systems studying ERG function in human (*Hs* ERG), medaka (*O/*ERG) and zebrafish (*Dr* ERG) (Hassel et al., 2008; Nakajima et al., 1999; Ng et al., 2012; Sanguinetti and Xu, 1999; Zhang et al., 2011). (c) Workflow. Control injections only contained the sgRNA. (d) Phenotypes observed following editing of *kcnh6a* were grouped by being either isolated “cardiac” phenotypes or with a “saturated” outcome of ventricular asystole concomitant with additional unspecific developmental defects (“global”). (e) Fraction of phenotype scores as a function of base editor generation. (f) Summary of editor type-specific C-to-T conversion efficiencies relative to the target C protospacer position for BE4-Gam (n=5), ancBE4max (n=4) and evoBE4max (n=7). To highlight the dinucleotide context, the nucleotide preceding the target C is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective C. (f') Example Sanger sequencing reads of single edited embryos with resulting missense and stop-gain mutations through editing at C3, C5, C7, and C8, respectively.

Figure 5 - Supplement 3

Figure 5-supplement 2. evoBE4max cytosine base editing enables efficient installation of PTCs or missense mutations at three additional *kcnh6a* loci. Summary of C-to-T conversion efficiencies relative to the target C protospacer position for (Sample number: n=6, 4, 6 respectively (b-d). To highlight the dinucleotide context, the nucleotide preceding the target C is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective C.

Figure 5 - Supplement 4

Figure 5-supplement 3. Analysis of BE4-Gam mediated PTC installation in *kcnh6a* in F1. (a) Injected embryos (*BE4-Gam* mRNA, pooled sgRNAs *kcnh6a*-Q11 and -R509) were raised to adulthood. Identified founders were crossed and F1 embryos further examined. (b-c) Analysis of F1 embryos at 6 dpf (b) and 4.5 dpf (c) from two independent founder couples.

Figure 5 - Supplement 4

Figure 5-supplement 4. Targeting the three *kcnh6a* loci simultaneously or *kcnh6a*-R512 alone in a different genetic background (*myl7::EGFP*, *myl7::H2A-mCherry*, *HdrR* strain) recapitulates phenotypic proportions.

Figure 6 - Supplement 1

Figure 6-supplement 1. CVD-associated SNVs can be mapped to the orthologous medaka peptide sequence with high conservation and are expressed during heart development. (a) Human (Hs) coding sequences for DAPK3, UBE2B, USP44 and PTPN11 around the missense mutations reported (Homsy et al., 2015; Zaidi et al., 2013), were obtained from the Ensemble genome browser and the orthologous medaka (OI) sequences were mapped to these (release 103). (b) Gene expression data for medaka *dapk3*, *ube2b*, *usp44* and *ptpn11* at embryonic stages was extracted from Li et al. (2020) and plotted to show presence of transcript during heart development and the contribution of maternal transcripts, which dominate before zygotic genome activation at around stages 7-8 (32 to 64-cell stages; Kraeussling et al., 2011).

Figure 6 - Supplement 2

Figure 6-supplement 2. Cytosine base editing allows the introduction of human CVD-associated missense mutations in medaka in F0 (continued on next page).

Figure 6-supplement 2. (a-d) Following microinjection at the 1-cell stage of evoBE4max together with the respective sgRNAs, Sanger sequencing results were quantified by EditR (Kluesner et al., 2018). Results are plotted for the genomic loci of *dapk3-P204* (n=7), *ube2b-R8* (n=5), *usp44-E68* (n=11), and *ptpn11-G504* (n=11) and are depicted for the underlying phenotype (square: normal; circle: cardiac) according to the phenotypes grouped in supplement 3. The standard base editing window is depicted in blue. To highlight the dinucleotide context, the nucleotide preceding the target C is shown by red (A), green (T), blue (C) and yellow (G) squares below the respective C. (a'-d') Sanger sequencing reads within the region of edited amino acids for the respective sgRNA experiment and locus. (b'-d') sgRNAs mediating the C-to-T transition map to the complementary strand. Note: nucleotide position 27 shows a G/A SNP in the Cab genetic background used in this study for *ube2b*.

Figure 6 - Supplement 3

Figure 6-supplement 3. Phenotypic categorisation of cytosine base edited embryos in medaka in F0. (a) Microinjection of evoBE4max with the respective sgRNA for *dapk3* (n=74), *ube2b* (n=91), *usp44* (n=79), and *ptpn11* (n=85) results in an increase in cardiovascular-associated phenotypes (“cardiac”) compared to *oca2* control editing (n=62). (b) Exemplary phenotypes of 4 dpf embryos for the categories given in a) (ventral view) with highlighting of the ventricle (V) and atrium (A). Here, embryos with morphological abnormalities or looping defects are shown in the “cardiac” category. Quantification of base editing efficiencies (supplement 2) for the individual embryos is shown for the *oca2* control (here: sequencing results for the four candidate loci given as range) and the “cardiac” examples of the respective experiment. Scale bar = 100 μ m. (c) Sub-categorisation of the “cardiac” and “global” phenotypic groups shown as percentage of the sum of the two categories. dpf=days post fertilization.

Figure 6 - Supplement 4

Figure 6-supplement 4. Confocal microscopy of evoBE4max validated CVD genes (continued on next page).

Figure 6-supplement 4. The full complement of analysed CVD associated phenotypes in cytosine base editing of *myl7::GFP*, *myl7::H2A-mCherry* reporter embryos at 7 dpf. Note the mirroring effect of the confocal microscope used, placing the ventricle to the right, and the atrium to the left in the control. Quantified C-to-T transversions are shown for each imaged embryo for the target codon. Scale bar = 100 μ m. dpf=days post fertilization.

Tables

Table 1. Cytosine base editing efficiencies. Shows nucleotide position (of CDS) and corresponding amino acid with changes. Note: only cytosines on the protospacer with clear editing are shown.

Genomic locus targeted	Edited cytosine (Amino acid, position, change)	Protospacer position (with dinucleotide context)	Editing efficiency (Mean $\pm$ SD)		
			BE4-Gam	ancBE4max	evoBE4max
<i>O. latipes oca2-Q333</i>	C995 > T (T332I)	aC5	29.3 $\pm$ 6.7	86.4 $\pm$ 11.5	20.7 $\pm$ 0.6
	C996 > T (T332)	cC6	38.0 $\pm$ 7.9	98.2 $\pm$ 2.7	100 $\pm$ 0
	C997 > T (Q333*)	cC7	29.3 $\pm$ 7.4	93.8 $\pm$ 7.9	93.3 $\pm$ 9.8
	est. INDEL freq. (%)		0.0	0.0	0.0
<i>O. latipes tnnt2a-Q114</i>	C337 > T (R113C)	gC5	0 $\pm$ 0	24 $\pm$ 14.3	26.8 $\pm$ 13.2
	C339 > T (R113)	gC7	0 $\pm$ 0	14.5 $\pm$ 15.1	65.8 $\pm$ 20.6
	C340 > T (Q114*)	cC8	0 $\pm$ 0	27.8 $\pm$ 16.8	85.9 $\pm$ 23.5
	est. INDEL freq. (%)		0.0	0.0	2.7
<i>O. latipes kcnh6a-Q11</i>	C26 > T (A9V)	gC3	0.6 $\pm$ 1.3	3 $\pm$ 2	29 $\pm$ 20.3
	C28 > T (L10F)	gC5	17.8 $\pm$ 2.7	45.8 $\pm$ 15.8	95.3 $\pm$ 6.7
	C30 > T (L10)	tC7	90.4 $\pm$ 13.8	92.5 $\pm$ 5.8	93 $\pm$ 7.6
	C31 > T (Q11*)	cC8	61.0 $\pm$ 10.4	92.8 $\pm$ 6.4	86.9 $\pm$ 15.6
	est. INDEL freq. (%)		0.0	7.7	9.7
<i>O. latipes kcnh6a-R509</i>	C1523 > T (A508V)	gC6	-	-	92.7 $\pm$ 12.8
	C1524 > T (A508)	cC7	-	-	93.3 $\pm$ 11.3
	C1525 > T (R509*)	cC8	-	-	39.7 $\pm$ 23.9
	est. INDEL freq. (%)		-	-	0.0
<i>O. latipes kcnh6a-L511</i>	C1531 > T (L511F)	gC5	-	-	47.5 $\pm$ 12.8
	C1533 > T (L511)	tC7	-	-	75.8 $\pm$ 9.3
	est. INDEL freq. (%)		-	-	6.5
	C1563 > T (D521)	aC6	-	-	0 $\pm$ 0

<i>O. latipes</i> <i>kcnh6a-R522</i>	C1564 > T (R522C)	cC7	-	-	54.2 ± 17.0
	est. INDEL freq. (%)				0.0
<i>O. latipes</i> <i>tnnt2c-R112</i>	C332 > T (A111V)	gC4	26 ± 7.1	-	60.4 ± 10.4
	C334 > T (R112*)	aC6	2 ± 2.8	-	67.6 ± 10.1
	est. INDEL freq. (%)		0.0	-	2.6
<i>O. latipes</i> <i>s1pr2-R150</i>	C449 > T (R150H)	gC8	-	-	21.5 ± 6.9
	C447 > T (E149)	gC10	-	-	30.2 ± 10.5
	C445 > T (E149K)	tC12	-	-	38 ± 20.5
	est. INDEL freq. (%)		-	-	0.0
<i>O. latipes</i> <i>s1pr2-R167</i>	C498 > T (C166)	gC3	-	-	24.7 ± 16.5
	C499 > T (R167C)	cC4	-	-	37.7 ± 6.7
	C501 > T (R167)	gC6	-	-	88 ± 12.2
	est. INDEL freq. (%)		-	-	0.0

1349 *sgRNA on complementary strand

Table 2. Adenine base editing efficiencies for ABE8e. Shows nucleotide position (of CDS) and corresponding amino acid with changes. Note: only adenines on the protospacer with clear editing are shown.

Genomic locus targeted	Edited adenine (Amino acid, position, change)	Protospacer position (with dinucleotide context)	Editing efficiency (Mean $\pm$ SD)
<i>O. latipes oca2-Q256</i>	A767 > G (Q256R)	cA6	100 $\pm$ 0
	A768 > G (Q256)	aA7	80 $\pm$ 1.7
	est. INDEL freq. (%)		0.0
<i>O. latipes oca2-T306</i>	A918 > G (T306)	cA4	52.3 $\pm$ 2.3
	A919 > G (I307)	aA5	4.7 $\pm$ 1.2
	est. INDEL freq. (%)		0.0
<i>O. latipes oca2-Q333[#]</i>	A994 > G (T332A)	aA4	96.5 $\pm$ 7.6
	A998 > G (Q333R)	cA8	89.7 $\pm$ 4.4
	est. INDEL freq. (%)		0.0
<i>D. rerio oca2-L293[*]</i>	A878 > G (L293P)	cA5	52.4 $\pm$ 32.9
	est. INDEL freq. (%)		0.0
<i>GFP-C71[*]</i>	A199 > G (V69A)	cA4	61.2 $\pm$ 9.4
	A211 > G (C71R)	cA9	97.0 $\pm$ 4.4
	est. INDEL freq. (%)		7.4
<i>O. latipes kcnh6a-T507</i>	A1517 > G (K506)	aA4	0.1 $\pm$ 0.4
	A1519 > G (T507A)	gA6	81.6 $\pm$ 7.2
	A1521 > G (T507)	cA8	44.1 $\pm$ 6.2
	est. INDEL freq. (%)		0.0
<i>O. latipes kcnh6a-R512</i>	A1534 > G (R512G)	cA8	92.1 $\pm$ 9.2
	est. INDEL freq. (%)		5.8
<i>O. latipes kcnh6a-D521</i>	A1562 > G (D521G)	gA5	61.0 $\pm$ 12.5
	A1568 > G (Y523C)	tA11	19.6 $\pm$ 2.8
	est. INDEL freq. (%)		7.4

*sgRNA on complementary strand

[#]averaged over single *oca2-Q333* and pooled injections

Table 3. Cytosine base editing efficiencies at GWAS validation genes. Shows nucleotide position (of CDS) and corresponding amino acid with changes. Note: only cytosines on the protospacer with clear editing are shown.

Genomic locus targeted	Edited cytosine (Amino acid, position, change)	Protospacer position (with dinucleotide context)	evoBE4max editing efficiency (Mean $\pm$ SD)
<i>O. latipes</i> <i>dapk3-P204</i> (n=7)	C610 > T (P204L)	gC5	64.7 $\pm$ 16.3
	C611 > T (P204L)	cC6	86.9 $\pm$ 8.7
	C613 > T (L205F)	gC8	61.7 $\pm$ 14.1
	est. INDEL freq. (%)		2.0
<i>O. latipes</i> <i>ptpn11-G504*</i> (n=11)	C1510 > T (G504R/K)	cC5	87.4 $\pm$ 7.7
	C1511 > T (G504E/K)	cC6	88.0 $\pm$ 7.3
	est. INDEL freq. (%)		0.0
<i>O. latipes</i> <i>ube2b-R8*</i> (n=5)	C20 > T (R7K)	tC8	67.4 $\pm$ 16.8
	C23 > T (R8Q)	tC5	54.2 $\pm$ 24.8
	est. INDEL freq. (%)		0.0
<i>O. latipes</i> <i>usp44-E68*</i> (n=11)	C201 > T (M67I)	cC8	97.0 $\pm$ 5.0
	C202 > T (E68K)	tC7	99.7 $\pm$ 0.9
	est. INDEL freq. (%)		0.0

*sgRNA on complementary strand

Table 4. Estimated editing windows and dinucleotide preference affecting editing efficiencies. Comparison of literature estimates (*in vitro*) and *in vivo* metrics observed in this study.

Base editor	Highest average editing efficiency (site tested)	Editing window on protospacer: overall (peak) activity		Dinucleotide sequence preference	
		<i>In vitro</i> *	This study	<i>In vitro</i> *	This study
BE4-Gam	61.0 $\pm$ 10.4 (<i>kcnh6a-p.Q11X</i>)	3-10 (4-8)	4-8 (NA)	canonical	canonical
ancBE4max	93.8 $\pm$ 7.9% (<i>oca2-p.Q333X</i>)	3-9 (4-7)	NA (5-8)	<<GC	<<GC
evoBE4max	99.7 $\pm$ 0.9% (<i>usp44-p.E68K</i>)	1-11 (4-8)	3-12 (5-8)	<<AC	<<AC
ABE8e	100% (<i>oca2-p.Q256R</i>)	3-11 (4-8)	3-11 (4-8)	-	<AC

NA – not sufficient data to estimate

*(Arbab et al., 2020; Huang et al., 2021; Richter et al., 2020)

Video legends

Video 1. evoBE4max introduced premature STOP codon in *O. latipes tnnt2a* results in silent heart phenotype. Time-lapse movie (10 seconds) of the beating medaka heart. Scale bar = 400 μm .

Video 2. evoBE4max introduced premature STOP codon in *O. latipes kcnh6a* results ventricular dyastole accompanied by morphological alterations. Time-lapse movie (10 seconds) of the beating medaka heart. Scale bar = 400 μm .

Video 3. ABE8e driven installation of the R512G missense mutation in *O. latipes kcnh6a* results ventricular dyastole accompanied by morphological alterations. Time-lapse movie (10 seconds) of the beating medaka heart. Scale bar = 400 μm .