Boosting the host immune system to fight tuberculosis

Boland, R.

2

Deep learning image recognition enables efficient genome editing in zebrafish by automated injections

Maria Lorena Cordero-Maldonado[1*], Simon Perathoner[1*], Kees-Jan van der Kolk[2*], **Ralf Boland**[3], Ursula Heins-Marroquin[1], Herman P. Spaink[3], Annemarie H. Meijer[3], Alexander D. Crawford[1], Jan de Sonneville[2]

1. Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belvaux, Luxembourg
2. Life Science Methods BV, Leiden, the Netherlands
3. Institute of Biology, Leiden University, Leiden, the Netherlands

* Equal contribution

Adapted from Cordero-Maldonado *et al* (2018)
DOI: https://doi.org/10.1371/journal.pone.0202377

Abstract

One of the most popular techniques in zebrafish research is microinjection. This is a rapid and efficient way to genetically manipulate early developing embryos, and to introduce microbes, chemical compounds, nanoparticles or tracers at larval stages. Here we demonstrate the development of a machine learning software that allows for microinjection at a trained target site in zebrafish eggs at unprecedented speed. The software is based on the open-source deep-learning library Inception v3. In a first step, the software distinguishes wells containing embryos at one-cell stage from wells to be skipped with an accuracy of 93%. A second step was developed to pinpoint the injection site. Deep learning allows to predict this location on average within 42 µm to manually annotated sites. Using a Graphics Processing Unit (GPU), both steps together take less than 100 milliseconds. We first tested our system by injecting a morpholino into the middle of the yolk and found that the automated injection efficiency is as efficient as manual injection (~ 80%). Next, we tested both CRISPR/Cas9 and DNA construct injections into the zygote and obtained a comparable efficiency to that of an experienced experimentalist. Combined with a higher throughput, this results in a higher yield. Hence, the automated injection of CRISPR/Cas9 will allow high-throughput applications to knock out and knock in relevant genes to study their mechanisms or pathways of interest in diverse areas of biomedical research.

Introduction

Microinjection is one of the most powerful techniques used in zebrafish (*Danio rerio*), as it allows to follow cell fate[1], evaluate pathogenesis of bacteria[2], produce chimeric individuals[3], study tumour progression[4,5], manipulate protein levels[6,7] and create genetically altered lines[8]. In addition, it is also a suitable technique to introduce chemical compounds that otherwise do not readily enter the embryo due to the compound lipophilicity properties and the protection function of the chorion[9]. More recently, microinjections have been also used in the fields of toxicology and nanomedicine to evaluate nanoparticles toxicity at different functional levels and to inject nanoparticles encapsulating genetic material or therapeutic drugs to specific tissues in older embryos and/or larvae[10-13].

The intrinsic biological properties of zebrafish make it particularly amenable to this technique, since these cyprinids are highly fecund, a spawning pair typically producing more than 400 eggs at a time. Moreover, fertilization is external and spawning is confined to a brief period at dawn (natural or artificial), allowing for timing of the experiments. Furthermore, the chorion of zebrafish eggs is supple and easy to pierce.

Classically, injection of tracer dyes is used to identify single cell populations[14,15], to follow cell lineages and to build fate maps in zebrafish[1,16]. The development of molecular methods for the zebrafish model enabled functional studies by manipulating the expression of specific genes. Injection of messenger RNA (mRNA) can be used to overexpress and misexpress a specific protein[17], while morpholino antisense oligonucleotides (MOs) can be employed to knock down a given target gene[18]. In zebrafish mRNA and MO injections are simply performed by introducing a fine-tipped needle into the yolk of one-cell stage eggs and delivering nanoliter volumes of the injection material into it[19]. As cytoplasmic streaming will move the mRNA or MOs into the cytoplasm, it is not necessary that the injection targets the cell. While injection into the yolk requires

some skill, it can usually be learned within a few weeks. Nevertheless, injections of mRNAs and MOs have their drawbacks. First of all, the effect is only transient, *i.e.* the injected molecules will be degraded and/or diluted with time. Moreover, in the case of mRNA injection, tissue-specific upregulation is not possible and a given mRNA will be expressed in all tissues indiscriminately. Also, the specificity of MO antisense technology has recently been questioned as MOs can sometimes lead to misleading results due to toxicity and off-target effects[20]. In a recent study[21], loss-of-function mutations for ten different genes previously thought to have an essential role in development failed to recapitulate the corresponding morpholino-induced phenotypes. In several cases, the discrepancy between mutant and morphant phenotypes, could be explained by genetic compensation mechanisms that occur in mutants[22], however, undoubtedly rigorous controls are required to ascertain the reliability of MO-induced phenotypes[20,23,24].

In the last years, with the implementation of targeted nuclease techniques in the zebrafish, the demand for genetic evidence to define gene function has greatly increased. Fortunately, after a somewhat slow start using zinc-finger nucleases (ZFNs)[25] and transcription activator-like effector nucleases (TALENs)[26], the adaptation of the prokaryotic CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/ CRISPR associated protein 9) defence system to engineer genomes has revolutionized reverse genetics in zebrafish[27].

Recently the CRISPR/Cas9 system was adapted and optimised to engineer genomes. A single synthetic guide RNA (gRNA) directs Cas9-mediated cleavage of target DNA [27,28], and the method was implemented in multiple systems including zebrafish[29,30], finally paving the road for knock-ins in this model[31]. Along with *Tol2* mediated transgenesis, a transposon system based on the *Tol2* element of medaka (*Oryzias latipes*) widely used in zebrafish to create transgenic lines[8], the CRISPR/Cas9 system has become an essential tool for genome editing in zebrafish. In this context microinjection is an essential technique. For the creation of genetically altered lines in zebrafish, be it through *Tol2* transgenesis or by means of zinc finger nucleases, TALEN or Cas9 nucleases, it is critical to inject the solution directly into the blastomere at the one-cell stage or at least at the interface between blastomere and yolk[32-35]. Contrary to RNA or MOs, DNA appears not to be transported into the blastomere via cytoplasmic streaming. Moreover, efficiency of all these genome editing techniques is much lower compared to mRNA or MOs injections. Therefore, in order to create genetically altered zebrafish lines it is essential to master microinjections into the cell. This can be challenging as this type of injection requires long training and excellent technical skills

Automated microinjection system

One of the first reports attempting to establish an automated microinjection system was published by Wang and colleagues in 2007[36]. This microrobotic system based on computer vision and motion control was able to inject zebrafish embryos at an average speed of 25 seconds per embryo. Although quite innovative, this system is limited by the low batch size (only up to 24 embryos per plate) and low injection speed compared to the first version of our microinjection system[33,37,38]. This automated microinjection system featured half-spherical wells, moulded in agarose gel, which allowed for high-throughput microinjection into the yolk of zebrafish eggs at fast speed (1 embryo in 1.8 seconds). This was used for microinjection of bacteria, morpholinos and cancer cells[37,38]. The great advantage over other systems was the higher batch size and speed of the injections allowing to inject up to 2000 embryos per hour and up to 2580 embryos per plate. As the initial cell division steps in zebrafish embryos occur in intervals of 20-40 minutes, speed is crucial for the accuracy, reproducibility and number of experiments.

In our experience, it is apparent that injections into the middle of the yolk are less suitable for DNA injections. Therefore as a first step, the program "click-to-inject" was developed to test the efficiency of injections closer to the first cell[38]. With this, we noticed that we could achieve a great increase in efficiency, similar to manual injections done into the first cell. Therefore, we set out to automate this procedure.

In this study we demonstrate the results of autonomous site selection and injection for CRISPR/Cas9 and DNA manipulation of the zebrafish genome.

Results and discussion

Manual and automated injections of slc45a2-MO

In order to test MO efficiency of manual and automated injections we employed a translation-blocking MO against *slc45a2* (solute carrier family 45 member 2). Downregulation of this gene induces albino and/or hypo-pigmented morphants, as the melanophores are unable to produce melanin[41]. Manual and automated yolk microinjections were performed in parallel, and in both cases the induced albino phenotype was assessed in larvae at 3 dpf (Figure 1). The results obtained with both microinjection approaches are comparable and show that downregulation of *slc45a2* is highly efficient using morpholino antisense technology (Figure 1A). Additionally, the manual injections were performed by two different experimentalists (Figure 1B) and this shows that efficiency and variation of efficiency obtained by manual morpholino injections differs from person to person and that the variation of the efficiency of the automated injections is slightly larger.

Semi-automated "click-to-inject"

After demonstrating that automated injection into the yolk is an efficient way to generate morphants, we sought to apply the robotic injector for generating CRISPR/

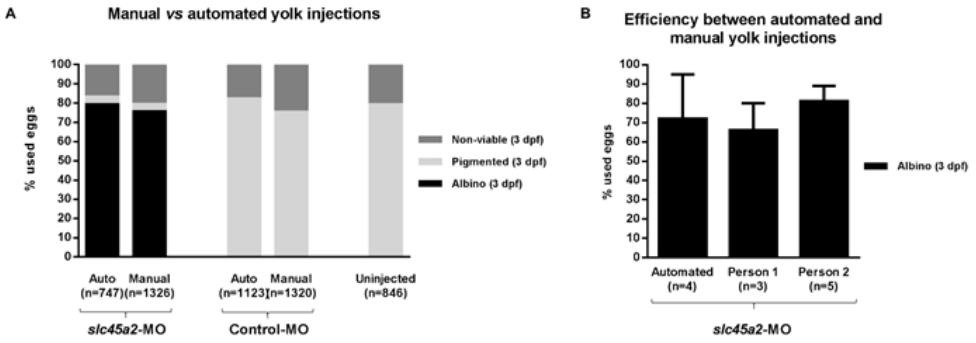

Figure 1. Morpholino knockdown efficiency with manual and automated injections.

A. The survival and knockdown efficiency of slc45a2-MO manual and automated (auto) microinjections were measured as the number of larvae displaying an albino phenotype at 3 days post-fertilization (dpf). Control-MO injected larvae and uninjected larvae were processed in parallel and the resulting pigmented (wild-type) larvae were also counted at 3 dpf. "n=" indicates the number of eggs used to obtain this cumulative result.

B. Efficiency comparison between the automated injection into the yolk and manual injections performed by two independent experimentalists (P1: experienced and P2: expert; not statistically significant). "n=" indicates the number of experiments used to calculate the average and standard deviation. Each experiment refers to different technical and biological experiments.

Cas9 mutants for *slc45a2*. To investigate the dependence on the injection location we used the "click-to-inject" program to test the efficiency of injections closer to the first cell. In the "click-to-inject" program the injection depth is set, but the (x y) position is chosen by the operator[38]. To inject, the operator moves the mouse pointer to a specific site (*e.g.* the first cell) and clicks to trigger an injection and a subsequent movement to the next egg. Based on this, next we set out to develop an automated image recognition to more precisely identify the first cell and to automate CRISPR/Cas9 injections.

Imaging conditions

In manual microinjection setups, as well as in standard microscopy, near-perfect imaging conditions are applied with lighting from the bottom and imaging from the top, or vice versa. As the zebrafish egg is very transparent, epi illumination from below is not suitable; most contrast and edges are then lost. As the egg is spherical, a ring-light displays a very bright circle on top of the egg. Therefore, to obtain better and more reproducible imaging conditions in different locations, we placed a large (L x B = 60 x 80 cm^2) diffuse light source above the robotic injector. Five different classes were used to annotate the images (Figure 2). In the "Inject" class the ideal injection position for automated microinjection is also annotated. Instead of injecting directly into the zygote, we have chosen to inject in the yolk, close to the visible zygote. The reason is that injections in a thin zygote (less ideal orientation, or very early stage) would often cause a rotation of the egg, and bounce the needle off. Injections into the yolk-blastomere boundary almost never show this problem, and thus gave a higher yield.

Machine learning

Initially, we tried a classic approach of machine vision on these images. The Hough

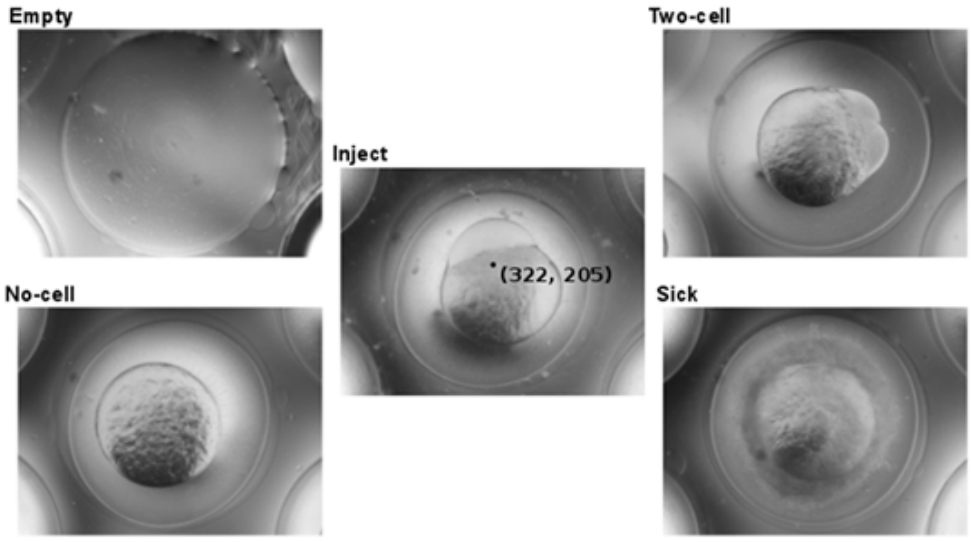

Figure 2. Imaging classification for injection.
Representative digital images measured from below of an agarose grid ("Empty") that supports zebrafish eggs with the first cell visible ("Inject") or not visible ("No cell"), eggs in a two- or higher cell stage ("Two cell") or non-viable eggs ("Sick"). In the "Inject" image an injection location is indicated by a black dot with (x y) coordinates.

circle transform allowed us to detect the yolk with an above 90% accuracy[36]. However, the next step to find the first cell was problematic. In cases where the shadow of the micromanipulator overlapped with the first cell, the edge detection algorithm failed. As an alternative to edge detection, we annotated a database of images with injection positions. We used a Fast Fourier Transform (FFT) algorithm to find a closest matching egg in this database and used that image to infer an injection position. This worked reasonably well with a peak error (distance between calculated position and annotated position) of 20 µm. However, when a good match could not be found, the error was quite high, and as a result the tail of the error was quite large. An explanation for the large variation in results is that there is also a large variation in first cell shapes, especially when looking from an arbitrary angle. It can be an early very thin line up to about a third of the yolk depending on the developmental stage and orientation of the egg. To overcome this variation, we could make the annotated database larger, to increase the chance of a close match. Nevertheless, the downside of this solution is that more images have to be compared, and this takes more processing time during injection. Thus, we sought to apply a better approach based on deep learning.

Using a database of annotated images as input, one can also train a deep learning network. Instead of comparing images during runtime, one trains an algorithm that is afterwards used to interpret new images. The execution time of this algorithm is independent of the size of the training image set. Thus, roughly speaking, the larger the number of annotated images, the higher the accuracy of the algorithm. We used the Inception v3 open source deep learning software[46]. This software has been built and tested to categorize images, based on a large training database of images, initially for the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC; www.image-net.org). The Inception v3 architecture uses a neural network that takes the pixels of images as input and extracts features. Many features are subsequently built on top of features, in different layers of neurons, in higher and higher levels of abstraction, until the neurons reach an output of defined categories[47]. One advantage of the Inception v3 software is that one can reuse the first layers of feature extraction for a different set of images. This is built on the idea that the basic features, *e.g.* lines and simple patterns, can be used in all higher-level features that are used to train new categories with new sets of images. Within eight hours of training time we reached a 93% accuracy, with an execution time in the order of tens of milliseconds.

After finding the images with a visible first cell, the next step was to determine the injection location. To enable the use of deep learning for this problem, we had to modify the output from categories into an ideal location. When just the pixel (x y) coordinate is used as output, only one pixel of the whole image is correct. With this output the neuronal network cannot easily distinguish between locations closer to the annotated location and further away, and this makes learning impossible. Therefore, we translated the (x y) coordinates to a barycentric coordinate system[48]. The Greek word "*barys*" means heavy and refers to the centre of gravity. In a barycentric coordinate system a grid of triangles is used, with a weight assigned to each vertex. This is used as follows. A chosen grid of triangles is placed on top of each image. The annotated injection position will fall within one triangle; then the weights of these triangle vertices are given a value according to the location within that triangle. These weights sum up to one, whereas the other vertices in the grid are all zero. This vertices output vector then represents the ideal outcome. The advantage is now that a small deviation from this ideal output vector can be scored gradually instead of binary. This then allows for efficient training. A more detailed explanation is available in S1 Text. After eight hours of training we created a table of (x y) coordinates using validation images. We calculated the distance between

the annotated injection position and the position as predicted by the deep learning network (Figure 3). The average distance is 42 µm, as depicted in Figure 3B, and for 83% of the images this distance is smaller than 60 µm.

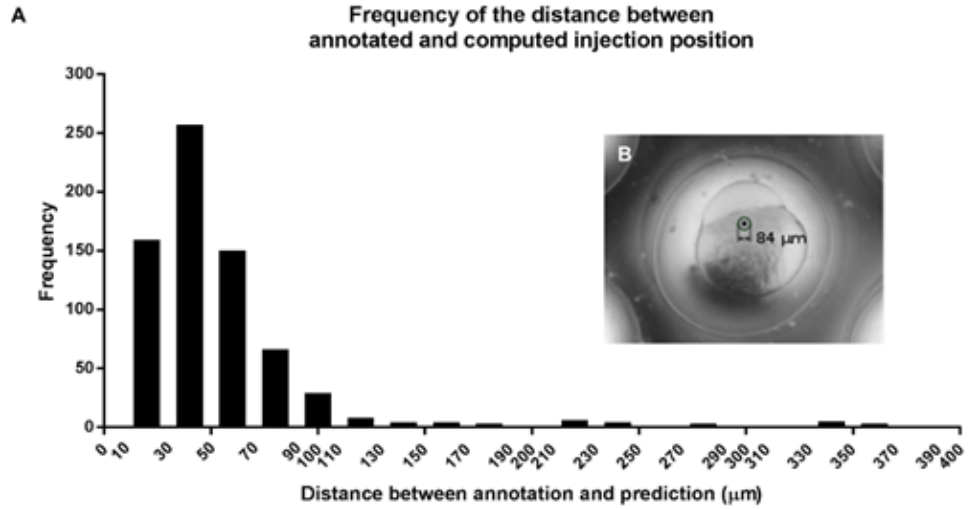

Figure 3. Distance between annotated and computed injection location.

A. Bar graph depicting the frequency of the distance between annotated and computed injection position (prediction).

B. Digital image with a circle around an annotated injection point to illustrate the average distance between annotation and prediction.

Automated injection of slc45a2 gRNA/Cas9

Trial and error in many laboratories have led to a best practice of injecting into the first cell for the application of the CRISPR/Cas9 editing technique. In our robotic microinjection system, injecting in the middle of the yolk gives the highest speed. Image recognition used to customise an injection location takes time but can increase the injection efficiency. To balance efficiency and speed, and to be able to monitor improvements of our image recognition model, we started by measuring efficiency of CRISPR/Cas9 injections performed in the yolk. Both manual and automated yolk injections gave a very low efficiency of 12% (Figure 4A). Then, with the "click-to-inject" program (semi-auto), resulting in injections closer to the first cell we could generate albino larvae at an almost three times higher efficiency than with the injections in the middle of the yolk (Figure 4A). Next, using deep learning (auto), we could automate this procedure and with this we reached a slightly lower efficiency when comparing it the "click-to-inject" injections but a higher efficiency than the one obtained with automated and manual injections in the middle of the yolk (Figure 4A). Still, manual injections into the first cell reached the highest efficiency of 43% (Figure 4A). Figure 4B shows that both the efficiency and the variation between experiments differs considerably depending on the experimentalist (displayed as P1: experienced, P2: expert and P3: novice). In contrast, here, the automated injections show relatively little variation, also when compare them to the "click-to-inject" (semi-auto) injections (Figure 4B). Also, it can be seen that the efficiency is quickly surpassed by humans given enough experience (P1 and P2). This lower efficiency achieved with the robot can be explained by the injection location – close-to-cell instead of into the zygote – and by the fact that not all the eggs

are oriented with a cell visible on the side, despite the fact that they are oriented in the agar grid. Hence, the automated injections tend to be a mixture of injections into the middle of the yolk, and close to the first cell, when the first cell was detected. With this we obtained an efficiency of 24% on average (Figure 4B).

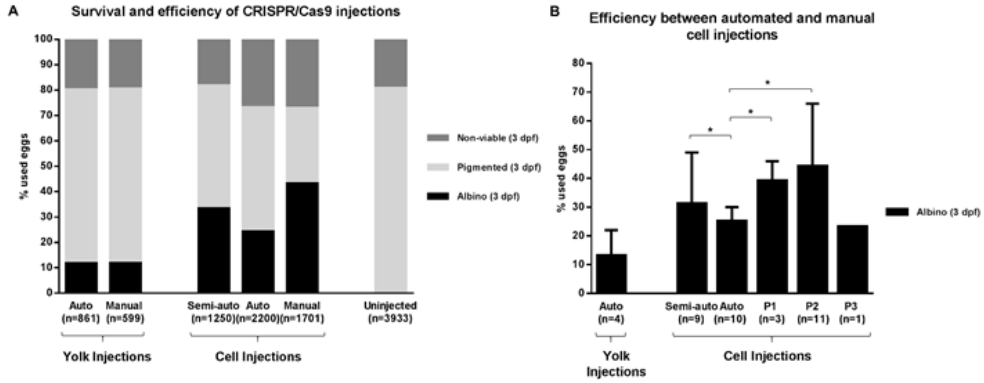

Figure 4. Automated injections of CRISPR/Cas9.

A. Survival and average efficiency of slc45a2 gRNA/Cas9 manual, click-to-inject (semi-auto) and automated (auto) microinjections both in the yolk and in the cell were measured as the number of larvae displaying an albino phenotype at 3 days post-fertilization (dpf). Uninjected larvae were processed as controls and the resulting pigmented (wild-type) larvae were also counted at 3 dpf. "n=" indicates the number of eggs that were used to obtain the cumulative results.

B. Comparison of the average efficiency and standard deviation between the automated (auto), click-to-inject (semi-auto) and manual injections performed by three independent experimentalists (P1: experienced, P2: expert and P3: novice). "n=" indicates the number of experiments that were used to calculate the average and standard deviation. Each experiment refers to different technical and biological experiments. (* $p<0.05$).

Automated injection of DNA

For the injections with DNA we used a COPAS (Complex Object Parameter Analyzer and Sorter) system to measure the efficiency of the injections (Figure 5A). For this, we first measured the highest red fluorescence signal of the uninjected control larvae and took the highest signal as a threshold at 5 dpf. Then we measured the DNA-injected larvae and counted the larvae that passed this threshold. The survival was measured at 1 dpf to focus on differences as a result of the injection. Prior to placing the larvae into the COPAS system, larvae with visible developmental defects were removed. Both the manual and automated injected eggs had a similar relative number of malformed embryos (4% on average).

These results show that DNA injections are less demanding in terms of injection location. Injections into the middle of the yolk reached an average efficiency of 32%. This can be improved by injecting close to the first cell, when possible, to reach an efficiency of 39%. Surprisingly, here manual injections close to the first cell (personal preference) had a lower efficiency than could be obtained by automated injections and gave on average the same efficiency as injections into the middle of the yolk.

Microinjection throughput

To calculate the microinjection throughput, we divided the average injection time by the average efficiency. This results in the average time needed for one successfully injected larva. We measured and compared the throughput for the different genetic

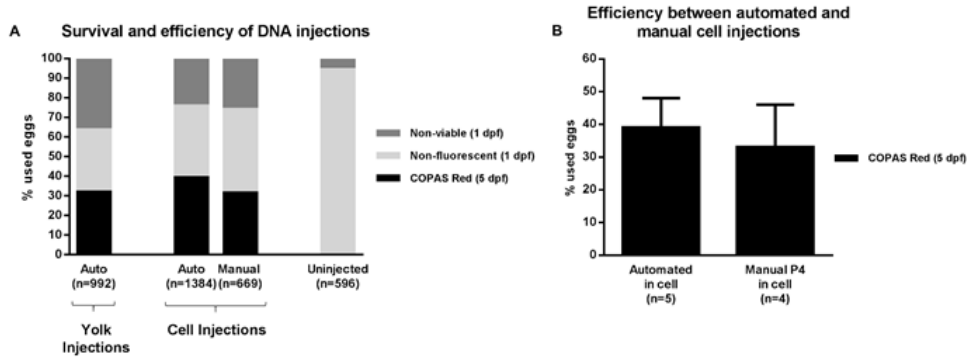

Figure 5. Automated injections of DNA.

A. Average survival and efficiency of DNA automated (auto) and manual injections as measured by the COPAS system. "n=" indicates the number of eggs that were used to obtain the cumulative results.

B. Comparison of the average efficiency and standard deviation between the automated and manual cell injections. P4 indicates a different experienced experimentalist and "n=" indicates the number of experiments that were used to calculate the average and standard deviation.

modifications and experimental setups described in this article, *i.e.* automated and manual injections for gene knockdown by morpholino antisense, gene knockout by CRISPR/Cas9 and transgenesis by *Tol2* (Figure 6).

In the case of the manual injections, the throughput differs greatly depending on the experimentalist, as experience can lead to a higher throughput by increasing both the efficiency and speed of the injection process. It can also be seen that the robot is on par with fast human performance in case of the morpholino injections, but 1.5 times as fast as average human performance.

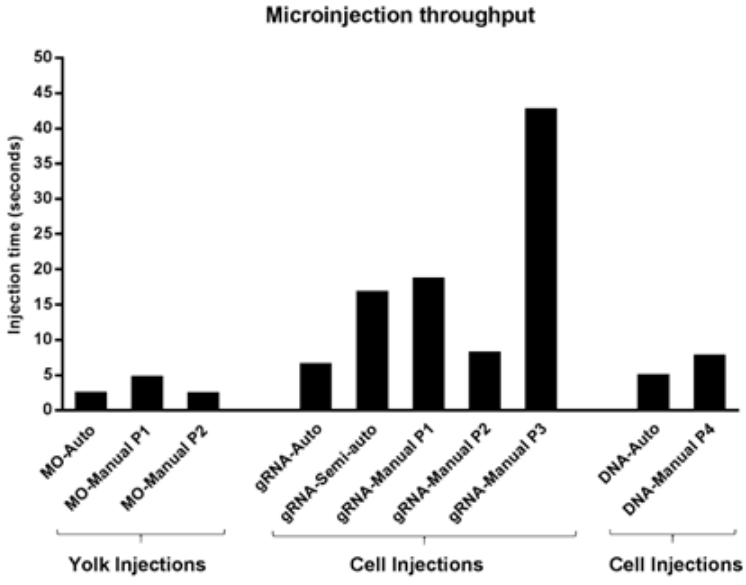

Figure 6. Average injection time required to obtain one positive genetically modified larva.

Abbreviations: MO, slc45a2 morpholino; gRNA, slc45a2 gRNA/Cas9; DNA, Tol2 construct; Auto, automated injections; P1-4, four different experimentalists.

With deep learning, a robot can outperform humans on the more complex cell injections. With CRISPR/Cas9 samples the robot needs 6 seconds of injection time to obtain a positive larva, and humans need 8 up to 43 seconds. On average, the robot is more than three times (3.6x) faster. Manual injections of DNA constructs close to the cell are faster to perform than injections into the cell (2.5 seconds *vs* 6.8 seconds). However, this also reduces the manual efficiency, resulting in a 1.5 times higher throughput of the robot. A movie showing the robotic injection process in real-time is available in S1 File. The movie shows that the time between capturing the image and placing the cross (demonstrating the calculated injection location) is only about 100 milliseconds.

Efficiency dependence on the injection location

Contrary to what might be expected, the efficiency of injections into the middle of the yolk to alter the genome were not negligible as the efficiency was 12% for CRISPR/Cas9, 32% for DNA injections and 80% for morpholino injections. Using the measured efficiencies and statistics of image classification we can calculate the efficiencies of injections close to the first cell. During the injections of CRISPR/Cas9, on average 65% of the eggs were oriented with a first cell visible, and 35% were injected into the middle of the yolk. The increase in efficiency, 24%, was caused by 65% of the eggs being injected with efficiency much higher than 12%. Using the efficiency of the yolk injections we can predict the efficiency of injections close to the first cell. Solving the equation 0.65*X+0.35*0.12 = 0.24 for X results in an efficiency of around 30% for injections close to the first cell. For DNA injections we have chosen to not orient the eggs after placing them in a grid, and therefore less eggs, 46%, were injected close to the first cell. Solving the equation 0.46*X + 0.54*0.32 = 0.39 for X results in a predicted efficiency of 47% for injections close to the first cell. Surprisingly, this is much higher than what was obtained by manual injections close to the first cell. The higher efficiency of DNA injections (47%) versus CRISPR/Cas9 injections (24%) can be partially explained by the fact that the integration of a single copy of DNA construct can still be detected, while the readout of the CRISPR/Cas9 injection requires a non-synonymous mutation to occur on both alleles in order to have a visible phenotype, the albino phenotype being recessive. A non-synonymous mutation in one copy of *slc45a2* would not be detected in our assay. Hence, we expect that the actual number of induced CRISPR/Cas9 mutations is underestimated.

These measured and calculated efficiencies can also be used to make a prediction of the positive embryos, directly after the injection.

Conclusion and perspectives

In this study we have demonstrated how we improved an automated injection robot to inject close to the first cell using image recognition in order to enable efficient genome editing in zebrafish embryos. This was accomplished using a modified open-source deep-learning software and annotation of thousands of images. A step-by-step approach of first testing an annotation strategy and efficiency helped to predict the increase in efficiency that can be obtained. Initially we tested the efficiency with a semi-automated click-to-inject program. This click-to-inject approach is also suitable as a first step for other microinjection applications, such as injections into older zebrafish larvae or different organisms.

Because of its transparency, rapid development and easy genetic manipulation, zebrafish

have become a key vertebrate model organism for the elucidation of developmental processes. With the advent of CRISPR/Cas9 technology, zebrafish are becoming an even more powerful tool for the study of diverse human disorders. The CRISPR/Cas9 system achieves mutagenesis rate of around 80% for generation of knockout lines[31], and has proven to have fewer side effects than other genome editing technologies. However, generation of specific hereditable mutations or epitope tagging of chromosomal genes in zebrafish is still challenging. Unfortunately, genome editing in zebrafish is unpredictable and efficiency sometimes drops to 3.5%[49]. Therefore, higher number of eggs should be injected for the generation of the expected mutation. Creation of zebrafish mutant lines using CRISPR/Cas9 requires precise injections into or close to the zygote. These types of injections take time to master and are tedious if many batches of hundreds of eggs have to be injected, particularly for the generation of knock-in lines. Our results have shown that efficiency and reproducibility of manual cell injections highly depend on the training stage of the person performing the experiment, making it more difficult to have this technique as a routine procedure in the laboratory. Here, we show the establishment of automated injections as a reliable tool for the generation of CRISPR/Cas9 mutants. Automated microinjections are simple to learn and allow the cell injection of 100 embryos in 2.5 minutes with comparable efficiency to manual cell injections. This method could also be used for high-throughput gene overexpression studies by microinjection of mRNA.

The need for high-throughput genome manipulation

To date there have been almost 9,000 morpholinos used in zebrafish research. In addition, the adaption of CRISPR/Cas9 editing technology is progressing faster than any other gene silencing method, and even faster than the adoption of morpholino knockdown technology (statistics on zfin.org). However, injections of mRNAs or DNA must be more precise and are more time consuming. Therefore injection can be a limiting step for high-throughput genetic studies. For the moment, there are about 30,000 known gene loci that could be interesting to manipulate in order to investigate their function in development, disease or expressed phenotype (zfin.org). Multiplied with 300 injections that are typically used to obtain a mutant, and multiple mutants per gene, this brings us to tens of millions of injections. Much time and efforts would be saved if this tedious but needed task can be performed mostly by robotic systems.

Acknowledgements

We would like to thank Dr. Uwe Irion for sharing materials and for his valuable input. We are also grateful to the fish caretaker teams from both Leiden University and the Luxembourg Centre for Systems Biomedicine for their daily valuable work.

Materials and methods

Animals

Wild type adult zebrafish (AB or TL strain) are maintained in the Aquatic Facility of the Luxembourg Centre for Systems Biomedicine and the Institute of Biology, Leiden University, according to standard protocols[39]. Zebrafish eggs were obtained by natural spawning on the day of each experiment, kept in 0.3X Danieau's solution (14 mM NaCl, 2 mM KCl, 0.12 mM $MgSO_4$, 1.8 mM $Ca(NO_3)_2$, 1.5 mM HEPES pH 7.5 and 0.03 M methylene

blue) or egg water (60 µg/ml sea salt, Sera Marin, Heinsberg, Germany), and staged by morphology (one-cell stage) for the injections. After each series of injections, the eggs were incubated at 28 °C (±0.5) and evaluated up to 5 days post-fertilization (dpf). Anaesthesia of larvae used for live imaging and COPAS analysis was done with 0.02% buffered Tricaine (3-aminobenzoic acid ethyl ester, Sigma-Aldrich) in egg water[37,40].

Ethics statement

The Luxembourg Centre for Systems Biomedicine is registered as an authorized breeder, supplier and user of zebrafish (*Danio rerio*) with Grand-Ducal decree of 20 January 2016. Zebrafish lines used at the Institute of Biology, Leiden University were handled in compliance with local animal welfare regulations as overseen by the Animal Welfare Body of Leiden University (license number: 10612). All practices involving zebrafish were performed in accordance with European laws, guidelines and policies for animal experimentation, housing and care (European Directive 2010/63/EU on the protection of animals used for scientific purposes). The present study did not involve any procedures within the meaning of Article 3 of Directive 2010/63/EU and as such it is not subject to authorization by an ethics committee.

Morpholino antisense oligonucleotide

The translation blocking morpholino for *slc45a2* (solute carrier family 45 member 2) was obtained from Gene Tools according to Dooley *et al*., 2012 with the following sequence: 5'-GCTGGTCCTCAGTAAGAAGAGTCAT-3' [41]. In addition, a 3' fluorescein modification was included, which allowed fluorescent differentiation of injected eggs. A standard MO with sequence 5'-CCTCTTACCTCAGTTACAATTTATA-3' was used as an injection control. In both cases, stock solutions (1 mM ~ 8 ng/nL) were prepared according to the specifications of the provider and titrated working solutions were freshly prepared for each experiment.

Preparation of Cas9 mRNA, slc45a2 sgRNA and DNA construct

Both *slc45a2* sgRNA and Cas9 mRNA were prepared according to Gagnon *et al*., 2014[34]. Briefly, the *slc45a2* DNA template was synthetized with T4 DNA polymerase (New England BioLabs) using the oligonucleotides: *slc45a2*-specific (taa tac gac tca cta taG GTT TGG GAA CCG GTC TGA Tgt ttt aga gct aga aat agc aag) and constant (AAA AGC ACC GAC TCG GTG CCA CTT TTT CAA GTT GAT AAC GGA CTA GCCTTA TTT TAA CTT GCT ATT TCT AGC TCT AAA AC). The sgRNA was synthetized using T7 RNA polymerase (Ambion MEGAscript) and then diluted to 400 ng/µl. Cas9 mRNA was synthetized using the pCS2-Cas9 plasmid[42], transcribed using the SP6 mMessage mMachine kit (Ambion) and finally diluted to 600 ng/µl.

The DNA plasmid was constructed using standard methods[43]. Briefly, a GFP reporter (*Tol*2kit construct 389) and mCherry reporter (*Tol2*kit construct 233) expressed under a constitutive promoter (*Tol2*kit construct 299) was constructed (final construct actb:-NLSmCherry-IRES-GFP) in the Gateway *Tol2* vector (pDestTol2pA2). The plasmid was transfected in *E. coli,* isolated from an overnight liquid culture and diluted to 25 ng/µl. The *Tol2* transposase RNA was synthetized using SP6 RNA polymerase (Ambion mMESSAGE mMACHINE) and then diluted to 25ng/µl.

Manual microinjections

Manual microinjections of *slc45a2*-MO, *slc45a2* sgRNA/Cas9 and the DNA construct in zebrafish embryos were performed following standard methods using Eppendorf FemtoJet microinjectors and both in-house pulled needles, prepared with thin-wall capillaries (World Precision Instruments) and a P-1000 Micropipette Puller (Sutter

Instrument, USA), and commercially available ready-to-use 10 µm tip needles (Qvotek, Mississauga, Canada)[19,44]. The needles for each round of injections were calibrated according to well-established methods using a stage micrometer slide (Carl Zeiss)[19]. Ultimately, the required bolus size for injection was achieved and controlled by regulating the pressure in the microinjector. After each series of microinjections, the embryos were incubated at 28 °C (±0.5) and evaluated daily until 5 dpf to record non-viable embryos, (*i.e.* non-fertilized, fluorescent negative, dead and dysmorphic embryos/larvae from 1 to 5 dpf), and the efficiency of injection. Table 1 shows the specifications of the different types of manual injections that were performed in this study.

	***slc45a2* MO & control MO**	***slc45a2* gRNA + Cas9 RNA**	**DNA construct + *Tol2* RNA**
Injection type	Manual	Manual	Manual
Developmental stage at injection	One- to two-cell	One-cell	One-cell
Injection location	Middle of the yolk	Middle of the yolk	Blastomere/yolk boundary
Sample concentration	230 µM	400 ng/µL + 600 ng/µL	25 ng/µL + 25 ng/µL
Injection volume	2 nL	4 nL	1 nL
Injection time(per 100 embryos)	P1: 5 min P2: 3 min	P1: 17 min P2: 10 min P3: 19 min	P4: 4 min
Microinjector type	EppendorfFemtoJet 4X	EppendorfFemtoJet 4X	Eppendorf FemtoJet
Evaluation lapse	From 6 hpf to 5 dpf	From 6 hpf to 5 dpf	From 6 hpf to 5 dpf
Sorting criteria	GFP positive and albino phenotype	Albino phenotype	GFP positive
Exclusion criteria	Non-fertilized, GFP negative, dysmorphic and dead embryos / larvae	Non-fertilized, no albino phenotype, dysmorphic and dead embryos / larvae	Non-fertilized, GFP negative, dysmorphic and dead embryos / larvae
Place of the experiments	LCSB	LCSB	Leiden University

Table 1. Technical specification for the manual microinjections
Abbreviations: P1, experienced experimentalist; P2, expert experimentalist; P3, novice experimentalist; P4, experienced experimentalist; hpf, hours post-fertilization; dpf, days post-fertilization; GFP, green fluorescent protein.

Automated microinjections

Automated microinjections of *slc45a2*-MO, *slc45a2* sgRNA/Cas9 and the DNA construct in zebrafish embryos were performed using the robotic injector (Life Science Methods BV) following guidelines described in Spaink *et al*. 2013[38]. Briefly, all components of the robotic injector are connected to a controlling computer that is equipped with a software control program written in Python. The robot uses Eppendorf FemtoJet 4X microinjectors in combination with both in-house pulled needles and commercially available ready-to-use 10 µm tip needles as described in the section above. The established parameters for each needle (*i.e.* pressure and time) were then used for the microinjector linked to the robot injector. Zebrafish embryos were carefully arranged in each well of a 1% agarose covered grid (9 blocks x 100 wells) with the help of an artist brush. Particularly for RNA injections, each embryo was also oriented (with the artist brush) to put the one-cell visible for injection. The grid was then placed in the motorized stage coupled to a controlled and motorized micro-manipulator. After the

robotic injector was properly set (position of grid and needle) automated injections occurred at high speed (Table 2). For RNA and DNA injections in the robotic injector we used an image classification algorithm (see section below) to recognize each embryo *i.e.* "First-cell", "No-cell", "Empty", "Two-Cell", and "Sick", and to decide if triggering an injection. The total count of injected embryos (*i.e.* classified as "First-cell" and "No-cell") and of non-injected wells (classified as "Empty", "Two-Cell", and "Sick") was obtained at the end of each injection round. After each series of microinjections, the embryos were incubated at 28 °C (±0.5) and evaluated daily until 5 dpf to record non-viable embryos, (*i.e.* non-fertilized and fluorescent negative at 6 hpf, dead and dysmorphic embryos/larvae from 1 to 5 dpf), and the efficiency of injection. Table 2 shows the specifications of the different types of automated injections that were performed in this study.

	***slc45a2* MO & control MO**	***slc45a2* gRNA + Cas9 RNA**	**DNA construct + *Tol2* RNA**
Injection type	Automated	Automated	Automated
Developmental stage at injection	One- to two-cell	One-cell	One-cell
Injection location	Middle of the yolk	Depending on image classification	Depending on image classification
Injection volume	2 nL	4 nL	1 nL
Injection concentration	230 µM	400 ng/µL + 600 ng/µL	25 ng/µL + 25 ng/µL
Microinjector type	Robotic injector + Eppendorf FemtoJet 4X	Robotic injector + Eppendorf FemtoJet 4X	Robotic injector + Eppendorf FemtoJet 4X
Injection time (per 100 embryos)	2 – 3 minutes	3 minutes	3 minutes
Evaluation lapse	From 6 hpf to 5 dpf	From 6 hpf to 5 dpf	From 6 hpf to 5 dpf
Sorting criteria	GFP positive and albino phenotype	Albino phenotype	GFP positive
Exclusion criteria	Non-fertilized, GFP negative, dysmorphic and dead embryos / larvae	Non-fertilized, no albino phenotype, dysmorphic and dead embryos / larvae	Non-fertilized, GFP negative, dysmorphic and dead embryos / larvae
Place of the experiments	LCSB	LCSB	Leiden University

Table 2. Technical specifications for the automated microinjections
Abbreviations: hpf, hours post-fertilization; dpf, days post-fertilization; GFP, green fluorescent protein.

Deep learning algorithm for image classification

As a first step we used the Inception v3 network to learn to distinguish between five different categories: "Empty", "No-Cell", "First-Cell", "Two-Cell", "Sick" (this term is used to refer to non-viable eggs). We used a total of 11,000 annotated images. To prevent overfitting, we artificially increased the number of training samples by performing four types of image transformation: *1)* rotations about the center of the image; *2)* zooming by a factor 0.9-1.1; *3)* shifting by 28 pixels orthogonally in the +/- x and y direction, and *4)* flipping the image horizontally. The neural network architecture consisted of: *1)* the top part of the Inception v3 network (containing all inception blocks); *2)* a 2D global spatial average pooling layer; *3)* a fully connected layer of 1024 nodes with ReLU activation function, and *4)* a fully connected layer of 5 nodes, with softmax activation function. Training of the classification step was done using the Adam stochastic optimizer[45], with a learning rate of 10^{-4}. For a more in-depth description see S1 Text.

Deep learning algorithm for finding the injection site

For the injection point determination, we translated the (x y) coordinates to a vector in a triangular mesh using a barycentric coordinate system. We let the outputs of the neural net correspond to vertices in the mesh. In our case, we used 160 vertices.

The neural network architecture consists of: *1)* the top part of the Inception v3 network (containing all inception blocks); *2)* a 2D global spatial average pooling layer; *3)* a fully connected layer of 1024 nodes with ReLU activation function, and *4)* a fully connected layer of 160 nodes, with softmax activation function. We used 2724 images for training and 674 images for validation (these are the same images as used for label "first-cell" in the classification step). Training of the injection point determination step was done using the Adam optimizer, with a learning rate varying from 10^{-3} to 10^{-5}. More details can be found in S1 Text.

Software and hardware

For deep learning and robot control we used a Shuttle SZ170R8 equipped with an Intel Core i3 6100 CPU, 16 GB kit Kingston DDR4 2133Mhz, ECC memory and an NVidia GeForce GTX 1070 GPU. Installed software are: Keras 1.2.2, Theano 0.9.0, NumPy 1.11.0, SciPy 0.17.0. For the analysis of the data, raw data for all the series of microinjections was processed in excel. Statistical analysis was done using excel and GraphPad Prism 6 followed by unpaired *t*-test with Welch's correction for single comparisons (when applicable). The criterion for statistical significance was $P<0.05$. Graphs were plotted using GraphPad Prism 6 and error bars on all graphs represent standard deviation.

Microscopy and fluorescent analysis

At the Institute of Biology, Leiden University, representative pictures were taken using a Leica M205 FA stereo fluorescence microscope equipped with a DFC345 FX monochrome camera. Fluorescent signal was quantified using a Complex Object Parameter Analyzer and Sorter (COPAS, Union Biometrica). At the Luxembourg Centre for Systems Biomedicine, fluorescent sorting of fluorescein positive embryos (for *slc45a2*-MO injections) was done using a Nikon SMZ25 stereomicroscope. Representative pictures of control larvae and injected larvae displaying an albino phenotype were taken using the Nikon SMZ25 stereomicroscope equipped with a Nikon Digital Sight DS-Ri1 camera

Supplementary information

S1 Text. Deep learning supplement.
S1 File. Movie demonstrating robotic injections with deep learning.
S2 File. Demo source code and demo images.
S3 File. Data set used for plotting the graphs.

The supporting information can be found here:
https://doi.org/10.1371/journal.pone.0202377

References

1. Kimmel CB, Warga RM, Schilling TF. Origin and organization of the zebrafish fate map. Development. 1990; 108(4): 581-594.

2. Takaki K, Davis JM, Winglee K, Ramakrishnan L. Evaluation of the pathogenesis and treatment of Mycobacterium marinum infection in zebrafish. Nature protocols. 2013; 8(6): 1114-1124.

3. Lin S, Long W, Chen J, Hopkins N. Production of germ-line chimeras in zebrafish by cell transplants from genetically pigmented to albino embryos. Proceedings of the National Academy of Sciences. 1992; 89(10): 4519-4523.

4. Jung DW, Oh ES, Park SH, Chang YT, Kim CH, Choi SY, et al. A novel zebrafish human tumor xenograft model validated for anti-cancer drug screening. Molecular BioSystems. 2012; 8(7): 1930-1939.

5. Nicoli S, Presta M. The zebrafish/tumor xenograft angiogenesis assay. Nature protocols. 2007; 2(11): 2918-2923.

6. Nasevicius A, Ekker SC. Effective targeted gene 'knockdown' in zebrafish. Nature genetics. 2000; 26(2): 216- 20.

7. Toyama R, O'Connell ML, Wright CV, Kuehn MR, Dawid IB. Nodal induces ectopic goosecoid and lim1 expression and axis duplication in zebrafish. Development. 1995; 121(2): 383-391.

8. Suster ML, Kikuta H, Urasaki A, Asakawa K, Kawakami K. Transgenesis in zebrafish with the tol2 transposon system. Transgenesis Techniques: Principles and Protocols. 2009: 41-63.

9. Braunbeck T, Boettcher M, Hollert H, Kosmehl T, Lammer E, Leist E, et al. Towards an alternative for the acute fish LC(50) test in chemical assessment: the fish embryo toxicity test goes multi-species -- an update. ALTEX. 2005; 22: 87-102.

10. Veneman W, Spaink HP, Brun NR, Bosker T, Vijver M. Pathway analysis of systemic transcriptome responses to injected polystyrene particles in zebrafish larvae. Aquatic Toxicology. 2017; 190: 112-120.

11. Hague E, Ward AC. Zebrafish as a model to evaluate nanoparticle toxicity. Nanomaterials. 2018; 8: 1-18.

12. Chen L, Watson C, Morsch M, Cole NJ, Chung RS, Saunders DN, et al. Improving the delivery of SOD1 antisense oligonucleotides to motor neurons using calcium phosphate-lipid nanoparticles. Frontiers in Neuroscience. 2017; 11: 1-12.

13. Craig MP, Gilday SD, Dabiri D, Hove JR. An optimized method for delivering flow tracer particles to intravital fluid environments in the developing zebrafish. Zebrafish. 2012; 9: 108-119.

14. Kaethner RJ, Stürmer C. Dynamics of process formation during differentiation of tectal neurons in embryonic zebrafish. Journal of Neurobiology. 1997; 32(6): 627-639.

15. Van Asselt E, De Graaf F, Smit-Onel MJ, Van Raamsdonk W. Spinal neurons in the zebrafish labeled with fluoro-gold and wheat-germ agglutinin. Neuroscience. 1991; 43(2): 611-622.

16. Stainier DY, Lee RK, Fishman MC. Cardiovascular development in the zebrafish. I. Myocardial fate map and heart tube formation. Development. 1993; 119(1): 31-40.

17. Kim CH, Bae YK, Yamanaka Y, Yamashita S, Shimizu T, Fujii R, et al. Overexpression of neurogenin induces ectopic expression of HuC in zebrafish. Neuroscience letters. 1997; 239(2): 113-116.

18. Gore AV, Maegawa S, Cheong A, Gilligan PC, Weinberg ES, Sampath K. The zebrafish dorsal axis is apparent at the four-cell stage. Nature. 2005; 438(7070): 1030-1035.

19. Rosen JN, Sweeney MF, Mably JD. Microinjection of zebrafish embryos to analyze gene function. Journal of visualized experiments: JoVE. 2009 (25).

20. Eisen JS, Smith JC. Controlling morpholino experiments: don't stop making antisense. Development. 2008; 135(10): 1735-1743.

21. Kok FO, Shin M, Ni CW, Gupta A, Grosse AS, van Impel A, et al. Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. Developmental cell. 2015; 32(1): 97-108.

22. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature. 2015; 524(7564): 230-233.

23. Stainier DY, Kontarakis Z, Rossi A. Making sense of anti-sense data. Developmental cell. 2015; 32(1): 7-8.

24. Stainier DY, Raz E, Lawson ND, Ekker SC, Burdine RD, Eisen JS, et al. Guidelines for morpholino use in zebrafish. PLOS Genetics. 2017; 13(10): e1007000. doi: 10.1371/journal.pgen.1007000

25. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nature Reviews Genetics. 2010; 11(9): 636-646.

26. Hwang WY, Peterson RT, Yeh JR. Methods for targeted mutagenesis in zebrafish using TALENs. Methods. 2014; 69(1): 76-84.

27. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337(6096): 816-821.

28. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013; 339(6121): 823-826.

29. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature biotechnology. 2013; 31(3): 227-229.

30. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nature biotechnology. 2014; 32(4): 347-355.

31. Auer TO, Del Bene F. CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. Methods. 2014; 69(2): 142-150.

32. Clark KJ, Voytas DF, Ekker SC. A TALE of two nucleases: gene targeting for the masses?. Zebrafish. 2011; 8(3): 147-149.

33. Foley JE, Yeh JR, Maeder ML, Reyon D, Sander JD, Peterson RT, et al. Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). PloS one. 2009; 4(2):e4348. doi: 10.1371/journal.pone.0004348

34. Gagnon JA, Valen E, Thyme SB, Huang P, Ahkmetova L, Pauli A, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PloS one. 2014; 9(5):e98186. doi: 10.1371/journal.pone.0098186

35. Gupta A, Hall VL, Kok FO, Shin M, McNulty JC, Lawson ND, et al. Targeted chromosomal deletions and inversions in zebrafish. Genome research. 2013; 23(6): 1008-1017.

36. Wang W, Liu X, Gelinas D, Ciruna B, Sun Y. A fully automated robotic system for microinjection of zebrafish embryos. PloS one. 2007; 2(9):e862.

37. Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, et al. A high-throughput screen for tuberculosis progression. PloS one. 2011; 6(2):e16779. doi: 10.1371/journal.pone.0016779.

38. Spaink HP, Cui C, Wiweger MI, Jansen HJ, Veneman WJ, Marín-Juez R, et al. Robotic injection of zebrafish embryos for high-throughput screening in disease models. Methods. 2013; 62(3): 246-254.

39. Westerfield, M. The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio). 4th ed., Univ. of Oregon Press, Eugene. 2000

40. Dove A. Screening for content—the evolution of high throughput. Nature biotechnology. 2003; 21(8): 859-864.

41. Dooley CM, Schwarz H, Mueller KP, Mongera A, Konantz M, Neuhauss S, et al. Slc45a2 and C-ATPase are regulators of melanosomal pH homeostasis in zebrafish, providing a mechanism for human pigment evolution and disease. Pigment Cell Melanoma Res. 2012; 26: 205-217.

42. Irion U, Krauss J, Nüsslein-Volhard C. Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. Development. 2014; 141, 4827-4830.

43. Kwan KM, Fujimoto E, Grabher C, Mangum BD, Hardy ME, Campbell DS, et al. The Tol2kit: A multisite gateway-based construction kit for Tol2 transposon transgenesis construct. Developmental Dynamics. 2007; 11: 3088-3099.

44. Lieschke G. Zebrafish: Methods and Protocols. Springer protocols, Heidelberg. 2008.

45. Kingma D, Ba J. Adam: A method for stochastic optimization. arXiv:1412.6980. 2014.

46. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016; 2818-2826.

47. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521(7553): 436-444.

48. Ungar AA. Barycentric calculus in Euclidean and hyperbolic geometry: A comparative introduction. 2010.

49. Hruscha A, Krawitz P, Rechenberg A, Heinrich V, Hecht J, Haass C, et al. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. Development. 2013; 140(24): 4982-4987.

Supplementary figures

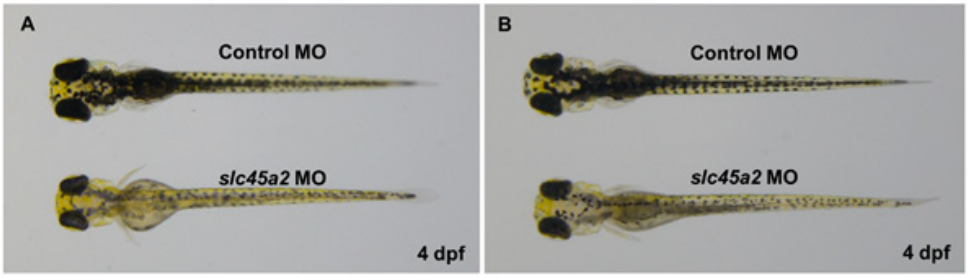

Figure S1. Representative images of manual and automated yolk injections with *slc45a2* morpholino.
The images show four days post-fertilization (dpf) larvae that were injected manually (A) or injected by the automated robot (B) with control morpholino (top) and *slc45a2* morpholino (bottom). In both cases the albino phenotype is evident in the morphant larvae (bottom), in which pigmentation is significantly reduced compared to the control morpholino (top).

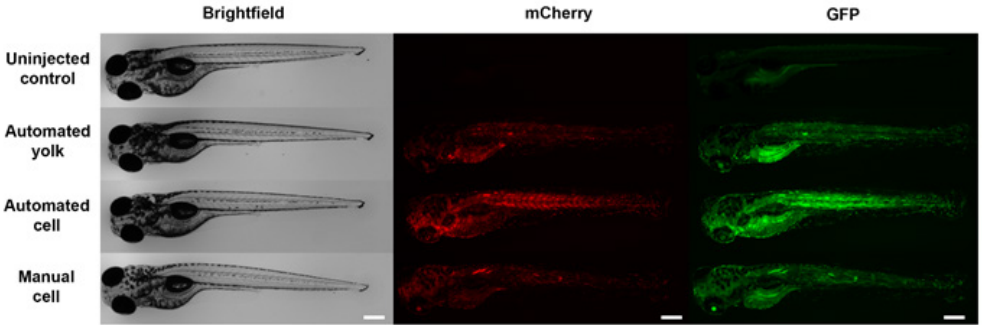

Figure S2. Representative images of manual and automated injections with (*act*:-NLSmCherry-IRES-GFP) *Tol2* construct.
The images show five days post-fertilization (dpf) larvae that were subjected to manual cell injections or automated yolk or cell injections with a (*act*:-NLSmCherry-IRES-GFP) *Tol2* DNA construct and *Tol2* transposase RNA. In all three cases the phenotype is evident as fluorescent signal is significantly increased compared to the uninjected control (top row). White scale bar = 250µm.

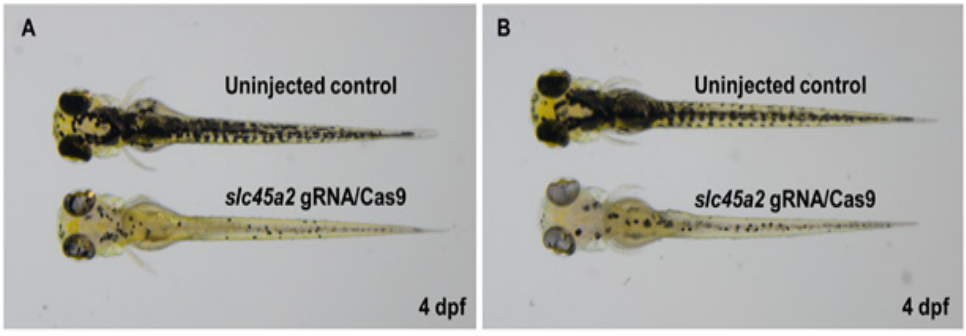

Figure S3. Representative images of manual and automated cell injections with *slc45a2* gRNA/Cas9.
The images show four days post-fertilization (dpf) larvae that were subjected to manual cell injections (A) or automated cell injections (B) with *slc45a2* gRNA/Cas9. In both cases the albino phenotype (bottom) is evident as pigmentation is significantly reduced compared to the uninjected controls (top).