PLOS GENETICS

RESEARCH ARTICLE

Multilevel regulation of the *glass* locus during *Drosophila* eye development

Cornelia Fritsch[1☯], **F. Javier Bernardo-Garcia**[1,2☯], **Tim-Henning Humberg**[1], **Abhishek Kumar Mishra**[1], **Sara Miellet**[1,3], **Silvia Almeida**[1], **Michael V. Frochaux**[4], **Bart Deplancke**[4], **Armin Huber**[5], **Simon G. Sprecher**[1]*

1 Department of Biology, University of Fribourg, Fribourg, Switzerland, **2** Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California, United States of America, **3** Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong, Australia, **4** Laboratory of Systems Biology, IBI, SV, EPFL, Lausanne, Switzerland, **5** Institute of Physiology, University of Hohenheim, Hohenheim, Germany

☯ These authors contributed equally to this work.
* simon.sprecher@unifr.ch

Citation: Fritsch C, Bernardo-Garcia FJ, Humberg T-H, Mishra AK, Miellet S, Almeida S, et al. (2019) Multilevel regulation of the *glass* locus during *Drosophila* eye development. PLoS Genet 15(7): e1008269. https://doi.org/10.1371/journal.pgen.1008269

Editor: Justin Kumar, Indiana University, UNITED STATES

Received: February 6, 2019

Accepted: June 23, 2019

Published: July 12, 2019

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This work was supported by the Swiss National Science Foundation (grant number 31003A_149499) and the Novartis Foundation for Biomedical Research (grant number 18A017) to SGS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Development of eye tissue is initiated by a conserved set of transcription factors termed retinal determination network (RDN). In the fruit fly *Drosophila melanogaster*, the zinc-finger transcription factor Glass acts directly downstream of the RDN to control identity of photoreceptor as well as non-photoreceptor cells. Tight control of spatial and temporal gene expression is a critical feature during development, cell-fate determination as well as maintenance of differentiated tissues. The molecular mechanisms that control expression of *glass*, however, remain largely unknown. We here identify complex regulatory mechanisms controlling expression of the *glass* locus. All information to recapitulate *glass* expression are contained in a compact 5.2 kb cis-acting genomic element by combining different cell-type specific and general enhancers with repressor elements. Moreover, the immature RNA of the locus contains an alternative small open reading frame (smORF) upstream of the actual glass translation start, resulting in a small peptide instead of the three possible Glass protein isoforms. CRISPR/Cas9-based mutagenesis shows that the smORF is not required for the formation of functioning photoreceptors, but is able to attenuate effects of *glass* misexpression. Furthermore, editing the genome to generate *glass* loci eliminating either one or two isoforms shows that only one of the three proteins is critical for formation of functioning photoreceptors, while removing the two other isoforms did not cause defects in developmental or photoreceptor function. Our results show that eye development and function is largely unaffected by targeted manipulations of critical features of the *glass* transcript, suggesting a strong selection pressure to allow the formation of a functioning eye.

Author summary

Changes of the genomic context can have a profound influence on gene expression. Addition or deletion of transcription factor binding sites can influence when and where a gene

Competing interests: The authors have declared that no competing interests exist.

is transcribed. Changes in exon/intron structure can affect protein length and composition. Stop codon readtrough results in the production of an elongated version of the protein. Such changes can also reduce the protein levels or even alter protein function. As a consequence, they are usually quickly removed from the genome. Thus, conservation of such traits over more than the most closely related species indicates that they are neutral or even beneficial. In the fruitfly *Drosophila melanogaster*, the *glass* gene, which is an important regulator of eye development, combines such features in its transcript, making it a good candidate to investigate these phenomena. In this study we analysed the role of the different Glass isoforms generated by intron retention and stop codon readthrough. We identified several cell- and tissue-specific enhancer elements in the *glass* regulatory sequence, and found a small open reading frame that interferes with Glass translation. Conservation of these features in other fly species suggests that their potential effects on Glass levels does not interfere with eye development.

Introduction

While genes of the retinal determination network (RDN) are necessary and sufficient for inducing eye tissue in the imaginal-disc, distinct transcription factors are subsequently involved in promoting the developmental program of cell fate determination as well as terminal differentiation. The zinc-finger transcription factor Glass provides a critical link between the RDN and terminal differentiation. Glass is required during eye development for the differentiation of photoreceptor neurons, patterning of the ommatidia, as well as for the differentiation of cone- and pigment-cells [1–4]. *glass* mutants were first discovered by H.J. Muller in 1918 and O. L. Mohr in 1919, and were named after their smaller eyes with smooth, glassy surface and altered pigmentation [5]. While it was initially assumed that photoreceptor precursors undergo apoptosis in *glass* mutants, we recently showed that these cells adopt a neuronal cell fate, extend axons and form synapses, but fail to express *rhodopsins* as well as phototransduction genes. For the determination of photoreceptor identity, *glass* promotes the terminal differentiation gene *hazy* [1, 6]. Interestingly *glass* acts in conjunction with distinct transcription factors to coordinate different cell fates during eye formation. For the specification of cone cells *glass* acts together with d*Pax2*, *eyes* absent and *lozenge*, while for the formation of pigment cells it requires *escargot* [4]. Thus, dependent on the cellular context Glass is likely to control distinct developmental programs. However, mechanisms that act to control expression of *glass* remain largely unknown.

We here provide insight into surprisingly diverse regulatory mechanisms acting to regulate the *glass* locus. By further dissecting a previously identified 5.2 kb genomic element we identified a set of regulatory core elements, including a general promotor, two pan-photoreceptor enhancer elements, a reciprocal enhancer element for non-photoreceptor cells, an element driving expression in a subset of photoreceptors as well as an ocelli-specific enhancer element. By analysing a GFP reporter including the 5'UTR we identified an alternative small open reading frame (smORF) upstream of the actual Glass translation start, resulting in a small peptide instead of the Glass protein. Interestingly, editing the corresponding genomic sequences to mutate the smORF did not cause any developmental defect nor photoattraction behaviour. However, when misexpressing Glass in a transcript including the smORF it attenuates developmental deficits, suggesting that while evolutionarily conserved within drosophilids the smORF is not essential for eye development, but may act to buffer Glass expression level. Moreover, we assessed the requirement and functionality of the three Glass

protein isoforms by CRISPR-mediated genome editing introducing deletions into the *glass* locus resulting in the loss of one or two of the three isoforms. We analysed these isoform mutants for the morphology of their eyes, the expression of photoreceptor markers that depend on Glass function, photoreceptor activity, and light preference behaviour. We found that the short Glass PB isoform is not able to confer normal eye development and function resulting in a *glass* mutant phenotype, while the Glass PA isoform alone is fully functional. Our results suggest that the expression of *glass* is tightly regulated as the development of a functional tissue surprisingly does not result in detectable change in the physiological response or alteration in photoattraction behaviour upon deletion of the smORF. Similarly, only one of the isoforms is critical for eye development. Since sequence comparison to closely related species show conservation of these features, such mechanisms may function to fine-tune gene expression.

Results

An overlapping upstream open reading frame inhibits the expression of a *glass* reporter

In the developing eye, expression of *glass* is initiated at the morphogenetic furrow in the eye-imaginal disc of third instar larvae and is detectable in the nuclei of all cells posterior to the morphogenetic furrow [7]. The same expression pattern is obtained with a reporter construct containing a 5.2 kb DNA fragment upstream of *glass* [8], spanning from -4190 bp to the AUG at +960 (Fig 1A) [8]. Surprisingly, using this 5.2 kb upstream genomic sequence to drive a GFP reporter we observed that GFP expression was barely detectable in the eye imaginal discs (Fig 1B and 1B”). By increasing the gain at the confocal microscope, we were able to detect a weak GFP signal posterior to the morphogenetic furrow, barely above background level (Fig 1B”’ and 1B””). A closer inspection of our reporter construct revealed the presence of two potential start codons in the 5’UTR of *glass*, that were also present in the GFP reporter construct, one at position +889 relative to the predicted transcription start, the other at position +955. Translation from the first start codon, if functional, may compete with the GFP start codon thus generating a protein that overlaps, but is not in frame, with the *GFP* coding sequence, resulting in the production of a 316 amino acid long protein (Fig 1B’).

To test whether translation of GFP in our reporter construct was affected by the presence of the upstream start codon(s), we generated two additional reporter constructs: one, in which the potential upstream start codons were deleted (Fig 1C’), and another, in which the GFP start codon was deleted and the GFP coding sequence was brought into frame with the upstream start codons (Fig 1D’). Both GFP reporter variants resulted in strong GFP expression posterior of the morphogenetic furrow (Fig 1C, 1C”, 1D and 1D”). Thus, the reduced GFP expression observed in the original reporter construct was caused by the translation of the reporter construct in a different reading frame due to the presence of additional start codons upstream of the GFP coding sequence. Since the GFP reporter we used does not contain a nuclear localization signal, GFP produced from its own start codon, as in construct C’, is mainly localized in the cytoplasm (Fig 1C”’ and 1C””). However, when GFP was fused in frame with the smORF, it showed strong nuclear localization (Fig 1D”’ and 1D””), suggesting that the first 24 amino acids added to the GFP coding sequence contain a nuclear localization signal. Indeed, amino acids 2 to 20 of this fusion protein are predicted to affect nuclear localization [9]. Thus, the translation of this fusion protein starts at the first AUG codon at position +889.

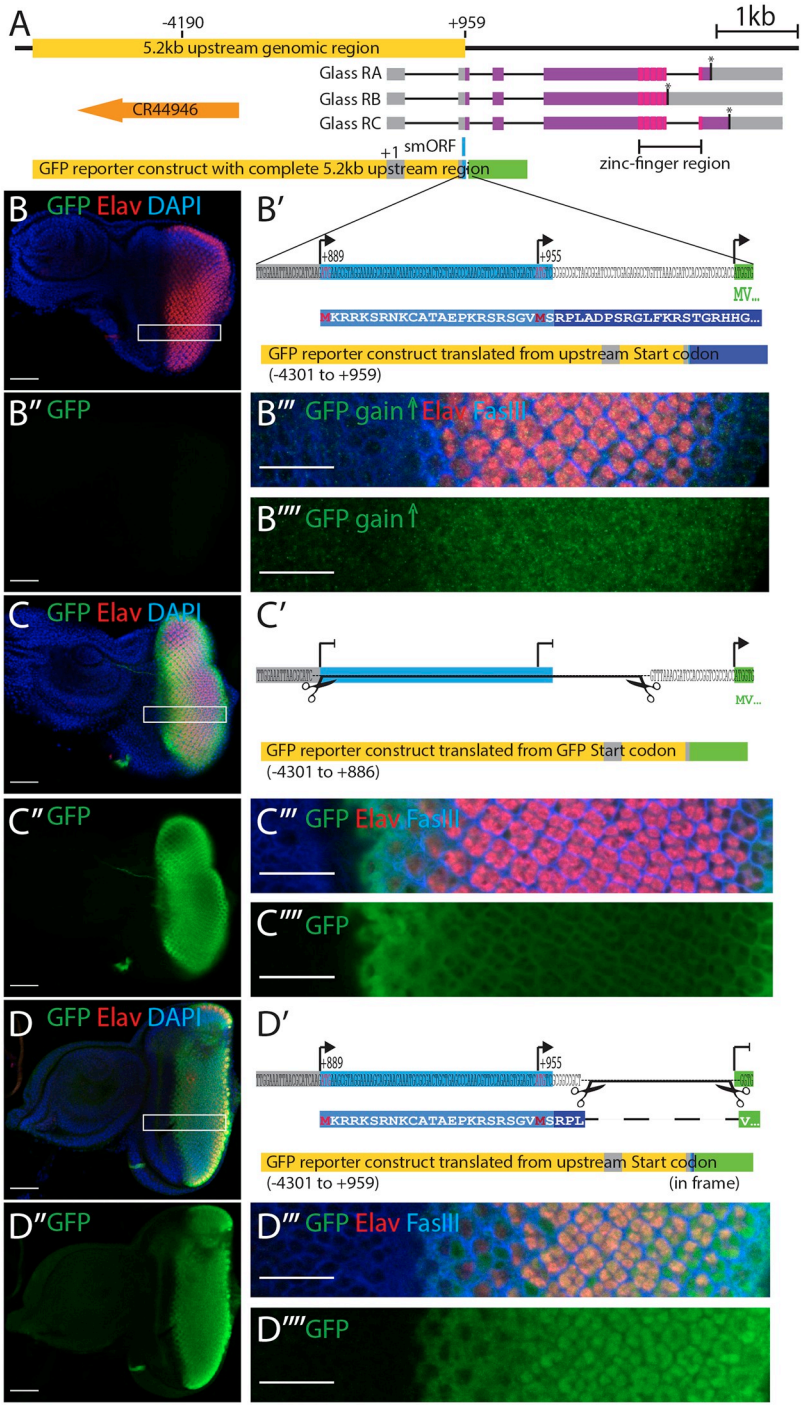

Fig 1. *glass* reporter constructs. A: genomic region of *glass* including the 5.2 kb regulatory region (yellow), the three glass isoforms (RA, RB, RC) with their intron-exon structures, stop codons (asterisks), the protein coding regions (purple), and the positions of the five zinc fingers (magenta). The position of the upstream overlapping open reading frame (smORF) is indicated in blue. A non-coding RNA (orange) is located upstream of the *glass* gene. The 5.2 kb upstream genomic region including the non-coding exon 1 and the 5'end of exon 2 were cloned in front of eGFP (green). B: eye imaginal disc of a fly transgenic for the *glass*-GFP reporter construct (-4301 to +959). The GFP expression level is very low. B': sequence fragment of the *glass*-GFP reporter construct including the 5'end of exon 2, the linker, and the first two codons of GFP (MV). The positions of the two upstream start codons and the GFP start codon are indicated by arrows. Translation from the upstream start codon results in the production of a protein encoded by the 3rd reading frame of eGFP (first 44 amino acids shown in blue box). B": GFP channel alone of the disc

shown in panel B. No GFP is detectable (gain: 621.7). B"': close up of the region indicated by the white box in B. B"": GFP channel alone of the region outlined in panel B imaged with a higher gain (827.1). GFP levels are slightly higher in the posterior region of the disc. C: eye disc of a transgenic fly expressing a *glass*-GFP reporter construct in which the two upstream start codons were deleted. GFP is expressed at high level in the posterior region of the disc that will give rise to the adult eye. C': Sequence fragment of the GFP reporter construct indicating the part that was excised to remove the upstream start codons (-4301 to +886). Translation can only start at the GFP start codon. C": GFP channel alone of the disc shown in panel C (gain: 621.7). C"': close up of the region indicated by the white box in C. C"": GFP channel alone of the region outlined in panel C (gain 621.7). eGFP is mainly cytoplasmic. D: eye disc of a transgenic fly expressing a *glass*-GFP reporter construct in which the GFP start codon was deleted and the GFP coding sequence is in frame with the upstream start codon(s). D' Sequence fragment of the GFP reporter construct indicating the part that was excised to remove the GFP start codon (-4301 to +959 in frame). The N-terminus of the resulting fusion protein between the upstream translation product and GFP is shown below). D": GFP channel alone of the disc shown in panel D (gain: 757.2). D"': close up of the region indicated by the white box in D. D"": GFP channel alone of the region outlined in panel D (gain 757.2). GFP shows nuclear localization. All discs are oriented with the posterior to the right. Discs were stained with antibodies against GFP (green), Elav (red), FasIII (blue in panels B"', C"' and D"') and with DAPI (blue in panels B, C and D). Scale bars in panels B, B", C, C", D and D" represent 50 μm, in panels B"', B"", C"', C"", D"' and D""they represent 20 μm.

https://doi.org/10.1371/journal.pgen.1008269.g001

Defined enhancer elements confine cell type specificity and temporal restricted expression

In order to understand the cis-regulatory logic of *glass* expression we further dissected this genomic region in the construct that does not contain the two upstream start codons (Fig 1C'). Using a number of restriction sites located in the upstream regulatory sequence, we generated truncations of our GFP reporter, similar to those used by Liu et al. [8], and also tested some deletions within this upstream sequence (Fig 2A). After deleting half of the 5.2 kb fragment (construct B: -1885 to +886), GFP expression is still restricted to the region posterior of the morphogenetic furrow (Fig 2B). Further deletion of a small fragment between the BamHI and EcoRI sites (construct C: -1598 to +886) shows patchy GFP expression in the developing photoreceptor precursors (Fig 2C). While construct B is expressed in all cell types forming the presumptive eye, the expression of construct C is restricted to presumptive photoreceptor cells with variable expression levels (S1A Fig), suggesting that the fragment from -1885 to -1598 might contain some non-photoreceptor specific enhancer. A fragment truncated at the XbaI site (construct D: -703 to +886) is expressed in all the photoreceptor precursors posterior of the morphogenetic furrow with the highest levels directly after the furrow and reduced levels towards the posterior end (Fig 2D). This construct also shows ectopic expression in a stripe anterior of the furrow (Fig 2D' arrowhead). This misexpression of GFP is spreading over the entire eye-antenna-disc in a construct starting at the XhoI site (construct E: -239 to +886, Fig 2E), suggesting that this fragment contains a minimal promoter whose activation is independent of eye specific enhancers. We used this minimal promoter region in combination with other fragments of the enhancer to analyse the expression patterns conferred by the 5' enhancer elements. We tested the 287 bp fragment between the BamHI and EcoRI sites that we suspected to drive expression specifically in non-photoreceptor cells based on the different expression patterns between constructs B and C. We found that this small fragment in combination with the minimal promoter (construct F: -1885 to -1598 / -239 to +886) can restrict GFP expression to the region posterior to the morphogenetic furrow (Fig 2F). With this enhancer fragment, the GFP signal is absent in the presumptive photoreceptor cells and restricted to the cells surrounding the photoreceptor precursors (Fig 2F"', S1B Fig). A complementary construct lacking only this small region (construct G: -4301 to -1906 / -1598 to +886), shows a reciprocal expression pattern posterior of the furrow with expression restricted to presumptive photoreceptors (Fig 2G). The 1.2 kb region located at the 5' end of the *glass* enhancer fragment in combination with the minimal promoter (construct H: -4301 to -3123 / -239 to

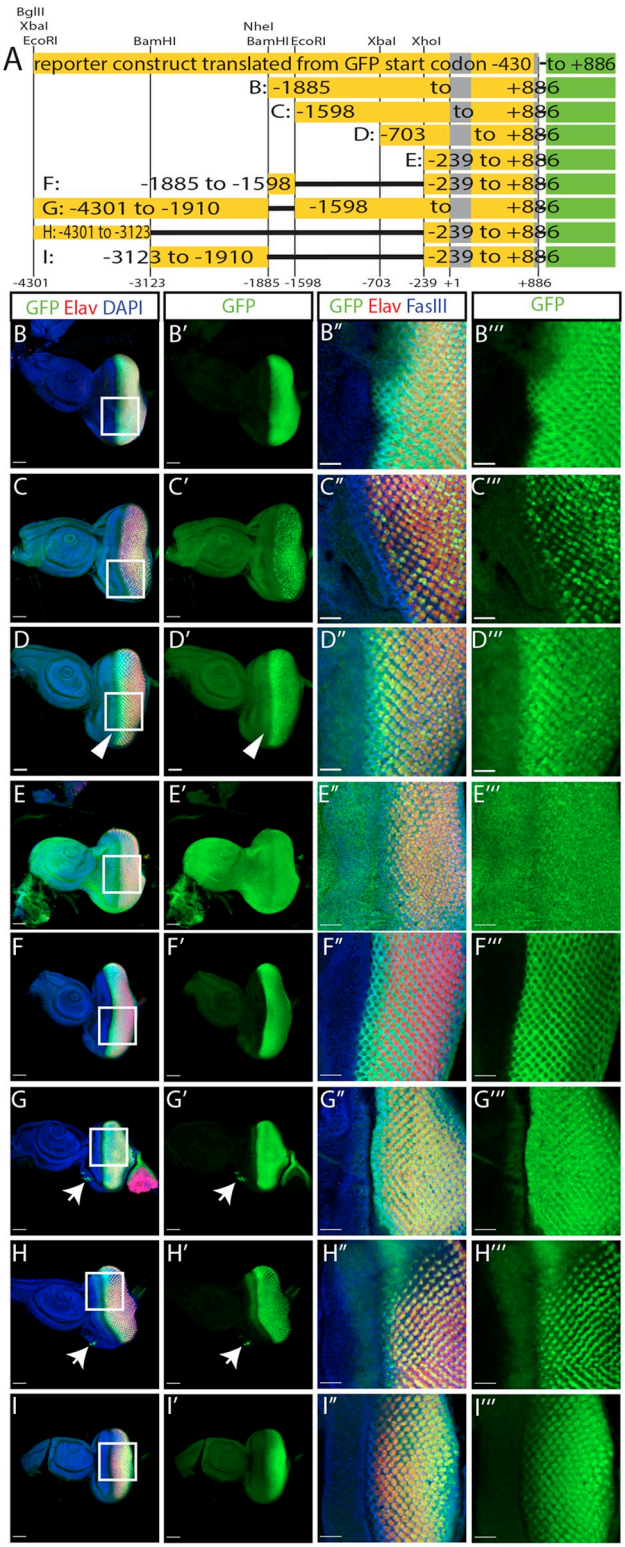

Fig 2. Functional analysis of the *glass* enhancer region. A: enhancer fragments used to drive GFP expression. The construct with the deleted upstream start codons (top) was fragmented using the restriction sites indicated above. The resulting reporter constructs are shown below. B-I: GFP expression patterns of constructs B to I. All discs are oriented with the posterior end to the right. Scale bars: 50μm. Arrow heads: GFP expression anterior of the morphogenetic furrow. Arrows: ocelli anlage. B"–I"': close ups of the regions marked by the rectangles in panel B to I. Scale bars:

20 μm. Discs were stained with antibodies against GFP (green), Elav (red), FasIII (blue in panels B' to I'), and with DAPI (blue in panels B-I).

https://doi.org/10.1371/journal.pgen.1008269.g002

+886), also restricts GFP expression to cells posterior of the morphogenetic furrow (Fig 2H). In this case GFP is only expressed in three of the eight presumptive photoreceptors (Fig 2H'''). We identified these as R2, R5, and R8 using defined markers [10] (S1C Fig). In addition, this part of the *glass* enhancer is required for expression in the ocelli anlage (Fig 2G and 2H arrows). Finally, the fragment between the two BamHI sites (construct I: -3123 to -1906 / -239 to +886) drives GFP expression in all presumptive photoreceptors (Fig 2I and 2I', S1D Fig), similar to construct D, but with lower expression levels directly after the furrow and increasing GFP levels towards the posterior end.

Taken together, the 5.2 kb *glass* regulatory region contains a general promoter region (-239 to +886), an ocelli enhancer region (-4301 to -3123), that also drives expression in a subset of photoreceptor precursors, a non-photoreceptor enhancer element (-1886 to -1598), two general photoreceptor enhancer elements (-3123 to -1906 and -1598 to -239), and a repressor region (-1598 to -703) (S2A Fig).

glass expression is directly regulated by Sine oculis, a member of the retina determination network [1]. The 5.2 kb enhancer region contains 31 potential Sine oculis binding sites (altogether 10 sites with a perfect AGATAC consensus sequence [11] and 22 sites with a more degenerate version YGATAY [12], S2A Fig). Expression of a reporter gene driven by the non-photoreceptor enhancer fragment (-1886 to -1598) is lost if the three Sine oculis binding sites present in this fragment are mutated [1]. Thus, Sine oculis might act as a general activator of *glass* expression by binding to all enhancer elements, while other transcription factors might be binding more specifically to an individual enhancer element conferring expression in only a subset of the cells. We performed an *in silico* analysis looking for additional transcription factor binding sites in the entire 5.2 kb region. We found potential binding sites for 263 different transcription factors including factors that are known to regulate eye development (p.e. Pph13, Optix, Otd, Hth, Ey, Dr, Kr, Ato) (S2B Fig, S1 Table). There are also several potential binding sites for Glass within the photoreceptor enhancer regions suggesting an auto-regulatory function. Our identification of the different enhancer regions will allow more specific testing of the role of these transcription factors in the cell-specific regulation of *glass*.

The overlapping upstream smORF is conserved and attenuates Glass misexpression

The upstream start codons in the 5'UTR of *glass* strongly reduced the expression of our original GFP reporter construct, presumably due to interference with GFP translation and production of a 316 amino acid long protein encoded in the 3rd frame of the eGFP sequence used here. In the *glass* transcript, translation from the upstream start codon might also interfere with Glass translation producing a 34 amino acid long peptide encoded by the smORF that overlapps with the Glass coding sequence (Fig 3A(b)). Interestingly, the 4 nucleotide sequence preceding the upstream start codon (CAAG) is more similar to the *Drosophila* consensus Kozak sequence (MAAM, whereby M stands for either A or C) [13] than the sequence upstream of the actual Glass start codon (TGTC) (Fig 3A). Sequence comparison with *glass* genes from other Diptera revealed that upstream start codons are present in all *glass* 5'UTRs of Drosophilidae as well as in *Lucilia*, *Musca*, and *Glossina*, possibly producing peptides that overlap with the Glass coding sequence (S3A Fig). Although the length of these peptides differs slightly due to insertion and deletion of nucleotide triplets, the frameshift relative to Glass and

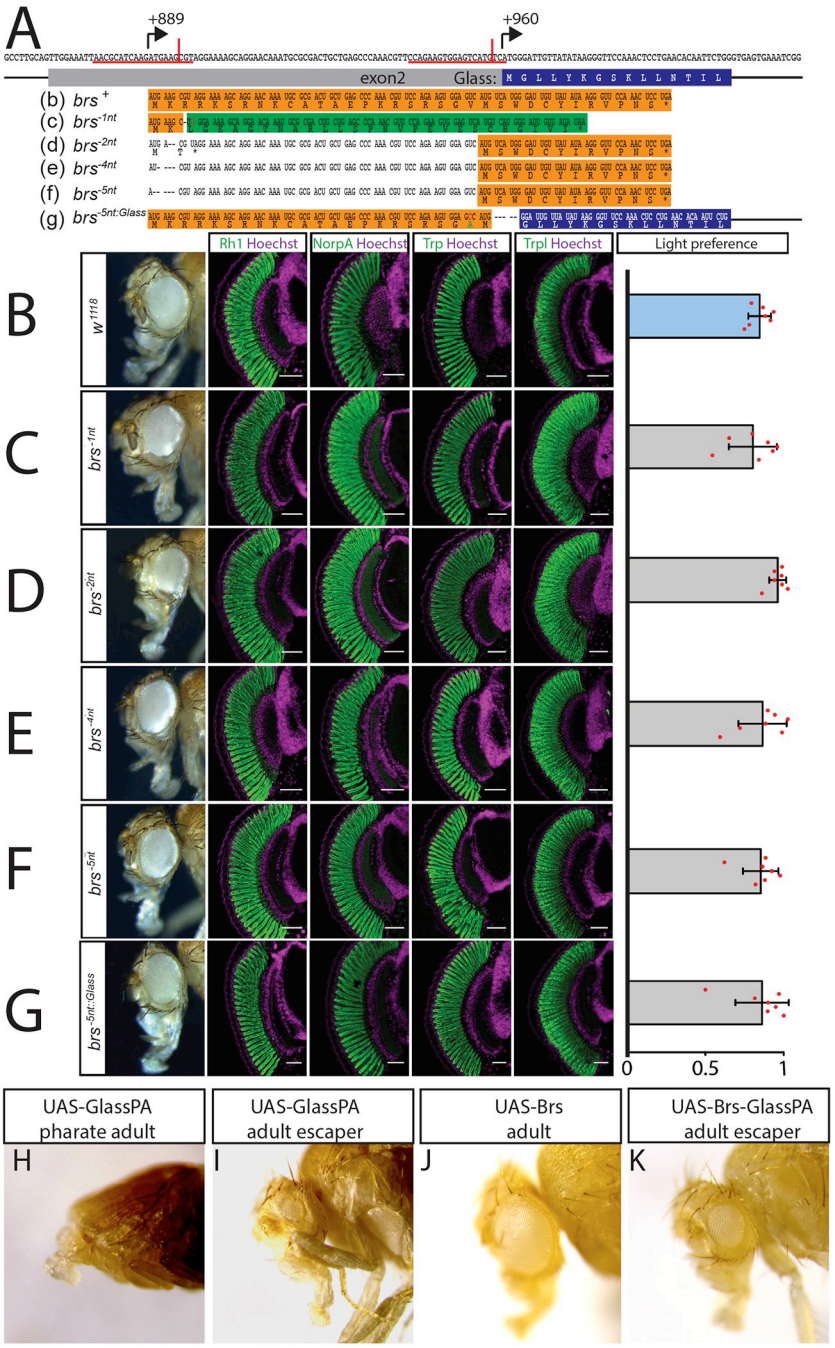

Fig 3. The upstream overlapping open reading frame is not required for eye development or function. A: sequence of *glass* exon 2. The positions of the upstream start codon (+889) and the Glass start codon (+960) are indicated by arrows. CRISPR sites are underlined in red, and the actual positions of the cuts are indicated by a vertical red line. The amino acid sequence of the Glass N-terminus is written in the blue box under the coding sequence. The *brs* nucleotide and amino acid sequence is shown in the orange box (b). A single nucleotide deletion was introduced in the brs^{-1nt} allele (followed by an A to T mutation to generate a HindIII restriction site (c)). This will result in a frameshift completely changing the amino acid sequence of the Brs peptide (shown in the green box) with only a minimal change in the nucleotide sequence of exon 2. The resulting peptide still overlaps with the Glass open reading frame. Cas9 also produced small deletions at the CRISPR site (d-f) resulting in a loss of the upstream start codon. Translation could still start from a second AUG just upstream of the Glass AUG interfering with Glass translation. A 5 nucleotide deletion was introduced in the $brs^{-5nt::Glass}$ allele between the second Brs AUG and the second codon of Glass, putting both sequences into the same frame as indicated by the orange and blue highlighted sequences (g) (a T to C mutation was introduced to generate an NcoI site, changing the valine at position 22 of this fusion protein into alanine). B-G: adult

eye, expression of the retinal markers Rh1, NorpA, Trp, and Trpl and light preference of control w^{1118} flies (B), brs^{-1nt} deletion flies (C), random deletion lines (D-F), and $brs^{-5nt::Glass}$ deletion flies (G). All antibody stainings are shown in green, DNA was stained with Hoechst (purple). Scale bars represent 40 μm. Flies of all tested genotypes are attracted by light. Two-tailed one sample *t* test followed by the Benjamini Hochberg procedure: For all data sets $n = 7$ experiments. CTRL: $p = 3.6^{*}10^{-8}$, $t_{(6)} = 31.09$; brs^{-1nt}: $p = 1^{*}10^{-5}$, $t_{(6)} = 13.88$; *brs-2nt*: $p = 3.6^{*}10^{-8}$, $t_{(6)} = 47.23$; *brs-4nt*: $p = 8.9^{*}10^{-6}$, $t_{(6)} = 14.81$; *brs-5nt*: $p = 2^{*}10^{-6}$, $t_{(6)} = 19.97$; $brs^{-5nt::Glass}$: $p = 1.1^{*}10^{-5}$, $t_{(6)} = 13.38$. The light preference index of all experimental groups is not different from the light preference index of the control group. One-way ANOVA of preference indices: $p = 0.495$, $F\text{-}Value_{(5,36)} = 0.895$. Data show mean and error bars show standard deviation. Red dots indicate means of individual experiments. H-K: overexpression of UAS constructs with a strong *ey*-Gal4 driver. H: Overexpression of the Glass PA protein leads to severe eye and head defects resulting in pharate lethality. I: the only *ey*-Gal4>*UAS-Glass-PA* fly that eclosed had very small eyes. J: Overexpression of the Brs peptide did not affect eye or head development. K: When expressed together on the same UAS-construct, Brs translation interferes with Glass translation resulting in a higher number of escapers that have small or even normal eyes.

https://doi.org/10.1371/journal.pgen.1008269.g003

the amino acid sequence are conserved within the Drosophilidae, suggesting that the encoded peptide itself might have a conserved function (S3B Fig). Interestingly, the N-terminal half of the peptide contains mainly basic residues that can provide a nuclear localization signal, as revealed in the GFP reporter construct that was cloned in frame with the upstream start codon (Fig 1D). The central part of the peptide sequence is more variable and truncated in *D. grimshawi*, *D. virilis*, and *D. mojavensis*, while it is extended in *D. wilistoni*, *L. cuprina*, *M. domestica*, and *G. morsitans* (S3B Fig). Not surprisingly, conservation is also high in the C-terminal part overlapping the Glass coding sequence. The N-termini of the Glass orthologs of other insects, including mosquitoes, are not conserved, and there are no upstream overlapping open reading frames in these transcripts.

Since the 34 amino acid long peptide is encoded by the *glass* mRNA, it might have a function in eye development. We used the CRISPR/Cas9 technique to introduce small deletions in the peptide coding sequence that will result in a frameshift of the peptide without affecting the Glass coding sequence. We named the resulting smORF alleles "*brainy smurf*" (*brs*) after the smurf with the glasses. We introduced a double strand break 6 nucleotides downstream of the start codon of the peptide and provided a template for repair that contained a single nucleotide change as well as a single nucleotide deletion (brs^{-1nt}) (Fig 3A(c)). Glass transcript levels were not significantly altered in these brs^{-1nt} mutants in comparison to the original *nos-Cas9* line (S4B Fig). In addition to the single nucleotide deletion provided by the template, we also found several lines that had small indels in the region of the CRISPR site used (Fig 3A(d-f)). Although after injection of the gRNA and crossing the G0 flies with a deficiency line that uncovers the glass locus we had selected F1 flies that showed a subtle rough eye phenotype, after establishing stable lines the eyes did not show any morphological defects (Fig 3B–3F). We therefore stained for the expression of different photoreceptor markers. We used w^{1118} flies as controls, since our brs^{-} lines were in a w^{-} background due to crossing the G0 flies with *w;; Df(3R)Exel6178* and the F1 flies with *w;; Dr e/TM3*. w^{1118} control flies have big round compound eyes expressing the phototransduction proteins Rhodopsin1 (Rh1), No receptor potential A (NorpA), Transient receptor potential (Trp), and Transient receptor potential-like (Trpl) (Fig 3B). Adult flies are attracted to light in phototaxis experiments [6]. This is also the case for white eyed mutants (Fig 3B). brs^{-1nt} homozygous flies have normal eyes, expressing all tested markers and show light preference comparable to wildtype flies (Fig 3C). We observed the same phototaxis behaviour and marker gene expression in the randomly generated *brs* mutations (Fig 3D–3F).

The GFP reporter constructs demonstrated that the presence of the upstream start codon can interfere with translation from the actual start codon. To test if this is also the case for the translation of Glass, we introduced a 5 nucleotide deletion at the Glass start codon, putting it into frame with the upstream start codon ($brs^{-5nt::Glass}$) (Fig 3A(g)). Again, no changes in glass

Table 1. Brs interferes with Glass translation. A strong *ey-Gal4/CyO* driver line was crossed with different UAS-constructs and the number of eclosed Cy^+ and *CyO* flies was determined. The *CyO* siblings do not contain the ey-Gal4 driver and therefore were taken as reference for the amount of Cy^+ flies expected in each experiment. The ratio of Cy^+ over *CyO* flies determines the survival rate of flies expressing the UAS-construct.

UAS-constructs	# of Cy^+ flies	# of *CyO* flies	Total	Survival rate
Glass-PA	1	267	268	0.4%
brs	409	408	817	100.2%
brs-glass-PA	134	661	795	20.3%
glass-PA; UAS-brs	2	702	704	0.3%

https://doi.org/10.1371/journal.pgen.1008269.t001

transcript levels were observed in this line in comparison to *nos-Cas9* flies (S4B Fig). From these transcripts, Glass translation starts at the upstream start codon that has the "better" Kozak sequence and fuses the nuclear localization signal encoded by the N-terminus of Brs to the Glass protein. We considered that this could result in higher levels of Glass activity that might interfere with eye development. However, we did not observe any changes in eye morphology, marker gene expression, photoreceptor shape, or light preference (Fig 3G). Thus, the potential increase of Glass protein either does not interfere with its function or is compensated by other mechanisms.

To test our hypothesis that the upstream start codon might interfere with Glass translation, we used an over-expression assay. Driving *UAS-glass-PA* expression with a strong *eyeless*-Gal4 enhancer results in lethality of the pharate flies (Fig 3H, Table 1). The flies have severe head defects that prevent them from eclosing with only a small number of escapers (0.4%) (Fig 3I). Overexpression of a *UAS-brs* construct did not affect viability of the flies or their eye shape (Fig 3J, Table 1), indicating that the small peptide produced does not interfere with eye development. Co-expression of *UAS-glass-PA* and *UAS-brs* inserted on different chromosomes showed a similar level of lethality as *UAS-glass-PA* alone (0,3% survival rate), indicating that the peptide itself does not interfere with Glass function. However, in a construct that contains the peptide coding sequence upstream and overlapping with the Glass PA coding sequence as in the endogenous transcript, the lethality caused by the over-expression of Glass protein was reduced, resulting in a 20.3% survival rate, where the adult escapers had normal or smaller eyes (Fig 3K). Therefore, the presence of the Brs peptide itself does not reduce Glass levels. *brs* only interferes with Glass translation, when directly placed as an upstream overlapping open reading frame in the *glass* mRNA.

The Glass PA protein isoform by itself is sufficient for photoreceptor differentiation and function

glass encodes a 604 amino acid protein containing a transcriptional activation domain and a DNA-binding domain that consists of five zinc-fingers of which the three C-terminal zinc-fingers were shown to be necessary and sufficient for DNA binding [7, 14]. However, according to FlyBase (flybase.org), the *glass* gene encodes three different protein isoforms (Fig 1A). The PA isoform contains a complete set of 5 zinc-fingers providing sequence specific DNA binding to the target genes of Glass [3, 15]. In addition to the Glass PA isoform, two other isoforms are predicted to exist based on expressed sequence tags and sequence conservation [16, 17]. Failure to splice out the last intron of the mRNA transcript, results in the production of a truncated 557 amino acid Glass PB isoform lacking the second half of the fifth zinc-finger. This version of Glass cannot bind specifically to its target sequence *in vitro* [14]. Of the 19 cDNA clones whose sequences are available on FlyBase (flybase.org), seven are covering the last and/or the second-last exon, and all seven still contain the last intron, suggesting that this intron is

frequently retained in the transcript. The position of the last intron (intron 4 in *Drosophila*) including the stop codon immediately following the exon-intron junction is only present in the *glass* orthologs of Diptera and Lepidoptera (S5A Fig). In the postman butterfly (*Heliconius melipone*) the stop codon is not located immediately after the exon intron junction but 17 basepairs into the intron. Other arthropods do not have an intron at this position.

An extended 679 amino acid long Glass PC isoform, containing all 5 zinc-fingers followed by additional 75 amino acids, is produced by a readthrough of the Glass PA stop codon. The prediction of this longer isoform is based on sequence conservation 3' of the regular stop codon [18] and was confirmed using the Coding Potential Assessment Tool (CPAT) that calculates the coding probability of a DNA sequence using *Drosophila* statistics [19]. A comparison of the sequence following the Glass stop codon within the higher Diptera shows conservation on the amino acid level suggesting that the extended protein is produced by a direct misinterpretation of the stop codon without shifting the reading frame (S5B and S5C Fig) [20, 21]. The amino acid sequences of the extended Glass proteins from different higher Diptera are highly conserved at their N- and C-termini, but have a central region that is rich in histidine residues of very variable length. Particularly, the *Musca* and *Lucilia* Glass PC versions contain a high number of additional amino acids in this central part. The PC sequence is not conserved in other insects including mosquitoes.

To test the requirement and function of the three different Glass isoforms *in vivo*, we introduced specific changes in the endogenous *glass* locus by CRISPR/Cas9- mediated genome editing, eliminating either one or two of the Glass isoforms (Fig 4A). We used w^{1118} flies as controls, since our *glass* deletion lines were in a w^{-} background due to crossing the G0 flies with *w;; Df(3R)Exel6178* and the F1 flies with *w;; Dr e/TM3* (Fig 4B). Flies expressing only the Glass PA+PC isoforms due to a deletion of the last intron had normal, functional eyes expressing phototransduction proteins like control flies (Fig 4C). In contrast, a deletion that allowed only the production of the truncated PB isoform phenocopied *glass* amorphic mutations, in which photoreceptors failed to differentiate as revealed by the loss of phototransduction proteins (Fig 4D) [1]. We also prevented the production of the PC isoform by adding two additional stop codons at the end of the Glass PA sequence (PA+PB). This had no effect on eye shape or on photoreceptor marker gene expression (Fig 4E). By deleting the last intron and adding stop codons at the end of the Glass PA sequence we generated flies that can only express the PA isoform. These flies also have normal functional eyes expressing all tested markers (Fig 4F). In addition to marker gene expression, we also measured photoreceptor activity by recording electroretinograms (ERGs). We found that all isoform mutants that had normal eye shape and were expressing phototransduction proteins, showed normal ERG responses [22], while the flies expressing only the Glass PB isoform did not produce any ERG signal in response to light (Fig 4G). When we tested the light preference of our different Glass isoform mutants, we found that all variations expressing the Glass PA isoform showed light preference comparable to wildtype flies (Fig 4H). In contrast, the flies expressing only the Glass PB isoform were photoneutral, with a light preference index that is not significantly different from chance, but significantly different from that of control flies and similar to that of *glass* mutants, which fail to detect light [6].

Glass is also required for the development of other light sensing organs in *Drosophila*. In addition to their compound eyes, adult fruitflies have three ocelli on top of their heads expressing Rhodopsin 2 (S6A Fig) [23, 24], and four photoreceptor cells on each side of the head forming the eyelet, that is located underneath the retina and required for regulation of the circadian rhythm (S6B Fig) [24, 25]. During larval stages 12 photoreceptor cells on each side of the head form the larval eyes (Bolwig organs), with four photoreceptor cells expressing Rhodopsin 5 and the other eight expressing Rhodopsin 6 (S6C Fig) [26]. All these visual organs are

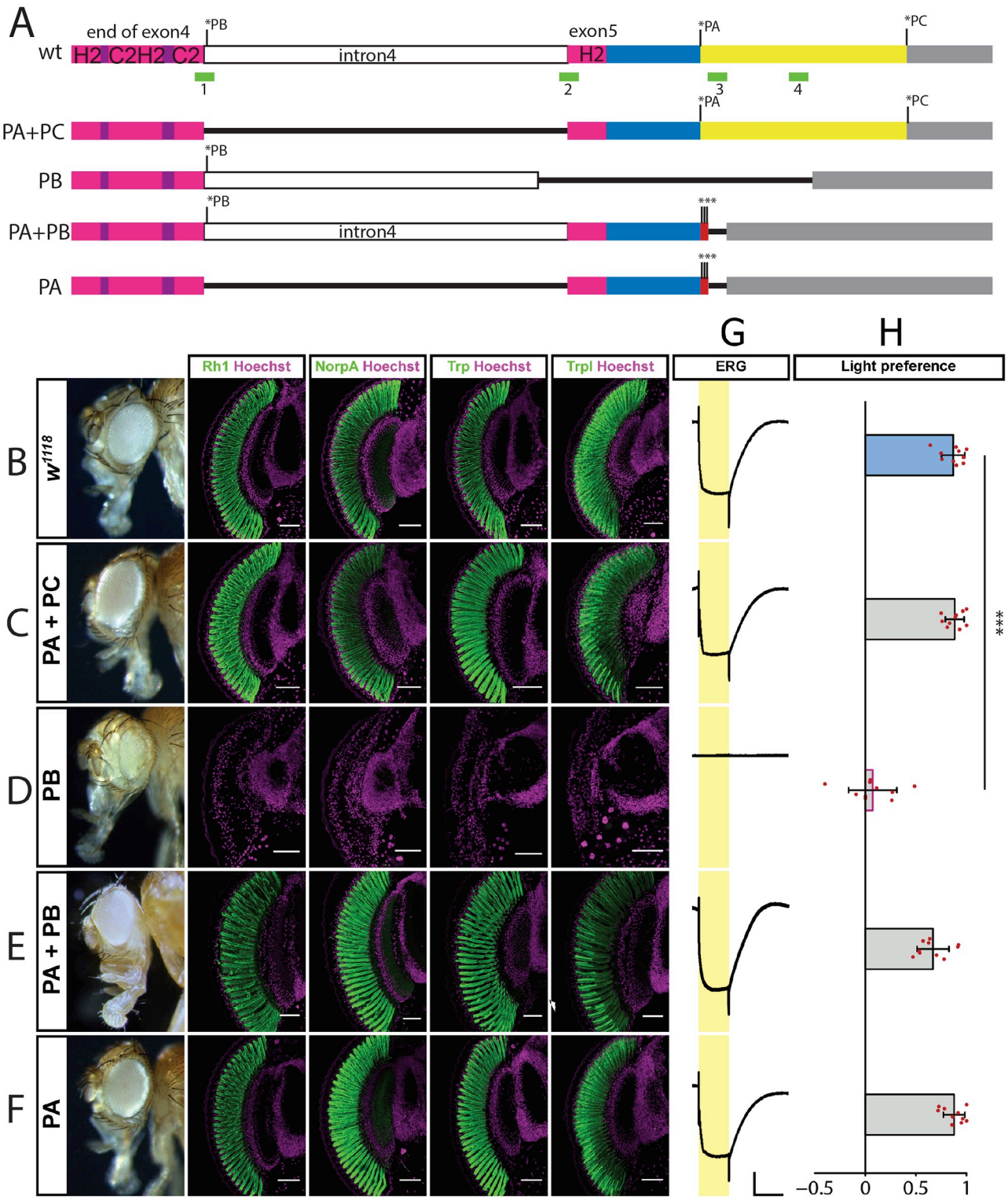

Fig 4. Isoform specific *glass* alleles. A: wildtype and mutated versions of *glass* from the end of exon 4 to the end of the transcript. Stop codons of isoforms PA, PB, and PC are indicated by asterisks. The C2H2-zinc-finger region is shown in purple and magenta. Intron 4 is shown in white. The C-terminus of the PA isoform is shown in blue, that of the PC isoform in yellow. The 3'UTR is grey. The positions of the CRISPR sites used for mutagenesis are shown as green boxes. Deletions are indicated as black lines. The triple stop codon introduced in the PA+PB and the PA alleles are indicated by red boxes and asterisks. B-F: adult eye and expression of the retinal markers Rh1, NorpA, Trp, and Trpl as indicated above in control w^{1118} flies (B), in flies expressing the Glass PA+PC isoforms (C), in flies expressing the Glass PB isoform (D), in flies expressing the Glass PA+PB isoforms

(E), in flies expressing the Glass PA isoform (F). All tested markers are expressed in the different alleles, except in the mutant only expressing the PB isoform, which has a *glass* mutant eye phenotype. All antibody stainings are shown in green, DNA was stained with Hoechst (magenta). Scale bars represent 40 μm G: ERGs of the different *glass* isoform alleles indicated at the left, scale bars represent 5 mV (vertical) and 5 seconds (horizontal). H: Light preference index of wildtype and flies that are mutant for specific Glass isoforms. Flies expressing only the PB isoform are photoneutral. Two-tailed one sample *t* test followed by the Benjamini Hochberg procedure: For all data sets $n = 10$ experiments. CTRL: $p = 5.8^{*}10^{-09}$, $t_{(9)} = 23.82$; PA+PC: $p = 2.1^{*}10^{-09}$, $t_{(9)} = 30.20$; PB: $p = 0.46$, $t_{(9)} = 0.97$; PA+PB: $p = 6.6^{*}10^{-07}$, $t_{(9)} = 13.43$; PA: $p = 3.8^{*}10^{-09}$, $t_{(9)} = 26.19$. Only flies expressing just the Glass PB isoform show a light preference which is different from the one of control flies. One-way ANOVA of preference indices: For all data sets $n = 10$ experiments, $p<2^{*}10^{-16}$, $F\text{-}Value_{(8,81)} = 44.04$. CTRL vs PA+PC: $p = 1$, $t = -0.15$; CTRL vs PB: $p<0.001$, $t = 8.11$; CTRL vs PA+PB: $p = 0.23$, $t = 2.03$; CTRL vs PA: p = 1, $t = -0.10$. Data show mean and error bars show standard deviation. Red dots indicate means of individual experiments. *** = $p<0.001$.

https://doi.org/10.1371/journal.pgen.1008269.g004

present in the flies expressing the Glass PA+PC isoforms (S6D–S6F Fig). These photoreceptors are also fully developed in flies expressing the Glass PA+PB isoforms (S6G–S6I Fig) and in flies expressing only the Glass PA isoform (S6J–S6L Fig). In contrast, flies expressing only the truncated Glass PB isoform not only show a *glass* mutant phenotype in their compound eyes, but also don't have ocelli. We were not able to identify fully differentiated photoreceptors of the Bolwig organ, the eyelet, or the ocelli based on their expression of Rhodopsins or Chaoptin. Therefore, we stained for Krüppel and Spalt, two transcription factors that are already expressed in the Bolwig organ precursors at embryonic stage 12 (S6M Fig), whose expression is maintained at larval stages in wildtype flies (S6N Fig). Krüppel is expressed in all twelve cells, while Spalt is only expressed in the four primary precursors that will later express Rhodopsin 5 (S6O Fig). In stage 12 embryos containing only the Glass PB isoform the Bolwig organ precursors are specified and two cells even express Spalt and Krüppel (S6P Fig). However, at larval stage neither these cell type specific transcription factors nor the rhodopsins are detectable suggesting that the cells lost their identity and probably never reached their final position (S6Q and S6R Fig).

To see if the expression of only one or two of the Glass isoforms would affect glass transcript levels, we performed qPCR on the different isoform mutant lines. We used *nos-Cas9* flies for comparison since the genomic changes had been introduced in this line using CRISPR. We used a forward primer located in exon 4 in combination with a reverse primer located in exon 5 to quantify the amount of spliced transcript and the same forward primer in combination with a reverse primer located in intron 4 to quantify the amount of unspliced transcript. Since the gl^{PB} line has a deletion of the part of exon5 that is bound by the reverse primer, we only tested unspliced transcript levels in this line. Similarly, since intron 4 was deleted in the gl^{PA+PC} and the gl^{PA} lines, we only tested spliced transcript levels in these lines. The *glass* transcript levels in the gl^{PA+PC} and in the gl^{PA} lines did not differ significantly from control flies (S4A and S4B Fig). We found a slight but significant increase of transcript levels in the gl^{PA+PB} line for the spliced as well as the unspliced version. This might be due to higher mRNA stability as result of the loss of translation of the PC isoform or due to higher transcription levels as response to the failure to produce the extended version of the protein. Finally, we found that in the gl^{PB} line the amount of unspliced transcript is significantly increased in comparison to the rather small amount found in wildtype flies and reaches a level in the same range as the spliced transcript found in wildtype flies, although due to the different reverse primers used, the expression levels cannot be directly compared. Thus, in the line that shows a *glass* mutant phenotype due to the loss of production of a functional protein, *glass* expression is not heavily upregulated.

A *glass* mutant phenotype was also observed in flies in which, after CRISPR-induced DNA double strand break, the DNA repair occurred in form of non-homologous end joining, either deleting the exon-intron junction and the stop codon located in the last intron, or introducing a frameshift at the beginning of the last exon (S7A Fig). Like flies expressing the truncated Glass PB isoform, these flies also have small eyes with a glassy surface. They do not express any

of the tested photoreceptor makers, have no ERG response and do not show phototaxis behaviour (S7B–S7G Fig).

Thus, although conserved, the extended version of Glass is dispensable in photoreceptors. In contrast, the truncated PB version alone cannot fulfil Glass function, while its absence does not interfere with Glass function in the eye.

Discussion

Changes of the genomic context in a given locus can have a profound influence on gene expression. Addition or deletion of transcription factor binding sites does not only affect when and where a gene is transcribed but can also determine the expression level. The addition of an exon containing a start codon upstream of the first coding exon can result in an N-terminal extension of the protein if the start codon is in the same reading frame as the coding sequence, or it might interfere with translation if it is in a different reading frame. Insertion of an intron within the coding sequence can result in the production of a truncated protein due to intron retention. Stop codon readthrough can lead to the production of an elongated version of the protein. Such changes can reduce the protein levels or even alter protein function meaning that they are usually quickly removed from the genome. Thus, conservation of such traits over more than the most closely related species indicates that they are neutral or even beneficial. The *Drosophila* transcription factor *glass* combines such features in its transcript, making it a good candidate to investigate these phenomena in the well-studied context of photoreceptor development and function. Here we identified an upstream overlapping open reading frame affecting Glass translation. We analysed the role of the different Glass isoforms generated by intron retention and stop codon readthrough. In addition, we dissected the *glass* regulatory sequence and identified several cell- and tissuespecific enhancer elements.

General and cell-specific enhancer elements regulate *glass* expression

The 5.2 kb region upstream of the *glass* start codon had previously been identified as the minimal sequence required for normal Glass expression [8]. The lacZ-reporter construct used in this paper also contained the upstream start codon located in intron 2. Due to the enzymatic activity of the β-galactosidase, sufficient signal was produced to detect reporter gene activity posterior of the morphogenetic furrow. However, further truncations of the upstream sequence only yielded transgenic lines with weak or variable expression or lines that did not show expression at all. Similarly, our original GFP-reporter construct showed only very weak expression levels even with the same 5.2 kb enhancer fragment. After removal of the upstream start codon, expression of our GFP-reporter construct was strongly enhanced allowing us to perform a classical enhancer bashing approach to further dissect the upstream region of *glass*. We were able to identify different enhancer regions that conferred reporter gene expression in cells posterior to the morphogenetic furrow. The retinal determination network consisting of several transcription factors, specifies the position of the eye field in many different organisms [27]. Sine oculis, a member of the retinal determination network, regulates *glass* expression by directly binding to sites in the enhancer sequence [1]. Given that all enhancer fragments we tested, showed GFP expression posterior to the morphogenetic furrow, we propose that Sine oculis binds to multiple sites in the 5.2 kb enhancer to activate *glass* expression. In addition to this general reporter gene activation we identified specific enhancer regions driving expression in distinct cell types. For example, the 5'-end of the enhancer that leads to expression in the ocelli anlage and in a subset of the photoreceptors, or the BamHI-EcoRI region that activates expression in non-photoreceptor cells. Thus, other transcription factors binding more

specifically to these enhancer elements, might regulate *glass* expression in a cell-type dependent manner.

The extended Glass PC isoform is dispensable for eye development and function

Stop codon readthrough is relatively abundant in *Drosophila* [28]. Especially genes expressed in the nervous system are putative candidates for this process [29]. Glass has also been listed as a candidate for this protein extension mechanism based on several criteria [18]: Sequence comparison of the amino acids following the regular stop codon shows a higher conservation within higher Diptera than what is found in the 3'UTR of non-readthrough transcripts. The pattern of nucleotide substitutions also suggests that there is no alteration of the reading frame as it might occur in the case of alternative splicing or ribosome hopping. The most frequent stop codon readthrough context (UGAC) [18], is also found at the Glass PA stop codon. Upon readthrough of a UGA stop codon either arginine, cysteine, serine, or tryptophan can be inserted by a near-cognate tRNA at this position [30]. In our isoform deletion experiments we did not test the conditions at which the stop codon is deleted or replaced by another codon (Glass PC and Glass PB+PC) because the function of the resulting protein might be affected by the type of alteration we introduce. Our results from the mutants expressing only the Glass PA isoform suggest, that under laboratory conditions, stop codon readthrough and production of the extended Glass PC version is not required for eye development and photoreceptor function.

Splicing of intron 4 is required to produce functional Glass protein

The Glass PB isoform alone is not functional. Our deletion mutants that can only express this truncated version as well as our other deletions that affect splicing and result in proteins terminating in intron 4, have a *glass* mutant eye phenotype, that is: they lack the expression of photoreceptor markers, show no photoreceptor activity and are photoneutral [1, 6, 31]. This corroborates previous results that showed that the last three zinc-fingers are essential for sequence specific DNA-binding and that a Glass protein lacking the C-terminal end shows no transcriptional activity [14]. The intron that is retained in the Glass RB transcript, is only found in Diptera and Lepidoptera, suggesting that it originated in the last common ancestor of flies and butterflies. Intron retention can be a means of regulating protein levels since they are usually degraded by nonsense-mediated decay [32]. One of the first examples for cell-type specific intron retention was the *Drosophila* P-element [33]. In germ cells intron 3 is spliced out resulting in functional transposase production. In contrast, in somatic cells intron 3 is retained resulting in a truncated protein that antagonizes the full-length protein. Intron retention can also generate new protein isoforms like the *Drosophila* Noble protein [34]. In addition, intron-retaining mRNA transcripts can remain in the nucleus and be spliced upon requirement providing a source of transcript that could be faster activated than by *de novo* transcription [35]. Recent RNAseq data suggests that intron 4 is not retained in the *glass* mRNA [4]. The authors found that expression of either the *glass* RA+RC or the *glass* RB transcript from transgenic constructs resulted in production of functional Glass protein, suggesting that in their ectopic expression experiments, intron 4 of the RB transcript was spliced out to produce full-length Glass PA (and PC) protein. Thus, we would consider the *glass* RB transcript as an intermediate stage that has been accumulated during cDNA preparation but that can be further processed to encode functional Glass protein. As we show here, the absence of intron 4 in the *glass* gene allowing only the production of the Glass PA (and PC) protein does not affect eye development and function.

Brs interferes with Glass translation

According to the scanning model of translation initiation [36], the 40S ribosomal subunit scans the mRNA from the 5'end until it encounters the first AUG codon. Translation will start at this codon, which in the case of *glass* mRNA would mean that only the Brs protein should be produced. However, under certain conditions, translation can also start at a later AUG codon [37]. One of these mechanisms, called leaky scanning, applies for an upstream AUG with a weak context, were the codon with the weak context is skipped by some ribosomes starting translation further downstream. However, this cannot be the case for Glass translation, since there are two AUG codons upstream of the Glass start codon within exon 2 and both have a better Kozak sequence than the actual Glass AUG. Another mechanism would be reinitiation, where after translation of a small upstream open reading frame, the ribosome can move on and re-acquire a Met-tRNA allowing it to reinitiate translation at the next AUG codon. However, since ribosomes cannot backup, the overlapping open reading frame should profoundly inhibit Glass translation [38, 39]. It was shown that overlapping upstream open reading frames; and particularly those that have an optimal AUG context, are efficiently removed from the *Drosophila* genome [40], suggesting that those that can be found and that are even conserved outside of the most closely related species, have been selected due to a specific function. In the case of Glass, we found evidence that translation from the upstream start codon strongly reduces GFP translation and also interferes with Glass translation when overexpressed. However, this suggests that the endogenous Glass protein would be expressed at very low levels. One possible way to overcome this problem, would be by splicing out exon 2 so that the two upstream AUGs and the Glass AUG would be removed from the transcript. In this case translation would start from an AUG codon in exon 3 (amino acid 26 of the predicted full-length protein). However, there is no evidence, that exon 2 is spliced out of the transcript to produce a truncated Glass version. Another hypothesis would be that Glass translation starts by reinitiation at the AUG codon in exon 3 after translation of Brs. In fact, the Glass proteins of *Anopheles darlingi* and of *Culex quinquefasciatus* are predicted to start at this position (with a conserved motif: MYISC), while *Anopheles gambiae* Glass is predicted to start in an exon located further upstream of the start codons of the other two mosquito species, but with an N-terminal sequence that is not related to that found in *Drosophila* and other higher Diptera. Also, other insects' Glass proteins start further downstream than in the Diptera. It could be possible that for most insects the actual Glass translation start has not yet been identified due to higher sequence divergence at the N-termini. This would suggest, that the first 25 amino acids mostly encoded by exon 2 of the *Drosophila* transcripts are dispensable for Glass function, or that they are only required in higher Diptera. In addition to regulating Glass protein levels by directly interfering with translation efficiency, the Brs peptide could have other functions in the developing eye, where it is expressed along with Glass. Small peptides can have important roles such as hormones, pheromones, transcriptional regulators, antibacterial peptides, etc. [41]. However, we could neither identify such a role for Brs by mutating it, nor by overexpressing it, suggesting that its main role is the regulation of Glass translation.

In summary, our results suggest that the removal of intron 4, which was added in the common ancestor of flies and butterflies, is essential for the production of a functional Glass protein. Stop codon readthrough resulting in an extension of the Glass protein that is conserved in higher Diptera seems to be dispensable for Glass function in photoreceptor development. The addition of an exon containing several AUGs upstream of the Glass start codon found in mosquitoes, can interfere with Glass translation. Nevertheless, conservation of the upstream start codon and sequence conservation of the Brs peptide suggest that higher Diptera have

found a way to overcome this interference and that Brs might even have adopted a beneficial function.

Materials and methods

Fly strains

Flies were reared at 25˚C on a cornmeal medium containing agar, fructose, molasses, and yeast. Strains for site directed integration (25709, 25710), w^{1118} mutants (3605), deficiency lines (4431, 7657) [42], balancer lines (36305, 8379), and *nos-Cas9* expressing flies (54591) [43] were obtained from the Bloomington Drosophila Stock Center. *ey-Gal4* expressing flies were a kind gift from R. Stocker.

Transgenic constructs

All oligos used for cloning and all sequencing reactions were purchased from microsynth. The following primer sequences show the annealing sequence part in capitals and any additional sequence in small letters. Restriction sites are underlined.

A 5257 basepair long PCR fragment was amplified from genomic DNA of CantonS flies using primers "glass -4301 Asc fw" (5´-ggcgcgccTAACCCGATACAAATGGAGAGG-3´) and glass 5'UTR Not re"(5´-gcggccgcGACATGACTCCACTTCTGGAAC-3´). The fragment was inserted into pCR-Blunt II-TOPO vector (Invitrogen). From there it was excised using the restriction enzymes AscI and NotI and cloned into a GFP reporter vector (pDVattBR, kind gift from Jens Rister) to generate the basic *glass*-GFP reporter construct (Fig 1A and 1B'). The plasmid was injected into y, w; attP2 embryos to produce transgenic flies (Genetic services Inc.).

To delete the two upstream start codons, a 1483 bp PCR product was amplified from the original glass-GFP reporter plasmid using primers "glass -597 fw" (5'-TAAAAACTACTGA AAACTGCTGCCGATG-3') and "glass exon2 noAUG Pme re" (5'-gcgtttaaacGATGCGT TAATTTCCAACTGCAAGGC-3'), TOPO cloned into pCRII, sequenced, digested XhoI-PmeI, and transferred into the original plasmid also cut with XhoI-PmeI, thereby removing the 104 basepairs encoding the N-terminus of Brs, and part of the multiple cloning site of the vector (Fig 1C').

To put the GFP coding sequence in frame with the upstream start codon, the GFP coding sequence was amplified by PCR using the primers "GFP noStart Not fw" (5'-tcgcggccgcgGG TGAGCAAGGGCGAGG-3') putting a NotI site in front of GFP (starting with the 3rd nucleotide of GFP) and "GFP down Fse re" (5'-GATTATGATCTAGAGTCGCGGCCG-3') covering an FseI and an XbaI site in the plasmid. The PCR product was cloned NotI-XbaI into pBluescript, sequenced, and transferred NotI-FseI into the original glass-GFP reporter plasmid deleting most of the multiple cloning site and the GFP start codon (Fig 1D').

For the enhancer analysis, the construct lacking the upstream start codons was digested with different combinations of restriction enzymes and religated. For construct B the plasmid was digested with BglII cutting in the multiple cloning site at the 5'-end of the enhancer and with BamHI cutting at position -3123. A second BamHI site at position -1885 was not in the database sequence but is present in the fragment amplified from the CantonS flies. Therefore, religation of the plasmid after BglII-BamHI digestion (the two enzymes producing compatible sticky ends) resulted in an enhancer fragment ranging from position -1885 to +886. For construct C the plasmid was digested with EcoRI cutting at the multiple cloning site at the 5'-end of the enhancer and at positions -1598 and -2040 in the enhancer. Religation resulted in an enhancer fragment ranging from position -1598 to +886. For construct D the plasmid was digested with XbaI cutting in the multiple cloning site at the 5'-end of the enhancer and at position -703. Religation resulted in an enhancer fragment ranging from -703 to +886. For

construct E, construct C was digested with XbaI cutting in the multiple cloning site at the 5'-end and at position -703 in the enhancer as well as with XhoI cutting at position -239 in the enhancer. The two ends were filled using Klenow polymerase and religated resulting in an enhancer fragment ranging from position -239 to +886. For construct F, construct B was digested with EcoRI cutting at position -1598 and XhoI cutting at position -239. The two ends were filled using Klenow polymerase and religated resulting in an enhancer fragment ranging from position -1885 to -1598 fused to the minimal promoter fragment from position -239 to +886. For construct G, the original -4301 to +886 plasmid was digested with NheI cutting at positions -1910, -1903, and -682, the site at position -2179 is missing in our enhancer fragment amplified from CantonS flies. In another reaction the original plasmid was digested with EcoRI cutting at positions -1598 and -2040. Both reactions were filled using Klenow polymerase, digested with BglII, and then the BglII-$NheI_{filled}$ fragment was ligated into the BglII-$EcoRI_{filled}$ fragment fusing the enhancer region from -4301 to -1910 to the region from -1598 to +886. For construct H, the original plasmid was digested BamHI-XhoI. The two ends were filled using Klenow polymerase and religated to fuse the fragment from -4301 to -3123 to the fragment from -239 to +886. For construct I, the original plasmid was digested with NheI and in an independent reaction with XhoI. Both digestion reactions were filled with Klenow polymerase. The $NheI_{filled}$ reaction was further digested with BamHI and the $XhoI_{filled}$ reaction was further digested with BglII. Then the BamHI-$NheI_{filled}$ fragment was ligated into the BglII-$XhoI_{filled}$ plasmid resulting in a fusion of an enhancer fragment ranging from -3123 to -1910 to the minimal promoter ranging from -239 to +886.

All constructs were injected into *nos-ΦC31;; attP2* flies for site directed integration. The G0 flies were crossed individually to w^{1118} flies to screen for w^+ offspring. w^+ F1 flies were crossed individually to 3rd chromosome balancer flies (*w;; Dr e/TM3*) and their balanced offspring was crossed *inter se* to produce stable lines.

For the UAS constructs we used the *glass* cDNA plasmid GH20219 as starting point. This cDNA still contains intron 4 due to incomplete splicing resulting in the RB transcript isoform. To remove the intron, two PCR reactions were set up. One with primers "glass 5'UTR BamHI fw" (5'-gaggatCCTCGCCAAAAGTCGCTTCTTG-3') and "glass exon4 re" (5'-ccccgactgcgaaaatCTGAGCAGGCAGAGCTTGCAC-3') resulting in a fragment ranging from the 5'-end of the 5'UTR to the end of exon 4, with the sequence given in small letters of the reverse primer overlapping with the beginning of exon 5. The other PCR reaction was done with primers "glass exon5 fw" (5'-gctctgcctgctCAGATTTTCGCAGTCGGGGAAC TTG-3') and "gl Stop Xho re" (5'-ggctcgaGTCATGTGAGCAGGCTGTTGCC-3'), resulting in a fragment ranging from the beginning of exon 5 to the PA stop codon, with the sequence given in small letters of the forward primer overlapping with the end of exon 4. Both PCR products were mixed together to provide the template for another PCR reaction with primers "gl 5'UTR BamHI fw" and "gl Stop Xho re". The resulting PCR product ranging from the 5'UTR to the PA stop codon without intron 4 was digested with BamHI-XhoI and cloned into pBluescript. After sequencing, different fragments were PCR amplified. The Glass PA coding sequence was amplified with primers "gl Start+Kozak attB1 fw" (5'-ggggacaagtttgtacaaaaaagcaggcttcaaCATGGGATTGTTATATAAGGGTTCCAAACT-3') and "gl Stop attB2 re" (5'-ggggaccactttgtacaagaaagctgggtcgTCATGTGAGCAGGCTGTTGCC-3'). The *brs* sequence was amplified with primers "glass+Smurf attB1 fw" (5'-ggggacaagtttgtacaaaaaagcaggcttcCGCATCAAGATGAAGCGTAGGAAAAGC-3') and "glass Smurf Stop attB2 re" (5'-ggggaccactttgtacaagaaagctgggtcTCAGGAGTTTGGAACCCTTATATAACAATCCC-3'). The *brs-glass-PA* sequence was amplified with primers "glass+Smurf attB1 fw" and "gl Stop attB2 re". The primer sequence in small letters are the attB parts used for gateway cloning. The PCR products were gateway cloned into pENTRY201, sequences, and transferred into

the vector *pUASg.attB* for injection (Genetic Services Inc.). *UAS-glass-PA* and *UAS-brs-Glass-PA* were injected into *nos-ΦC31; attP40* flies, while the *UAS-brs* plasmid was injected into *nos-ΦC31;; attP2* flies. After balancing the transgenic flies, *UAS-glass-RA*^{attP40} and *UAS-brs*^{attP2} were combined in a single line: *w; UAS-GlassRA*^{attP40}*; UAS-Brs*^{attP2}. The different UAS-construct bearing flies were crossed to *ey-Gal4/CyO* flies, and the number of offspring with *Cy*$^{+}$ versus the number of offspring with *CyO* wings was determined. For calculation of the survival rate the number of *Cy*$^{+}$ flies was divided by the number of *CyO* flies (Table 1).

CRISPR

For the alterations of the endogenous glass locus by CRISPR/Cas9 genome editing, we assembled the different templates in pBluescript. For the Glass PA+PC variant, we needed to remove intron 4 from the genomic DNA without changing the sequence at the Glass PA stop codon. Since there were no useful restriction sites between the intron 4 / exon 5 junction and the Glass PA stop codon, we decided to introduce an NdeI site in this sequence by altering a single nucleotide in the third position of the codon for the first histidine residue of the last zinc-finger (histidine 567 of Glass PA: CAC to CAT). We PCR amplified a 932 bp fragment from the genomic DNA of *nos-Cas9* flies using primers "glass ex5 R1 Nde fw" (5'-gagaattcatatgCGCGTCCACGGCAAC-3') and "glass 3'UTR re" (5'-GATCAAAGCACCTGTCTTACATCTACGTCTAG-3'), and a 1529 bp fragment from the intronless glass version assembled in pBluescript for generation of the *UAS-glass-PA* construct using primers "glass ex4 R1 fw" (5'-cggaattcAAGAGTGCGCCGCTTCC-3') and "glass ex5 Nde re" (5'-CGCATatgCCGATTCAAGTTCCCCGAC-3'). Both PCR products were combined in pBluescript vector by digesting the one covering the C-terminus from the NdeI site introduced in exon 5 to an endogenous HindIII site in the 5'UTR with NdeI-HindIII, and the one covering exon 4 and part of exon 5 with EcoR-I-NdeI (due to an endogenous NdeI site in the middle of exon4 this part was cloned in two steps). The sequence of the resulting fragment ranging from the beginning of exon 4 to the 5'UTR and lacking intron 4 was confirmed.

For the Glass PB variant, we introduced a deletion ranging from the end of intron 4 to the middle of the sequence added in the extended Glass PC protein version (Fig 4A). We amplified two PCR products from the genomic DNA of *nos-Cas9* flies. A 1866 bp fragment spanning exon 4 and most of intron 4 was amplified with primers "glass ex4 R1 fw" and "glass int4 Xho re" (5'-AActcgagGTATAACGTTCCAGGACTGCTC-3'), and a 1287 bp fragment ranging from the middle of the Glass PC encoding sequence to a place located around 500 bp downstream of the *glass* gene was amplified with primers "glass 3'UTR Xho fw" (5'-aactcgagCATCGGCGATTATACTCCACC-3') and "glass down Kpn re" (5'-agggtaccTTTATGGTGGCCTCCCAGG-3'). Both fragments were subcloned, sequenced and combined in pBluescript by digesting the one covering exon 4 and intron 4 with EcoRI-XhoI, and the one covering part of the PC coding region, the 3'UTR and donwnstream genomic sequence with XhoI-Acc65I.

For the Glass PA+PB variant, we introduced two additional stop codons at the end of the Glass PA encoding part and deleted 29 of the nucleotides following the stop codon. We PCR amplified two PCR products from the genomic DNA of *nos-Cas9* flies. A 2034 bp fragment spanning exon 4, intron 4, and the PA encoding part of exon 5 was amplified with primers "glass ex4 fw" and "glass RA **3xStop** Xho re" (5'-ctctcgag**ctattatca**TGTGAGCAGGCTGTTGCCAC-3'). A 1365 bp fragment spanning most of the Glass RC specific sequence, the 3'UTR and around 500 nucleotides of downstream genomic sequence was amplified using primers "glass RC Xho fw" (5'- ttctcgAGCATTACCACCCCCGC-3') and "glass down Kpn re". Both fragments were subcloned, sequenced, and combined in pBluescript by digesting the one covering exon 4, intron 4, and the Glass PA encoding part plus stop codons with EcoRI-XhoI,

and the one covering the Glass PC region, 3'UTR, and downstream genomic sequence XhoI-Acc65I.

For the Glass PA variant, we performed a PCR amplification of exon 4 and the 5'-end of exon 5 with the same first primer pair as for the Glass PA+PB template, but using the intronless glass version assembled in pBluescript for generation of the *UAS-glass-PA* construct as a template. We subcloned this 1638 bp fragment EcoRI-XhoI, sequenced it, and combined it with the 1365 bp fragment spanning most of the Glass RC specific sequence, the 3'UTR and around 500 nucleotides of downstream genomic sequence.

For expression of the CRISPR guideRNAs, sense and antisense oligos with overhangs fitting the sticky ends of the BbsI digested vector were annealed and ligated into pU6-BbsI-chiRNA plasmid (a gift from Melissa Harrison & Kate O'Connor-Giles & Jill Wildonger, Addgene plasmid # 45946 [44]). Site 1 (ctgctcaggtgagtccg/gga) is located at the junction between exon 4 and intron 4. Site 2 (gtccacagattttcgca/gtc) is located at the junction between intron 4 and exon 5. Site 3 (agg/agtgcaggaggtttcca) guides Cas9 to cut 12 bp downstream of the Glass PA stop codon. Site 4 (aca/tgggtaactacgactac) is located in the middle of the Glass PC encoding region (Fig 4A). CRISPR sites were selected based on their position in the glass genomic sequence using a CRISPR site prediction program (http://tools.flycrispr.molbio.wisc.edu/targetFinder/ [45]).

The templates for the different Glass isoform variants were co-injected with sgRNA expression plasmids into embryos of *nos-Cas9* flies. The Glass PA+PC variant was co-injected with the sgRNA plasmids for sites 1 and 2. The template for the Glass PB variant was co-injected with sgRNA plasmids for sites 2 and 4. The template for the Glass PA+PB variant was co-injected with the sgRNA plasmid for site 3. The template for the Glass PA variant was co-injected with the sgRNA plasmids for sites 1 and 3. The resulting G0 flies were crossed individually with deficiency lines uncovering the *glass* locus. Some of the offspring resulting from Cas9 cutting at CRISPR site 1 (and 3) showed a *glass* mutant phenotype due to CRISPR induced non-homologous end joining (S7 Fig). Irrespective of the eye phenotype the offspring was crossed individually with third chromosome balancer flies (*w;; Dr, e/TM3*) and analyzed by PCR for the introduced deletions and sequence alterations. The *glass* genes of those lines that showed changes in the PCR analyses, were sequenced to confirm the introduced changes and identify other alterations that resulted in *glass* mutant phenotypes.

For the deletions in the *brs* sequence, template sequences were assembled in pBluescript. For the 1nt deletion two PCR products were amplified. A 2759 bp fragment ranging from position -1863 in the glass enhancer region to position +896 in the *brs* coding sequence was amplified using primers "glass -2700 Kpn fw" (5'-tgggtaccGGCAGCAGAGACAGGCTC-3') and "smurf -1nt H3 re" (5'-ccaagcttCATCTTGATGCGTTAATTTCCAACTGC-3'). The resulting PCR product was cloned Acc65I-HindIII into pBluescript and sequenced. A 2493 bp fragment ranging from position +899 in the *brs* coding sequence to position +3392 at the end of exon 4 was amplified from *nos-Cas9* genomic DNA using primers "smurf -1nt H3 fw" (5'-gaaagcttG GAAAAGCAGGAACAAATGCGCG-3') and "glass exon4 Pst re" (5'-ctctgcagGCAGAGCT TGCACTGG-3'). The resulting PCR product was cloned HindIII-PstI into pBluescript, sequenced, and combined with the 5'-fragment. For the 5 nt deletion fusing Brs with Glass, A 2821 bp fragment ranging from position -1863 to +953 in the *brs* coding sequence just before the second AUG codon was amplified from *nos-Cas9* genomic DNA using primers "glass -2700 Kpn fw" and "smurf -5nt BamH Nco re" (5'-ccggatcccatggCTCCACTTCTGGAACGTT TGGGC-3'). The resulting PCR product was cloned Acc65I-BamHI into pBluescript and sequenced. A 1632 bp fragment ranging from position +959 just before the Glass start codon to position +2591 in exon 4 was amplified from *nos-Cas9* genomic DNA using primers "smurf -5nt Nco fw" (5'-ctccatggGATTGTTATATAAGGGTTCCAAACTCCTG-3') and "glass ex4

Not re" (5'-aagcggccgcatggtgcatggtcatgttcatgc-3'), cloned $NcoI_{filled}$-NotI into pBluescript $BamHI_{filled}$-NotI, sequenced, and combined KpnI-NcoI with the 5'-fragment. For expression of the CRISPR guideRNAs, sense and antisense oligos with overhangs fitting the sticky ends of the BbsI digested vector were annealed and ligated into the BbsI digested pCDF4-U6:1_U6:3tandemgRNAs plasmid (gift from Simon Bullock, addgene plasmid # 49411). The site for the -1nt deletion (aacgcatcaagatgaag/cgt) is located at the upstream start codon, the site for the -5nt deletion (ccagaagtggagtcatg/tca) is located at the Glass start codon (Fig 3A). The templates for the *brs* mutations were co-injected with the corresponding sgRNA expression plasmid into embryos of *nos-Cas9* flies. The resulting G0 flies were crossed individually with deficiency lines uncovering the glass locus. Some of the F1 flies had a very subtle rough eye phenotype over the glass deficiency–but also over the TM6b balancer. The F1 flies were crossed individually with 3^{rd} chromosome balancer flies to establish stocks, and tested by PCR and restriction digest with either HindIII or NcoI for presence of the introduced changes. Genomic DNA of the *glass* locus from homozygous candidates was PCR amplified and sent for sequencing. Additional sequence changes due to non-homologous end joining were also identified (Fig 3A(d-f)).

Transcription factor binding site scanning

We used FIMO from the meme suite version 4.9.1 [46, 47] with the standard parameter. *Drosophila* only Position Weight Matrices (PWMs) were collected and carefully curated from flyfactorsurvey [48], Jaspar [49] and Transfac [50].

Quantitative PCR

Eye-antennal-discs were dissected from 25–30 larvae per genotype and experiment. Total RNA was extracted using Trizol (QIAzol, QIAGEN) and chloroform. After precipitation the RNA pellets were resuspeded in nuclease-free water and the RNA content determined using NanoDrop (Thermo Fisher Scientific). cDNA was generated using the GoScript reverse transcription system (Promega) with oligo(dT) primers. qPCR was performed on a Rotor Gene Q cycler (QIAGEN) using the KAPA SYBR FAST qPCR Kit (KAPA biosystems). The following primer combinations were used:

- "act42c qP fw" (GCAGCGGATAACTAGAAACTACTCC) located at the end of the first exon of the *Drosophila actin42c* gene, and "act42c qP re" (CGACCACTAAAGCTGCAA CCTC) located at the beginning of the second exon of the *Drosophila actin42c* gene as housekeeping gene.
- "gl qP ex4 fw" (TCACCAAGCACCTGCGAATC) located at the end of exon 4 of the *Drosophila glass* gene, and "gl qP ex5 re" (CGCATGTGCCGATTCAAGTTCC) located at the beginning of exon 5, to detect spliced *glass* cDNA.
- "gl qP ex4 fw" and "gl qP int4 re" (AGGACGGTGTAAGTAAATCCCCC) located at the beginning of intron 4, to detect *glass* cDNA that still contains intron 4.

qPCR Primers were synthetized and HPLC purified at microsynth. All qPCR reactions were set up as duplicates and three biological replicates were performed for each genotype. The thresholds were set manually and the resulting Ct values were analysed using Microsoft Excel to calculate average, standard deviation and statistical significance (student's t-test).

Light preference assay

The light preference assays were prepared and performed under red light conditions during the subjective day. *Drosophila melanogaster* adults of both sexes were used. Given that light

perception in *Drosophila* is affected by age [6], for consistency, we used <1 day old flies in all our experiments. Without anaesthesia 20–40 flies were taken from food vials and loaded into the elevator chamber of a T-maze. The elevator chamber was descended, and flies were allowed to move freely between the elevator chamber and two plastic tubes for 2 min. A white LED was placed at one end of one testing tube. Thus, only one testing tube was illuminated. After 2 min the elevator chamber was ascended, and flies were not able to move between testing tubes and elevator chamber. We determined the number of flies in the illuminated testing tube (L) and the number of flies in the dark testing tube (D) as well as the number of flies in the elevator chamber (E). We calculated a preference index as follows:

$$\text{Preference index} = (L - D)/(L + D + E)$$

Light intensity was measured from the distance of the elevator chamber to the LED. The light intensity was 1338 μW/cm^2 with a first maximum intensity peak of 16.6 μW/cm^2/nm at 443 nm with half-widths of around 11 nm and a second maximum intensity peak of 6.8 μW/cm^2/nm at 545 nm with half-widths of around 62 nm.

Electroretinogram

ERG recordings were obtained as previously described [51]. Briefly, we mounted living flies inside a pipette tip, leaving their heads outside, and immobilised them with a mixture of bee wax and colophony 3:1, which worked as a glue. We placed the flies inside a dark chamber and applied two electrodes: a ground electrode was positioned inside the head of the fly, and a recording electrode was introduced into the retina. In our stimulation protocol, we illuminated the compound eye with orange light for 5 seconds to transform all metarhodopsin to rhodopsin, switched off the light for 10 seconds, and illuminated again a second time with orange light for another 5 seconds. We recorded the response of photoreceptors to the second stimulus.

Immunohistochemistry

Eye imaginal discs were dissected from third instar larvae and fixed in 3.7% formaldehyde dissolved in phosphate buffer (PB) for 20 minutes. Bolwig organs from third instar larvae and ocelli and eyelets in the adult brain were dissected in PB and fixed in 3.7% formaldehyde prepared in PB for 25 minutes.

For cryosections, we dissected the heads of the flies and fixed them for 20 minutes with 3.7% formaldehyde dissolved PB, as previously described [1]. We washed these samples by using phosphate buffer with Triton 0.3% (PBT) and incubated them overnight in cryoprotected solution (25% sucrose in PB). Then, we embedded the heads in OCT, froze them, and took 14 μm sections by using a cryostat.

All tissue samples (eye imaginal discs, bolwig organs, adult brains with ocelli, eyelets and cryosections) were washed with PBT at least 3–4 times and incubated sequentially in primary and secondary antibodies (each antibody incubation step was performed overnight, washing with PBT (3–4 times) between and in the end of these steps). We used Vectashield as a mounting medium.

As primary antibodies, we used rabbit anti-GFP (1:1000, Molecular probes, A-6455), rabbit anti-Rh2 (1:100) [24], rabbit anti-Rh6 (1:10,000) [52], rabbit anti-Sal (1:500) [53], guinea pig anti-Kr (1:200, a gift from J. Jaeger), mouse anti-Rh5 (1:50) [54] and guinea pig anti-Sens (1:800, courtesy of H. Bellen [55]). To stain against phototransduction proteins we used antibodies generated in C. Zuker's lab: anti-NorpA (1:100), rabbit anti-Trpl (1:100), both of which were kindly provided by N. Colley. The following antibodies were obtained from Developmental Studies Hybridoma Bank (DSHB) at The University of Iowa: mouse anti-Rh1 (1:20, 4C5), mouse anti-Trp (1:20, MAb83F6), mouse anti-FasII (1:20, No. 1D4), mouse anti-Chp (1:20, No. 24B10) and rat anti-Elav (1:30, No. 7E8A10).

Secondary antibodies were obtained from Molecular probes, and we used the conjugated with the following Alexa fluor proteins: 488, 568, and 647. In addition, we also used Hoechst 33258 (1:100, Sigma, No. 94403).

Supporting information

S1 Fig. *glass*-GFP reporter gene expression patterns. GFP (green), Senseless (Sens, red), and Elav (blue) expression in the eye region of larval imaginal discs of transgenic flies expressing construct C (-1598 to +886) (A), construct F (-1885 to -1598 / -239 to +886) (B), construct H (-4301 to -3123 / -239 to +886) (C), or construct I (-3123 to -1906 / -239 to +886) (D) (compare to Fig 2A for the individual constructs). Scale bar: 40 μm A' to D': magnification of areas in panels A to D. Scale bar: 5 μm. A" to D": GFP channel alone.
(TIF)

S2 Fig. Potential transcription factor binding sites in the *glass* regulatory region. A: The 5.2 kb *glass* regulatory region contains different enhancer elements: an ocelli enhancer that also drives expression in a subset of photoreceptors (-4301 to -3123; purple), two photoreceptor enhancers (-3123 to -1886 and -1598 to -239; orange), a non-photoreceptor enhancer (-1886 to -1598, green), a repressor element (-1598 to -703; blue) and a promoter region (-239 to +886; yellow). There are 31 potential binding sites for the transcription factor Sine oculis within the entire 5.2 kb upstream genomic region; 10 sites with the consensus AGATAC (red bars) and 21 sites with the consensus YGATAY (green bars). B: *in silico* analysis of potential transcription factor binding sites. Binding sites are spread along the entire 5.2 kb region (coloured bars) with several clusters in the different enhancer fragments.
(TIF)

S3 Fig. Sequence conservation upstream of the *glass* start codon. A: nucleotide sequence of glass exon 2 of different higher *Diptera*. The position of the glass start codon is outlined in green. The position of the upstream start codon is outlined in blue. The black vertical lines show the triplets following the upstream start codon that result in the amino acid sequences shown in B. B: Alignment of the amino acid sequences resulting from translation beginning at the start codon upstream of Glass of different higher Diptera. *Drosophila melanogaster (D mel)*, *Drosophila sechellia (D sec)*, *Drosophila simulans (D sim)*, *Drosophila erecta (D ere)*, *Drosophila yakuba (D yak)*, *Drosophila takahashii (D tak)*, *Drosophila eugracilis (D eug)*, *Drosophila rhopaloa (D rho)*, *Drosophila ficusphila (D fic)*, *Drosophila elegans (D ele)*, *Drosophila ananassae (D ana)*, *Drosophila persimilis (D per)*, *Drosophila pseudoobscura (D pse)*, *Drosophila grimshawi (D gri)*, *Drosophila virilis (D vir)*, *Drosophila mojavensis (D moj)*, *Drosophila willistoni (D wil)*, *Lucilia cuprina (L cup)*, *Musca domestica (M dom)*, *Glossina morsitans (G mor)*.
(TIF)

S4 Fig. qPCR analysis of transcript levels. A: wildtype and mutated versions of *glass* from the end of exon 4 to the end of the transcript. The C2H2-zinc-finger region is shown in purple and magenta. Intron 4 is shown in white. The C-terminus of the PA isoform is shown in blue, that of the PC isoform in yellow. The 3'UTR is grey. Deletions are indicated as black lines. The triple stop codon introduced in the PA+PB and the PA alleles are indicated by red boxes. The positions of the primers used for qPCR mutagenesis are shown as arrows. Black arrow: primer "gl qP ex4 fw"; orange arrow primer "gl qP int4 re"; blue arrow: primer "gl qP ex5 re". B: Relative amounts of spliced (blue) and unspliced (orange) *glass* transcripts in different CRISPR generated lines. Paired samples two-tailed *t* test: For all data sets $n = 3$ experiments. brs^{-1nt}: $p = 0.268$; $brs^{-5nt::Glass}$: $p = 0.193$; PA+PC spliced: $p = 0.345$; PA+PB spliced: $p = 0.026$; PA spliced: $p = 0.265$; PA+PB unspliced: $p = 0.046$; PB unspliced: p = 0.004. Data show mean and

error bars show standard deviation. ns = not significant; * = $p<0.05$; ** = $p<0.01$.
(TIF)

S5 Fig. Sequence conservation of intron 4 and Glass PC. A: nucleotide sequence of the transition between exon 4 and intron 4 and between intron 4 and exon 5 of *Drosophila melanogaster*. Other Diptera and Lepidoptera also contain this intron followed by a stop codon black box) immediately after the exon intron junction (black line), except in *Heliconius*, where the stop codon is located 17 pb downstream of the exon intron junction. Although the intron is absent in other insect species, the amino acid sequence flanking the intron and forming part of the Glass zinc-finger is highly conserved as indicated by the translation below the alignment. B: amino acid alignment of the Glass C-termini of different higher Diptera. The position of the Glass PA stop codon is marked by an asterisk (arrow). The amino acid sequence directly following the end of the PA isoform is highly conserved. There is also high sequence conservation at the C-terminus of the PC isoform. The central region, which is rich in histidine residues, is more variable. *Drosophila melanogaster (D mel)*, *Drosophila simulans (D sim)*, *Drosophila sechellia (D sec)*, *Drosophila erecta (D ere)*, *Drosophila yakuba (D yak)*, *Drosophila ananassae (D ana)*, *Drosophila pseudoobscura (D pse)*, *Drosophila persimilis (D per)*, *Drosophila willistoni (D wil)*, *Drosophila virilis (D vir)*, *Drosophila mojavensis (D moj)*, *Drosophila grimshawi (D gri)*, *Musca domestica (M dom)*, *Glossina morsitans (G mor)*, *Lucilia cuprina (L cup)*, *Anopheles darlingi (A dar)*, *Anopheles gambiae (A gam)*, *Culex quinquefasciatus (C qui)*, *Culex pipiens (C pip)*, *Danaus plexipus (D ple)*, *Heliconius melpomene (H mel)*, *Apis melifera (A mel)*, *Bombus impatiens (B imp)*, *Nasonia vitripennis (N vit)*, *Tribolium castaneum (T cas)*, *Pediculus humanus (P hum)*, *Acyrthosiphon pisum (A pim)*, *Ixodes scapularis (I sca)*, *Drosophila ficusphila (D fic)*, *Drosophila eugracilis (D eug)*, *Drosophila biarmipes (D bia)*, *Drosophila takahashii (D tak)*, *Drosophila elegans (D ele)*, *Drosophila bipectinata (D bip)*, *Drosophila kikkawai (D kik)*. C: WebLogo image of the 5' end of the PC specific coding region (first 18 codons from the PA stop codon). Highest nucleotide variability is found in the third positions (9, 12, 15, etc.) suggesting that the reading frame has not been shifted from the end of the PA sequence.
(TIF)

S6 Fig. Glass isoform function in other visual systems. A-R: antibody staining for photoreceptor markers in ocelli, eyelet, and Bolwig's organ of w^{1118} and different isoform lines as indicated on the left. Scale bars represent 10μm. A, D, G, J: ocelli stained for Rhodopsin2 (Rh2, blue), Chaoptin (Chp, green), and Elav (red) which also stains neurons of the brain located underneath the ocelli. B, E, H, K: eyelet photoreceptors stained for Chp (green) and Elav (red), which also stains neurons in the lamina (row of cells in the left part of the panels). C, F, I, L, O, R: Bolwig organ photoreceptors stained for Rhodopsin 5 (Rh5, blue), Rhodopsin 6 (Rh6, green), and Elav (red). The markers are expressed in flies expressing the Glass isoforms PA +PC, PA, and PA+PB, but not in flies expressing only the PB isoform indicating that in this case the cells are not fully differentiated or are missing completely. M, P: Bolwig organ precursors are detectable at embryonic stage 12 due to their expression of Krüppel (Kr, green), Fasciclin II (FasII, red), and Spalt (Sal, blue, arrowheads), which is only expressed in the four primary photoreceptor precursors giving rise to the Rh5 expressing photoreceptors. The precursor cells are present in embryos expressing only the Glass PB isoform, but only two of them express the markers Kr and Sal (arrowheads). N, Q: Bolwig organs stained for Sal (blue), Kr (green) and Elav (red). In wildtype Kr and Elav are expressed in all photoreceptors, while Sal is a marker for the four photoreceptors expressing Rh5. In larvae that have only the Glass PB isoform, none of the markers are expressed.
(TIF)

S7 Fig. Additional *glass* alleles generated by non-homologous end joining. A: wildtype and mutated versions of *glass* from the end of exon 4 to the end of the transcript. Stop codons of isoforms PA, PB, and PC are indicated by asterisks. The C2H2-zinc-finger region is shown in magenta. Intron 4 is shown in white. The C-terminus of the PA isoform is shown in blue, that of the PC isoform in yellow. The 3'UTR is grey. The sequences at the exon intron and intron exon junctions are given as letters with a black vertical line depicting the position of the junction. The stop codon at the beginning of exon 4 is highlighted in blue. The positions of the CRISPR sites used for mutagenesis are underlined in green. Deletions are indicated as black lines in the schemes, and as red letters in the sequences. Due to the deletions of the exon intron junction and the stop codons, additional amino acids are added to the Glass PB sequence until they reach the next stop codon in intron 4 (purple boxes). Due to the frameshift caused by the single nucleotide deletion in exon 5, the amino acid sequence of the Glass PA isoform gets shifted in $gl^{del21.3}$ (orange box). B-E: Adult eye phenotype and expression of the retinal markers Rh1, NorpA, Trp, and Trpl of the $gl^{del1.1}$ mutant (B), of the $gl^{del22.4}$ mutant (C), of the $gl^{del31.9}$ mutant (D), of the $gl^{del21.3}$ mutant (E). All antibody stainings are shown in green, counterstaining of DNA with Hoechst (magenta). None of the tested photoreceptor makers is expressed in these glass alleles. Scale bars: 40 μm F: ERGs of the different deletion alleles show no response to light; scale bars represent 5 mV (vertical) and 5 seconds (horizontal). G: Flies homozygous for the different small deletions in the *glass* locus are photoneutral. Two-tailed one sample *t* test followed by the Benjamini Hochberg procedure: For all data sets n = 10 experiments. $gl^{del1.1}$: $p = 0.90$, $t_{(9)} = 0.13$; $gl^{del21.3}$: $p = 0.53$, $t_{(9)} = -1.11$; $gl^{del22.4}$: $p = 0.49$, $t_{(9)} = -1.03$; $gl^{del31.9}$: $p = 0.66$, $t_{(9)} = -0.57$. The light preference index of all experimental groups is different from the light preference index of the control group shown in Fig 3. One-way ANOVA of preference indices: For all data sets n = 10 experiments, $p<2^{*}10^{-16}$, $F\text{-}Value_{(8,81)} = 44.04$. CTRL vs $gl^{del1.1}$: $p<0.001$, $t = 8.74$; CTRL vs $gl^{del21.3}$: $p<0.001$, $t = 9.81$; CTRL vs $gl^{del22.4}$: $p<0.001$, $t = 10.12$; CTRL vs $gl^{del31.9}$: $p<0.001$, $t = 9.20$. Data show mean and error bars show standard deviation. Red dots indicate means of individual experiments. ns = not significant.
(TIF)

S1 Table. Transcription factor binding sites in the 5.2 kb upstream region of glass. Results of the *in silico* analysis for transcription factor binding sites upstream of *glass* ranked according to their p-values. Highlighted in red are factors that are reported to function in the eye. Factors that bind only to one single site in the entire 5.2 kb region are highlighted depending on the enhancer region they bind to: ocelli enhancer in purple; photoreceptor enhancers in orange; non-photoreceptor enhancer in green; repressor region in blue; promoter in yellow.
(XLSM)

S2 Table. qPCR results and analysis. Ct values of spliced and unspliced *glass* transcript and of *actin42c* transcript of three independent qPCR experiments (with 2 replicates, each). Comparative quantification and statistical analysis.
(XLSX)

S3 Table. Light preference assay results and calculation of preference indices.
(XLSX)

Acknowledgments

We thank J. Rister and S.Bullock for plasmids; M.Harrison, K. O'Connor-Giles, and J.Wildonger for the pU6-BbsI-gRNA plasmid and the CRISPR target finder website; R. Stocker for the *ey-Gal4* flies; and N. Colley, H. Bellen, J.Jaeger and the DSHB for antibodies. We thank J. E.

Treisman for helpful comments and discussions, Désirée König for help with qPCR, and Laurent Falquet for help with DNA sequence analysis. We are also grateful to our colleagues from the Sprecher and Huber labs for valuable discussions, particularly thanks to B. Egger, J. Kaldun, T. Smylla and O. Voolstra.

Author Contributions

Conceptualization: Cornelia Fritsch, F. Javier Bernardo-Garcia, Armin Huber, Simon G. Sprecher.

Data curation: Cornelia Fritsch, F. Javier Bernardo-Garcia, Tim-Henning Humberg.

Formal analysis: Cornelia Fritsch, F. Javier Bernardo-Garcia, Armin Huber.

Funding acquisition: Simon G. Sprecher.

Investigation: Cornelia Fritsch, F. Javier Bernardo-Garcia, Tim-Henning Humberg, Abhishek Kumar Mishra, Sara Miellet, Silvia Almeida, Michael V. Frochaux, Bart Deplancke.

Project administration: Armin Huber, Simon G. Sprecher.

Supervision: Simon G. Sprecher.

Visualization: Cornelia Fritsch, F. Javier Bernardo-Garcia, Tim-Henning Humberg, Abhishek Kumar Mishra, Simon G. Sprecher.

Writing – original draft: Cornelia Fritsch, F. Javier Bernardo-Garcia, Simon G. Sprecher.

Writing – review & editing: Cornelia Fritsch, Simon G. Sprecher.

References

1. Bernardo-Garcia FJ, Fritsch C., Sprecher S.G. The transcription factor Glass links eye field specification with photoreceptor differentiation in Drosophila. Development. 2016 143:1413–23. https://doi.org/10.1242/dev.128801 PMID: 26952983
2. Liang X, Mahato S., Hemmerich C., Zelhof A.C. Two temporal functions of Glass: Ommatidium patterning and photoreceptor differentiation. Dev Biol. 2016; 414(1):4–20. https://doi.org/10.1016/j.ydbio.2016.04.012 PMID: 27105580
3. Moses K, Ellis M.C., Rubin G.M. The glass gene encodes a zinc-finger protein required by Drosophila photoreceptor cells. Nature. 1989; 340:531–6. https://doi.org/10.1038/340531a0 PMID: 2770860
4. Morrison CA, Chen H., Cook T., Brown S., Treisman J.E. Glass promotes the differentiation of neuronal and non-neuronal cell types in the Drosophila eye. PLoS Genet. 2018; 2018(14):1.
5. Bridges CB, Morgan T.H. The third-chromosome group of mutant characters of Drosophila melanogaster. Washington: Carnegie Institution of Washington 1923; 1923.
6. Bernardo-Garcia FJ, Humberg T.H., Fritsch C., Sprecher S.G. Successive requirement of Glass and Hazy for photoreceprot specification and maintenance in Drosophila. Fly (Austin). 2017; 11(2):112–20. https://doi.org/10.1080/19336934.2016.1244591 PMID: 27723419
7. Moses K, Rubin G.M. glass encodes a site-specific DNA-binding protein that is regulated in response to positional signals in the developing Drosophila eye. Genes & Development. 1991; 5(4):583–93.
8. Liu H. M C, Moses K. Identification and functional characterization of conserved promoter elements from glass: a retinal development gene of Drosophila. Mech Dev. 1996; 56(1–2):73–82. PMID: 8798148
9. Kosugi s, Haseve M., Tomita M., Yanagawa H. Systematic identification of yeast cell cycle-dependent nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proc Natl Acad Sci USA. 2009; 106(10171–10176). https://doi.org/10.1073/pnas.0900604106 PMID: 19520826
10. Ready DF, Hanson T.E., Benzer S. Development of the Drosophila retina, a neurocrystalline lattice. Dev Biol. 1976; 53(2):217–40. PMID: 825400
11. Jusiak B, Karandikar U.C., Kwan S.J., Wang F., Wang H., Chen R., Mardon G. Regulation of Drosophila eye development by the transcription factor Sine oculis. PLoS One. 2014; 9(2). https://doi.org/10.1371/journal.pone.0089695 PMID: 24586968

12. Jemc J, Rebay I. Identification of transcriptional targets of the dual-function transcription factor/phosphatase eyes absent. Dev Biol. 2007; 310(2):416–29. https://doi.org/10.1016/j.ydbio.2007.07.024 PMID: 17714699

13. Cavener DR. Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. Nucleic Acid Res. 1987; 15(4):1353–61. https://doi.org/10.1093/nar/15.4.1353 PMID: 3822832

14. O'Neill EM, Ellis M.C., Rubin G.M., Tjian R. Functional domain analysis of glass, a zinc-finger-containing transcription factor in Drosophila.. Proc Natl Acad Sci U S A. 1995; 92(14):6557–61. https://doi.org/10.1073/pnas.92.14.6557 PMID: 7604032

15. Naval-Sanchez M, Potier D., Haagen L., Sanchez M., Munck S., Van de Sande B., Casares F., Chrisiaens V., Aerts S. Comparative motif discovery combined with comparative transcriptomics yields accurate targetome and enhancer predictions. Genome Res. 2012; 23(1):74–88. https://doi.org/10.1101/gr.140426.112 PMID: 23070853

16. FlyBase. FlyBase inference based on genome sequence analysis. 1996.

17. Annotators FG. Changes affecting gene model number or type in release 3.2 of the annotated D.melanogaser genome. 2004.

18. Jungreis I, Lin M.F., Spokony R., Chan C.S., Negre N., Victorsen A., White K.P., Kellis M. Evidence of abundant stop codon readthrough in Drosophila and other metazoa. Genome Res. 2011; 21(12):2096–113. https://doi.org/10.1101/gr.119974.110 PMID: 21994247

19. Wang L, Park H.J., Dasari S., Wang S., Kocher J.P., Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acid Res. 2013; 41(6). https://doi.org/10.1093/nar/gkt006 PMID: 23335781

20. Crooks GE, Hon G., Chandonia J.M., Brenner S.E. WebLogo: A sequence logo generator. Genome Res. 2004; 14(6):1188–90. https://doi.org/10.1101/gr.849004 PMID: 15173120

21. Schneider TD, Stephens R.M. Sequence Logos: A New Way to Display Consensus Sequences. Nucleic Acid Res. 1990; 18(20):6097–100. https://doi.org/10.1093/nar/18.20.6097 PMID: 2172928

22. Belusic G. ERG in Drosophila2011.

23. Pollock JA, Benzer S. Transcript localization of four opsin genes in the three visual organs of Drosophila; RH2 is ocellus specific. Nature. 1988; 333(6175):779–82. https://doi.org/10.1038/333779a0 PMID: 2968518

24. Mishra AK, Bargmann B.O.R., Tsacjaki M., Fritsch C., Sprecher S.G. Functional genomics identifies regulators of the phototransduction machinery in the Drosophila larval eye and adult ocelli. Dev Biol. 2016; 410(2):164–77. https://doi.org/10.1016/j.ydbio.2015.12.026 PMID: 26769100

25. Helfrich-Forster C, Edwards T., Yasuyama K., Wisotzky B., Schneuwly S., Stanewsly R., Meinertzhagen I.A., Hofbauer A. The extraretinal eyelet of Drosophila: development, ultrastructure, and putative circadian function. J Neuroscience. 2002; 22(21):2955–66.

26. Sprecher SG, Pichaud F., Desplan C. Adult and larval photoreceptors use different mechanisms to specify the same rhodopsin fates. Genes & Development. 2007; 21:2182–95. https://doi.org/10.1101/gad.1565407 PMID: 17785526

27. Silver S.J RI. Signaling circuitries in development: insights from the retinal determination gene network. Development. 2005; 132(1):3–13. https://doi.org/10.1242/dev.01539 PMID: 15590745

28. Jungreis I, Chan C.S., Waterhouse R.M., Fields G., Lin M.F., Kellis M. Evolutionary Dynamics of Abundant Stop Codon Readthrough. Mol Biol Evol. 2016; 33(12):3108–32. https://doi.org/10.1093/molbev/msw189 PMID: 27604222

29. Lin MF, Carlson J.W., Crosby M.A., Matthews B.B., Yu C., Park S., Wan K.H., Schroeder A.J., Gramates L.S., St Pierre S.E., Roark M., Wiley K.L Jr., Kulathinal R.J., Zhang P., Myrick K.V., Antone J.V., Celniker S.E., Gelbart W.M., Kellis M. Revisiting the protein-coding gene catalog of Drosophila melanogaster using 12 fly genomes. Genome Res. 2007; 17(12):1823–36. https://doi.org/10.1101/gr.6679507 PMID: 17989253

30. Chittum HS, Lane W.S., Carlson B.A., Roller P.P., Lung F.D., Lee B.J., Hatfield D.L. Rabbit beta-globin is extended beyond its UGA stop codon by multiple suppressions and translational reading gaps. Biochemistry. 1998; 37(31):10866–70. https://doi.org/10.1021/bi981042r PMID: 9692979

31. Liang X, Mahato S, Hemmerich C, Zelhof AC. Two temporal functions of Glass: ommatidium patterning and photoreceptor differentiation. Developmental biology. 2016; 414(1):4–20. Epub 2016/04/24. https://doi.org/10.1016/j.ydbio.2016.04.012 PMID: 27105580;

32. Lykke-Andersen S, Jensen T.H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol. 2015; 16(11):665–77. https://doi.org/10.1038/nrm4063 PMID: 26397022

33. Rio DC, Laski F.A., Rubin G.M. Identification and immunochemical analysis of biologically active Drosophila P element transposase. Cell. 1986; 44(1):21–32. https://doi.org/10.1016/0092-8674(86)90481-2 PMID: 2416475

34. Gontijo AM, Miguela V., Whiting M.F., Woodruff R.C., Dominguez M. Intron retention in the Drosophila melanogaster Rieske Iron Sulphur Protein gene generated a new protein. Nat Commun. 2011; 2(323). https://doi.org/10.1038/ncomms1328 PMID: 21610726

35. Wong JJ, Au A.Y., Ritchie W., Rasko J.E. Intron retention in mRNA: No longer nonsense: Known and putative roles of intron retention in normal and disease biology. Bioessays. 2016; 38(1):41–9. https://doi.org/10.1002/bies.201500117 PMID: 26612485

36. Kozak M. Pushing the limits of the scanning mechanism for initiation of translation. Gene. 2002; 299:1–34. https://doi.org/10.1016/s0378-1119(02)01056-9 PMID: 12459250

37. Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene. 2005; 361:13–37. https://doi.org/10.1016/j.gene.2005.06.037 PMID: 16213112

38. McGeachy AM, Ingolia N.T. Starting too soon: upstream reading frames repress downstream translation. EMBO J. 2016; 35(7):699–700. https://doi.org/10.15252/embj.201693946 PMID: 26896443

39. Johnstone TG, Bazzini A.A., Giraldez A.J. Upstream ORFs are prevalent translational repressors in vertebrates. EMBO J. 2016; 35(7):706–23. https://doi.org/10.15252/embj.201592759 PMID: 26896445

40. Hsu MK, Chen F.C. Selective constraint on the upstream open reading frames that overlap with coding sequences in animals. PLoS One. 2012; 7(11). https://doi.org/10.1371/journal.pone.0048413 PMID: 23133632

41. Ladoukakis E, Pereira V., Magny E.G., Eyre-Walker A., Couso J.P. Hundreds of putatively functional small open reading frames in Drosophila. Genome Biol. 2011; 12(11):R118. https://doi.org/10.1186/gb-2011-12-11-r118 PMID: 22118156

42. Parks AL, Cook K.R., Belvin M., Dompe N.A., Fawcett R., Huppert K., Tan L.R., Winter C.G., Bogart K. P., Deal J.E., Deal-Herr M.E., Grant D., Marcinko M., Miyazaki W.Y., Robertson S., Shaw K.J., Tabios M., Vysotskaia V., Zhao L., Andrade R.S., Edgar K.A., Howie E., Killpack K., Milash B., Norton A., Thao D., Whittaker K., Winner M.A., Friedman L., Margolis J., Singer M.A., Kopczynski C., Curtis D., Kaufman T.C., Plowman G.D., Duyk G., Francis-Lang H.L. Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. Nat Genet. 2004; 36(3):288–92. https://doi.org/10.1038/ng1312 PMID: 14981519

43. Port F, Chen H.M., Lee T., Bullock S.L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proc Natl Acad Sci U S A. 2014; 111(29):E2967–76. https://doi.org/10.1073/pnas.1405500111 PMID: 25002478

44. Gratz SJ, Cummings A.M., Nguyen J.N., HAmm D.C., Donohue L.K., Harrison M.M., Wildogner J., O'Connor-Giles K.M. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics. 2013; 194(4):1029–35. https://doi.org/10.1534/genetics.113.152710 PMID: 23709638

45. Gratz SJ, Ukken F.P., Rubinstein C.D., Thiede G., Donohue L.K., Cummings A.M., O'Connor-Giles K. M. Highly specific and efficienct CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics. 2014; 196(4):961–71. https://doi.org/10.1534/genetics.113.160713 PMID: 24478335

46. Bailey TL, Boden M., Buske F. A., Frith M., Grant C. E., Clementi L., Ren J. Li W.W:, Noble W.S. MEME Suite: Tools for motif discovery and searching. Nucleic Acid Res. 2009; 37:202–8. https://doi.org/10.1093/nar/gkp335 PMID: 19458158

47. Grant CE, Bailey T. L., Noble W. S. FIMO: Scanning for occurrences of a given motif.. Bioinformatics. 2011; 27(7):1017–8. https://doi.org/10.1093/bioinformatics/btr064 PMID: 21330290

48. Zhu LJ, Christensen R. G., Kazemian M., Hull C. J., Enuameh M. S., Basciotta M. D., Brasefield J.A., Zhu C., Asriyan Y., Sinha S., Wolfe S.A., Brodsky M. H. FlyFactorSurvey: a database of Drosophila transcription factor binding specificities determined using the bacterial one-hybrid system. Nucleic Acid Res. 2011; 39. https://doi.org/10.1093/nar/gkq858 PMID: 21097781

49. Khan A, Fornes O., Stigliani A., Gheorghe M., Castro-Mondragon J. A., Van Der Lee R., Bessy A., Chèneby J., Kulkarni S.R., Tan G., Baranasic D., Arenillas D.J., Sandelin A., Vandepoele K., Lenhard B., Ballester B., Wasserman W.W., Parcy F., Mathelier A. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acid Res. 2018; 46. https://doi.org/10.1093/nar/gkx1126 PMID: 29140473

50. Wingender E, Dietze P., Karas H., & Knüppel R. TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acid Res. 1996; 24(1):238–41. https://doi.org/10.1093/nar/24.1.238 PMID: 8594589

51. Belusic G. ERG in Drosophila. In: Belusic DG, editor. Electroretinograms IntechOpen; 2011.

52. Tahayato A, Sonneville R., Pichaud F., Wernet M.F., Papatsenko D., Beaufils P., Cook T., Desplan C. Otd/Crx, a dual regulator for the specification of ommatidia subtypes in the Drosophila retina. Dev Cell. 2003; 5(3):391–402. PMID: 12967559

53. Kühnlein RP, Frommer G., Friedrich M., Gonzalez-Gaitan M., Weber A., Wagner-Bernholz J.F., Gehring W.J., Jäckle H., Schuh R. spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. EMBO J. 1994; 13(1):168–79. PMID: 7905822

54. Chou WH, Hall K.J., Wilson D.B., Townson S.M., Chadwell L.V., Britt S.G. Identification of a novel Drosophila opsin reveals specific patterning of the R7 and R8 photoreceptor cells. Neuron. 1996; 19 (6):1101–15.

55. Nolo R, Abbott LA, Bellen HJ. Senseless, a Zn finger transcription factor, is necessary and sufficient for sensory organ development in Drosophila. Cell. 2000; 102(3):349–62. Epub 2000/09/07. https://doi.org/10.1016/s0092-8674(00)00040-4 PMID: 10975525.