

# GENETICS

## Supporting Information

<http://www.genetics.org/cgi/content/full/genetics.109.101329/DC1>

## Genetic Analysis of Zinc-finger Nuclease-induced Gene Targeting in *Drosophila*

Ana Bozas, Kelly J. Beumer, Jonathan K. Trautman and Dana Carroll

Copyright © 2009 by the Genetics Society of America  
DOI:10.1534/genetics.109.101329

**Deletions**

4696

|  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT wt  
 ACC TAT AGC TAC TAC ACG -AT GGC GTG GGA GTC ACT (7)  
 ACC TAT AGC TAC TAC ACG AA- GGC GTG GGA GTC ACT (7)  
 ACC TAT AGC TAC TAC ACG A-T GGC GTG GGA GTC ACT (4)  
 ACC TAT AGC TAC TAC -CG AAT GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC ACG --T GGC GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAC A-- AAT GGC GTG GGA GTC ACT (3)  
 ACC TAT AGC TAC TAC --G AAT GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC A-- -AT GGC GTG GGA GTC ACT (4)  
 ACC TAT AGC TAC TAC ACG AA- --C GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC ACG AA- --- GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAT AC- --- GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC ACG --- --C GTG GGA GTC ACT  
 ACC TAT AGC TAC TA- --- --T GGC GTG GGA GTC ACT (3)  
 ACC TAT AGC TAC TAC --- --- GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC ACG AA- --- --G GGA GTC ACT  
 ACC TAT AGC TAC TAC ACG --- --- -TG GGA GTC ACT  
 ACC TAT AGC TAC T-- --- --- GGC GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAC --- --- --- GTG GGA GTC ACT (3) \*  
 ACC TAT AGC TAC TAC A-- --- --- -TG GGA GTC ACT  
 ACC TAT AGC TA- --- --- --- -GC GTG GGA GTC ACT

ACC TAT AGC TAC TAC A-- --- --- --- GGA GTC ACT  
 ACC TAT AGC TAC T-- --- --- --- --G GGA GTC ACT  
 ACC TAT AGC --- --- --- --- --- GTG GGA GTC ACT  
 ACC TAT AGC --- --- --- --- --- GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC --- -- (803 bp Δ) - --- ---

### Insertions

4696

|  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT wt  
 |A  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT (4)  
 |C  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT (6)  
 |CG  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT  
 |AA  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT  
 |CGA  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT (2)  
 |CGAA  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT (18) †  
 |C  
 ACC TAT AGC TAC TAC ACG A-T GGC GTG GGA GTC ACT  
 |(200 bp) ‡  
 ACC TAT AGC TAC TAC A-G AAT GGC GTG GGA GTC ACT  
 |GT  
 ACC TAT AGC TAC TAC A-- -AT GGC GTG GGA GTC ACT  
 |TG TAGTCCCACG  
 ACC TAT AGC TAC TAC A-- -AT GGC GTG GGA GTC ACT

|CATGTTTCGCATGTTACTAC  
 ACC TAT AGC TAC TAC --- -AT GGC GTG GGA GTC ACT

|TACTACACCT  
 ACC TAT AGC TAC TAC --- -AT GGC GTG GGA GTC ACT

|TCCCAC  
 ACC TAT AGC TAC TAC AC- --- GGC GTG GGA GTC ACT

|G  
 ACC TAT AGC TAC TAC --- --T GGC GTG GGA GTC ACT (6)

|GTGGCTATAGCTACG  
 ACC TAT AGC TAC TAC --- --T GGC GTG GGA GTC ACT

|TACAC  
 ACC TAT AGC TAC TAC AC- --- -GC GTG GGA GTC ACT

|GTCACT  
 ACC TAT AGC TAC TAC ACG A-- --- GTG GGA GTC ACT

|GTGGGCGTG  
 ACC TAT AGC TAC TAC --- --- GGC GTG GGA GTC ACT

|GGCGTAGTAGCTACTACGTAGTGACTACTACAC  
 ACC TAT AGC TAC TAC --- --- GGC GTG GGA GTC ACT

|GGCGTAGGAGTCATGTAGTACCAC  
 ACC TAT AGC TAC TAC AC- --- --- GTG GGA GTC ACT

|G  
 ACC TAT AGC TAC --- --- --T GGC GTG GGA GTC ACT

|ACA  
 ACC TAT AGC TAC --- --- --- -GC GTG GGA GTC ACT

|TA  
 ACC TAT AGC TA- --- --- --- -GC GTG GGA GTC ACT

|G  
 ACC TAT AGC --- --- --- --T GGC GTG GGA GTC ACT

|(399 bp) \*\*  
 ACC TAT AGC TAC TA- --- --- --- --G GGA GTC ACT

|TGT

```

ACC TAT AGC TAC TAC A-- --- --- --- --A GTC ACT
                                     |GCTAGAGATC
ACC TAT AGC TAC TAC A-- --- --- --- --- ---
                                     |TATTGCCTGAC
ACC TAT AGC TAC TAC --- --- (44 bp Δ) -- --- ---
                                     |TGAA
ACC TAT AGC TAC TAC ACG A-- --- (915 bp Δ) - ---

```

†The 4-base fill-in and blunt join of the overlap created by ZFN cleavage.

‡This insertion corresponds to sequences in the 18S rRNA gene.

\*\* This insertion matches sequences from the histone gene cluster.

FIGURE S1.—NHEJ mutations at *ry* from *lig4<sup>+</sup>* flies. Sequences of the ZFN-induced mutants in the *ry* gene. The wild type sequence is shown above the deletions and the insertions for comparison. The first position in the sequence is numbered from start of transcription. ZFN recognition sequences are in red; insertions and substitutions are in blue. All these sequences were isolated independently – i.e., they came from different heat-induced parents. When a mutant sequence was isolated more than once, the number of observations is shown in parentheses to the right. Many of the short insertions are partially homologous to *ry* sequences in the immediate vicinity of the ZFN cut. \*A 9-bp deletion found more commonly from *lig4<sup>-</sup>* parents.

**Deletions**

4696

|  
 ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT wt  
 ACC TAT AGC TAC TAC ACG AA- GGC GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAC ACG A-T GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC --G AAT GGC GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAC A-- AAT GGC GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAC ACG --T GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC A-- -AT GGC GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAC A-- --T GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC A-- --- GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC --- --T GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC ACG AA- --- --G GGA GTC ACT (2)  
 ACC TAT AGC TAC T-- --- --T GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TA- --- --- GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC --- --- -GC GTG GGA GTC ACT  
 ACC TAT AGC TAC T-- --- --- GGC GTG GGA GTC ACT (2)  
 ACC TAT AGC TAC TAC --- --- --- GTG GGA GTC ACT (7) \*  
 ACC TAT AGC TAC TAC A-- --- --- -TG GGA GTC ACT  
 ACC TAT AGC TA- --- --- --T GGC GTG GGA GTC ACT  
 ACC TAT AGC TAC TA- --- --- --C GTG GGA GTC ACT  
 ACC TAT AGC TAC TAC AC- --- --- --- -GA GTC ACT  
 ACC TAT AGC --- --- --- --- GTG GGA GTC ACT

**Insertions**

4696

|  
ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT wt

|A  
ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT (2)

|TA  
ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT

|CGAA  
ACC TAT AGC TAC TAC ACG AAT GGC GTG GGA GTC ACT (2) †

|CAGTATGCC  
ACC TAT AGC TAC TAC ACG --T GGC GTG GGA GTC ACT

|GCTAT  
ACC TAT AGC TAC TAC A-- -AT GGC GTG GGA GTC ACT

|TACGA  
ACC TAT AGC TAC TAC A-- -AT GGC GTG GGA GTC ACT

|CT  
ACC TAT AGC TAC TAC --- -AT GGC GTG GGA GTC ACT

|T  
ACC TAT AGC TAC TAC --- -AT GGC GTG GGA GTC ACT

|ATAGCTACTAC  
ACC TAT AGC TAC TAC A-- --T GGC GTG GGA GTC ACT

|G  
ACC TAT AGC TAC TAC --- --T GGC GTG GGA GTC ACT (2)

|GTAGCTACTCCTAC  
ACC TAT AGC TAC TAC --- --T GGC GTG GGA GTC ACT

|GTGGTAGTACG  
ACC TAT AGC TAC TAC --- --T GGC GTG GGA GTC ACT

|GTAGCGTGGGAG  
ACC TAT AGC TAC TAC --- --T GGC GTG GGA GTC ACT

|GTCACT

```

ACC TAT AGC TAC TAC ACG A-- --- GTG GGA GTC ACT
                |TATAGTTAC
ACC TAT AGC TAC TAC --- --- GGC GTG GGA GTC ACT
                |GTG
ACC TAT AGC TAC TAC --- --- GGC GTG GGA GTC ACT
                |GGTAGCG
ACC TAT AGC TAC TA- --- --T GGC GTG GGA GTC ACT
                |ATACC
ACC TAT AGC TAC --- --- -AT GGC GTG GGA GTC ACT
                | (39 bp)
ACC TAT AGC TAC TA- --- --- GGC GTG GGA GTC ACT
                |G
ACC TAT AGC TAC T-- --- --- GGC GTG GGA GTC ACT
                |CACTACTAC
ACC TAT AGC --- --- --- -AT GGC GTG GGA GTC ACT
                |GTATCACTGTGGGA
ACC TAT AGC TAC TA- --- --- --- GTG GGA GTC ACT
                | (31 bp)
ACC TAT AGC TAC TAC --- --- --- --G GGA GTC ACT
                |GTGG
ACC TAT AGC TAC TAC --- --- --- --G GGA GTC ACT
                | (720 bp) †
ACC TAT AGC --- --- --- --- -GC GTG GGA GTC ACT
                |AG
ACC --- --- --- --- --- --T GGC GTG GGA GTC ACT
                | (64 bp) **
ACC TAT AGC TAC TAC A-- --- -- (71 bp Δ) -- ---

```

\*A 9-bp deletion found more commonly from *lig4* parents.

†The 4-base fill-in and blunt join of the overlap created by ZFN cleavage.

‡This insertion corresponds to *gish* mRNA (see Fig. 5 in main text).

\*\* This insertion matches sequences downstream of the *ry* gene.

FIGURE S2.—NHEJ mutations at from *lig4* flies. Other features as in Figure S1.



**TABLE S1**  
**Statistical analysis of gene targeting parameters**

Males		----- p values -----		
Genotypes		%Yielders	%ry	%HR
M1 : M2	wt vs. <i>spnA</i> <sup>+/-</sup>	0.81	0.623	0.14
M1 : M3	“	0.19	0.126	0.40
M1 : M4	wt vs. <i>spnA</i> <sup>-/-</sup>	0.69	0.024	9 x 10 <sup>-6</sup>
M1 : M5	“	0.028	0.51	0.00064
M1 : M6	“	0.00042	0.222	0.00059
M1 : M7	wt L vs. C	0.76	0.118	0.30
M1 : M12	wt L2 vs. L3	1.0	0.017	0.67
M7 : M8	wt vs. <i>spnA</i> <sup>+/-</sup> C	0.094	0.00025	0.013
M7 : M9	“	0.50	0.381	0.22
M7 : M10	wt vs. <i>spnA</i> <sup>-/-</sup> C	0.010	0.011	3 x 10 <sup>-5</sup>
M7 : M11	“	0.53	0.261	3 x 10 <sup>-5</sup>
M12 : M13	wt vs. <i>okr</i> <sup>+/-</sup>	0.21	0.614	0.010
M12 : M14	wt vs. <i>okr</i> <sup>-/-</sup>	0.47	0.194	0.071
M1 : M15	wt vs. <i>lig4</i>	1.6 x 10 <sup>-5</sup>	0.000425	7 x 10 <sup>-11</sup>
M1 : M16	wt vs. <i>lig4 spnA</i> <sup>-/-</sup>	0.0019	0.219	0.028
M7 : M17	wt vs. <i>lig4</i> C	0.0056	7.7 x 10 <sup>-8</sup>	0.019
M7 : M18	wt vs. <i>lig4 spnA</i> <sup>-/-</sup> C	0.00087	0.38	0.0025
M15 : M16	<i>lig4</i> vs. <i>lig4 spnA</i> <sup>-/-</sup>	0.83	0.0016	9.5 x 10 <sup>-14</sup>
M17 : M18	<i>lig4</i> vs. <i>lig4 spnA</i> <sup>-/-</sup> C	1.0	0.00048	4.4 x 10 <sup>-7</sup>
M5 : M10	<i>spnA</i> <sup>-/-</sup> L vs. C	0.84	0.061	0.060
M6 : M11	“	0.011	0.439	0.0031
M15 : M17	<i>lig4</i> L vs. C	0.83	0.278	0.0043
M16 : M18	<i>lig4 spnA</i> <sup>-/-</sup> L vs. C	1.0	0.487	0.014

  

Females		----- p values -----		
Genotypes		%Yielders	%ry	%HR
F1 : F2	wt vs. <i>spnA</i> <sup>+/-</sup>	0.68	0.8	0.19
F1 : F3	“	0.026	0.223	1.0
F1 : F6	wt vs. <i>spnA</i> <sup>-/-</sup>	8 x 10 <sup>-9</sup>	0.485	5 x 10 <sup>-8</sup>
F1 : F7	wt L vs. C	0.62	0.731	0.32
F1 : F12	wt L2 vs. L3	0.84	0.588	0.0040
F7 : F8	wt vs. <i>spnA</i> <sup>+/-</sup> C	0.14	0.0623	0.033
F7 : F9	“	1.0	0.587	0.14

F7 : F11	wt vs. <i>spnA</i> <sup>-/-</sup> C	0.00026	0.95	4 x 10 <sup>-7</sup>
F1 : F19	wt vs. <i>lig4</i> <sup>+/-</sup>	0.69	0.0923	0.025
F1 : F20	wt vs. <i>lig4</i> <sup>-/-</sup>	0.012	0.00046	8 x 10 <sup>-25</sup>
F7 : F21	wt vs. <i>lig4</i> <sup>-/-</sup> C	0.0053	0.00395	4 x 10 <sup>-5</sup>
F20 : F21	<i>lig4</i> <sup>-/-</sup> L vs. C	0.86	0.803	6.5 x 10 <sup>-8</sup>
F19 : F20	<i>lig4</i> <sup>-/-</sup> L vs. <i>lig4</i> <sup>+/-</sup>	0.0042	6.7 x 10 <sup>-6</sup>	9.4 x 10 <sup>-11</sup>

Males vs. Females		----- p values -----		
Genotypes		%Yielders	%ry	%HR
M1 : F1	wt L	0.54	0.191	0.035
M2 : F2	<i>spnA</i> <sup>+/-</sup>	1.0	0.158	0.15
M3 : F3	“	0.0035	9.6 x 10 <sup>-9</sup>	0.046
M6 : F6	<i>spnA</i> <sup>-/-</sup>	0.00042	0.49	0.11
M7 : F7	wt C	0.72	0.524	0.22
M8 : F8	<i>spnA</i> <sup>+/-</sup> C	1.2 x 10 <sup>-5</sup>	1.2 x 10 <sup>-10</sup>	0.024
M11 : F11	<i>spnA</i> <sup>-/-</sup> C	0.0027	0.825	1.0
M12 : F12	wt AB3	0.46	0.0274	4.6 x 10 <sup>-6</sup>
M13 : F13	<i>okr</i> <sup>+/-</sup>	0.30	2.52 x 10 <sup>-5</sup>	1.0
M15 : F20	<i>lig4</i>	0.63	0.394	1.9 x 10 <sup>-12</sup>
M17 : F21	<i>lig4</i> C	0.63	0.089	0.0030

Genotypes are as in Table 2, with M indicating males and F indicating females of specified class. %Yielders, %ry and %HR have the same meanings as in Table 2 of the main text. L denotes linear donor, C denotes circular donor. As stated in the main text, p values were calculated with Fisher's exact test for %Yielders and %HR, and with the glm function of the R software package for %ry.