Assessment of Factors for Successful Gene Editing using CRISPR-Cas9 Technology

Elif F. Acar

Objective

The primary objective in this study is to determine the experimental factors affecting the germline transmission status and the founder rate of mutant mouse lines produced using the CRISPR-Cas9 technology.

Data

The study data after filtering consist of 4473 experiments conducted at eight IMPC centres. The number of experiments at each centre, along with the number of genes edited in these experiments, is provided in Table 1. We observe that seven out of eight centres had multiple experimental attempts for some genes. Reported also are the proportions of experiments and genes repeated within each centre. In total, there are 3973 unique genes that were edited using the CRISPR-Cas9 technology. All genes were edited in only one center, except for two genes: `Rpl5` was edited twice at BCM using electroporation and twice at WTSI using cytoplasmic injection; `Ebf3` was edited twice at BCM using cytoplasmic injection and three times at Harwell using pronuclear injection twice and electroporation once.

Table 1: The number of experiments and of genes across the eight IMPC centers (in alphabetical order), as well as in total. Reported also are the proportions of experiments and genes repeated within each centre.

	BCM	**CCP-IMG**	**Harwell**	**ICS**	**JAX**	**TCP**	**UCD**	**WTSI**	**Total**
Number of Experiments	679	339	410	60	1430	386	701	468	4473
Number of Genes	585	278	305	60	1340	341	683	383	3973
% Repeated Experiments	25.8	32.7	41.0	0.0	12.4	22.5	5.1	31.2	20.1
% Repeated Genes	13.8	18.0	20.7	0.0	6.6	12.3	2.6	15.9	10.1

The success of Cas9 gene editing is evaluated based on the germline transmission (GLT) status and the founder rate (FR). There are 104 experiments that are in progress, hence their GLT status is unknown. We exclude these experiments from our investigations and focus on the complete cases for which the GLT status is known.

Evaluation of Factors Associated with Success in Gene Editing

We investigate the impact of genomic and experimental factors on the FR and the GLT status using the data on the 3869 genes. For genes with multiple attempts at the same center, we consider only the latest attempt with a known GLT status.

Evaluation of Genomic Factors

The genomic factors of interest are the cellular essentiality and lethality of genes. We first examine whether the FR differs between essential and non-essential genes, and between lethal and non-lethal genes. The essentiality and lethality information was available for 3428 and 804 genes, respectively. Table 2 gives the number of essential and lethal genes at each chromosome.

Table 2: The number of genes across chromosomes and their distribution based on essentiality and lethality.

		Essentiality			Lethality		
Chromosome	**Num.of Genes**	**Essential**	**Non-essential**	**Unknown**	**Lethal**	**Non-lethal**	**Unknown**
1	251	24	205	22	11	39	201
2	277	30	240	7	6	45	226
3	212	21	173	18	5	35	172
4	212	23	180	9	7	38	167
5	254	38	200	16	10	38	206
6	192	15	155	22	5	39	148
7	311	36	237	38	12	57	242
8	208	28	173	7	7	40	161
9	251	26	206	19	10	42	199
10	170	24	132	14	10	28	132
11	262	31	216	15	8	51	203
12	129	17	99	13	3	19	107
13	145	19	106	20	10	32	103
14	144	12	125	7	5	18	121
15	150	19	120	11	8	25	117
16	126	15	106	5	4	15	107
17	172	24	131	17	9	27	136
18	112	15	85	12	6	26	80
19	131	14	108	9	4	23	104
x	159	0	0	159	1	26	132
y	1	0	0	1	0	0	1
Total	3869	431	2997	441	141	663	3065

Comparisons of Founder and GLT Rates of Essential and Non-essential Genes

Figure 1 displays the distribution of the founder rates (FR) for non-essential and essential genes. The Wilcoxon rank sum test suggests a significant difference (p-value = 0) between the median FR of non-essential (median = 0.061) and essential genes (median = 0.033). Hence, we conclude that gene editing using Cas9 is more successful for non-essential genes than essential genes. We also compared the germline transmission (GLT) rates of non-essential and essential genes focusing on the experiments where a positive FR is obtained.However, the Pearson chi-square test indicated no significant difference (p-value = 0.15) between the GLT rates of non-essential (GLT = 0.968) and essential (GLT = 0.952) genes.

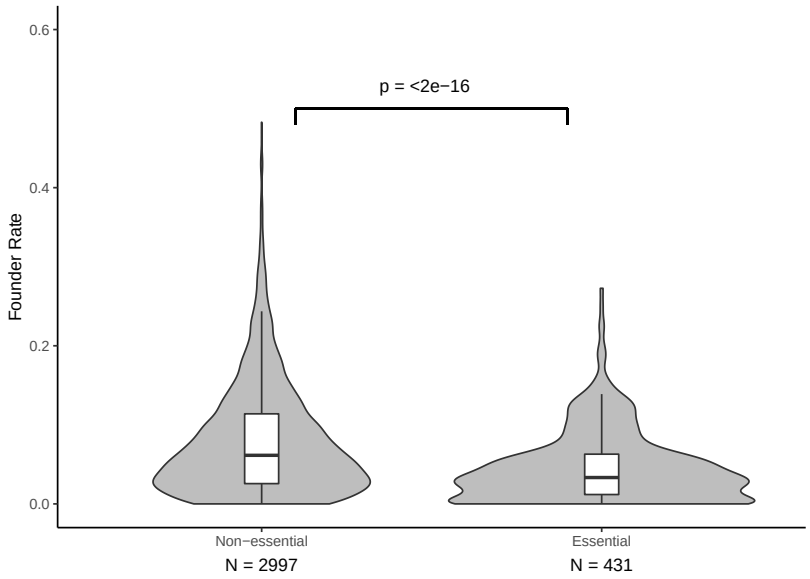

Figure 1: FR of non-essential and essential genes.

Comparisons of Founder and GLT Rates of Lethal and Non-lethal Genes

Similarly, we compared the FR for non-lethal and lethal genes. As can be seen in Figure 2, the median FR for non-lethal genes (median = 0.063) is higher than that of lethal genes (median=0.036), with a significant difference based on the Wilcoxon rank sum test (p-value = 0). However, the Pearson chi-square test indicated no significant difference (p-value = 1) between the GLT rates of non-lethal (GLT $= 0.997$) and lethal (GLT $= 0.993$) genes. Note that there are only very few genes that had the viability information but did not have GLT. As a result, the estimated GLT rates are both very high (over 99%).

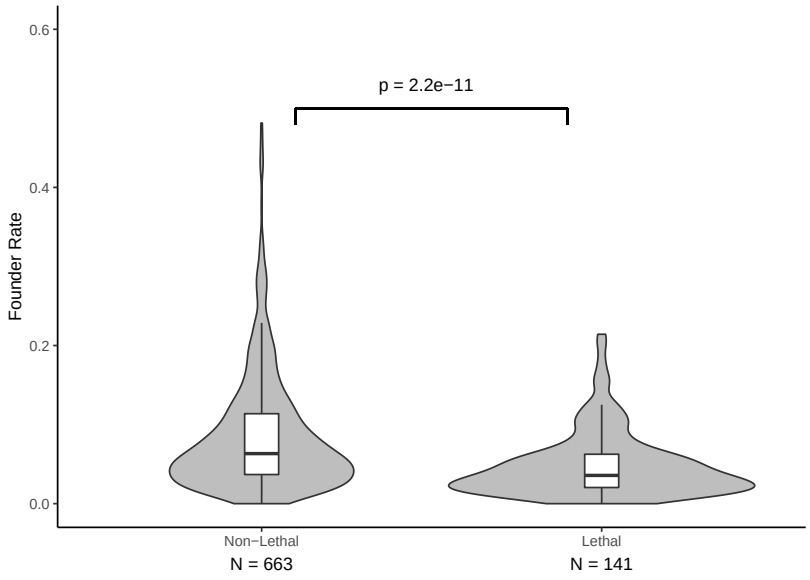

Figure 2: FR of non-lethal and lethal genes.

Comparisons of Birth Rates based on Gene Essentiality and Gene Lethality

In addition to the FR and GLT rate, we also compared the birth rates (BR) of essential and non-essential genes, and of lethal and non-lethal genes. We concluded from Figure 3 that non-essential genes (median = 0.229) have a significantly higher birth rate than essential genes (median = 0.189). Similarly, non-lethal genes (median = 0.231) were found to have a higher birth rate than lethal genes (median = 0.184).

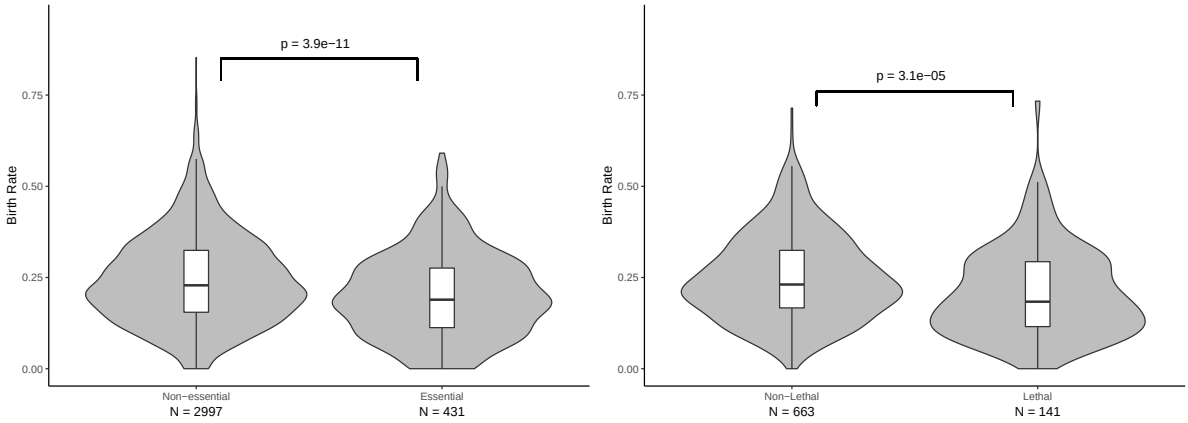

Figure 3: BR of non-essential and essential genes (left) and of non-lethal and lethal genes (right).

Evaluation of Experimental Factors

The experimental factors that are of interest are the delivery method (cytoplasmic injection, electroporation and pronuclear injection), number of guides (2 Guides versus 4 Guides), deletion size (ranging from 35 to 9487) and number of founders selected for breeding (ranging from 0 to 16).

Comparisons of Founder and GLT Rates based on Delivery Method

We examine the FR for the three delivery methods in Figure 4. The pairwise comparisons of the FR via the Wilcoxon rank sum test indicate a significant difference between cytoplasmic injection (median = 0.057) and pronuclear injection (median = 0.029) and between electroporation (median = 0.062) and pronuclear injection. The FR does not significantly differ between cytoplasmic injection and electroporation. The GLT rates for the three delivery methods are all higher than 95% (0.954, 0.965 and 0.973, respectively) and do not show any significant difference in the pairwise comparisons via the Pearson chi-square test (all three p-values > 0.15).

Comparisons of Founder and GLT Rates based on Number of Guides

The FR of genes edited using 2 guides versus 4 guides are displayed in Figure 5. Based on the Wilcoxon rank sum test, we concluded that there is no significant difference between the median FR of genes edited using 2 guides (median = 0.056) and 4 guides (median = 0.056). The Pearson chi-square test also indicated no significant difference (p-value = 0.096) between the GLT rates of genes edited using 2 guides (GLT = 0.968) and 4 guides (GLT = 0.957) genes.

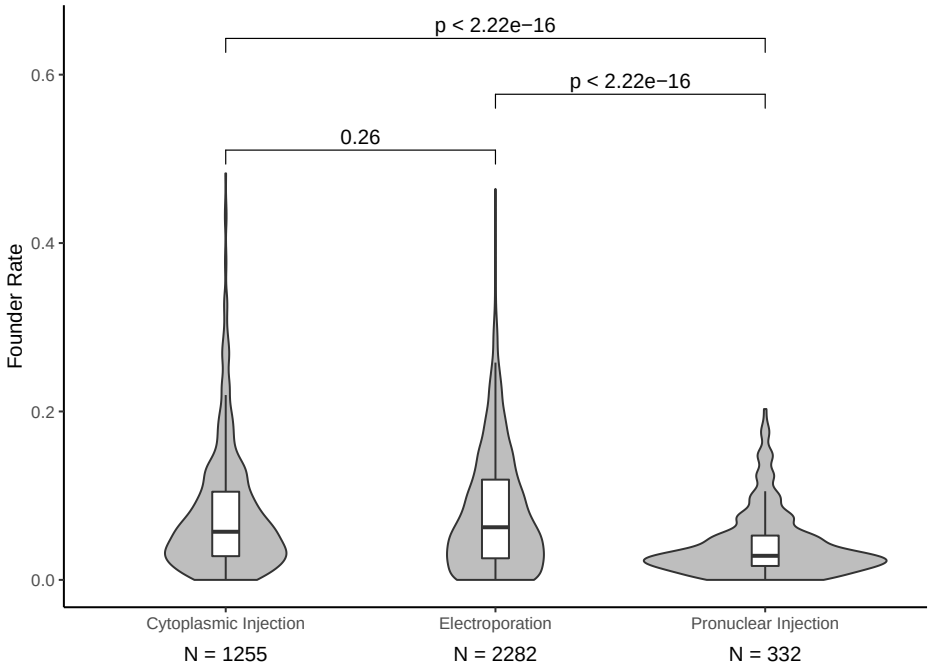

Figure 4: FR of genes under different delivery methods.

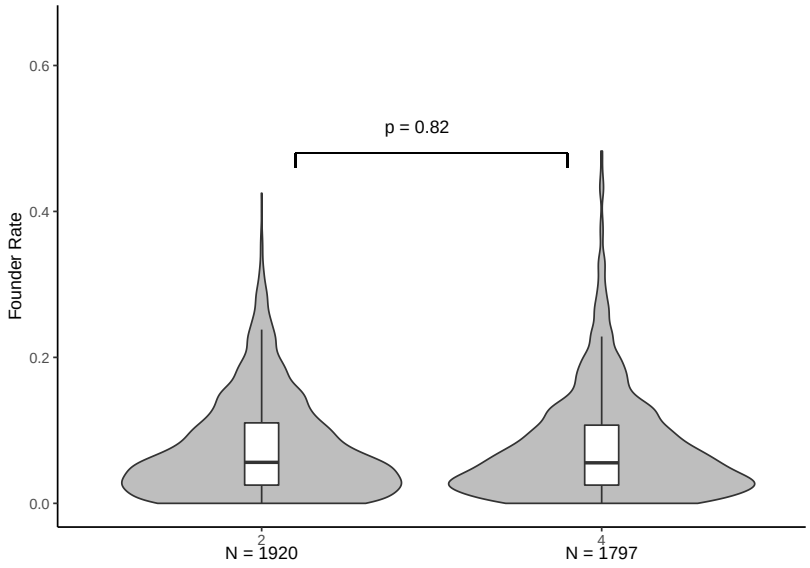

Figure 5: FR of genes under different number of guides.

Comparisons of Founder and GLT Rates based on Deletion Size

The deletion size is another experimental factor that may influence the success of gene editing. The FR of genes across deletion size of six bins of roughly equal size are displayed in Figure 6. Based on the Kruskal-Wallis test for comparing the medians of six groups, we conclude that the FR does not differ significantly across different deletion sizes. The GLT rate comparison of the six groups also yielded no significant difference (p-value = 0.597).

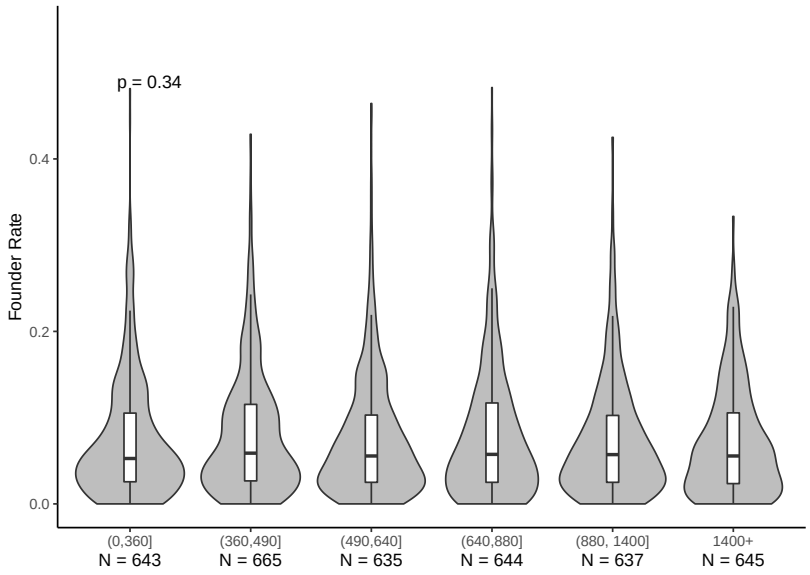

Figure 6: FR of genes across different deletion size (base pairs).

Comparisons of Founder and GLT Rates based on Number of Founders Selected for Breeding

For the genes where there was a founder selected for breeding (444 out of 3869 genes had no founders), we examine the relationships of the number of founders selected for breeding with the GLT rate.

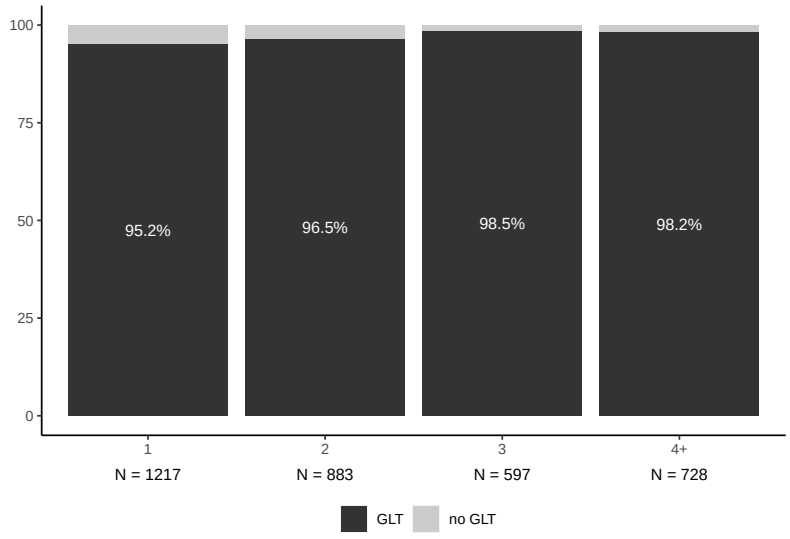

Figure 7: GLT rate across different number of founders selected for breeding.

The GLT rate comparisons in Figure 7 indicated a significant difference (p-value = 0) across the four groups, where the GLT rates are $(0.952, 0.965, 0.985, 0.982)$. The pairwise comparisons of the GLT rates (with multiple comparison correction based on the Holm method, which is the default method) suggest a significant difference only for the cases with 3 founders (p-value = 0.004) and with 4 or more founders (p-value = 0.004) compared to the case where one founder was selected for breeding.

Evaluation of Success Rates in Repeated Attempts

We next examine the success rates of Cas9 gene editing in repeated attempts by production centre. Figure 8 displays the number and proportion of genes that were successful with GLT, aborted upon no GLT and repeated upon no GLT in the first, second and third attempts. For four or more attempts, we summarize the cases with and without GLT only. Among the 3871 unique gene-centre combinations, 3028 were successful with GLT in the first attempt, 440 were aborted and 403 were repeated for at least a second time. The high success rate in the first attempt drops drastically in the second attempt and is neglible in the follow-up attempts.

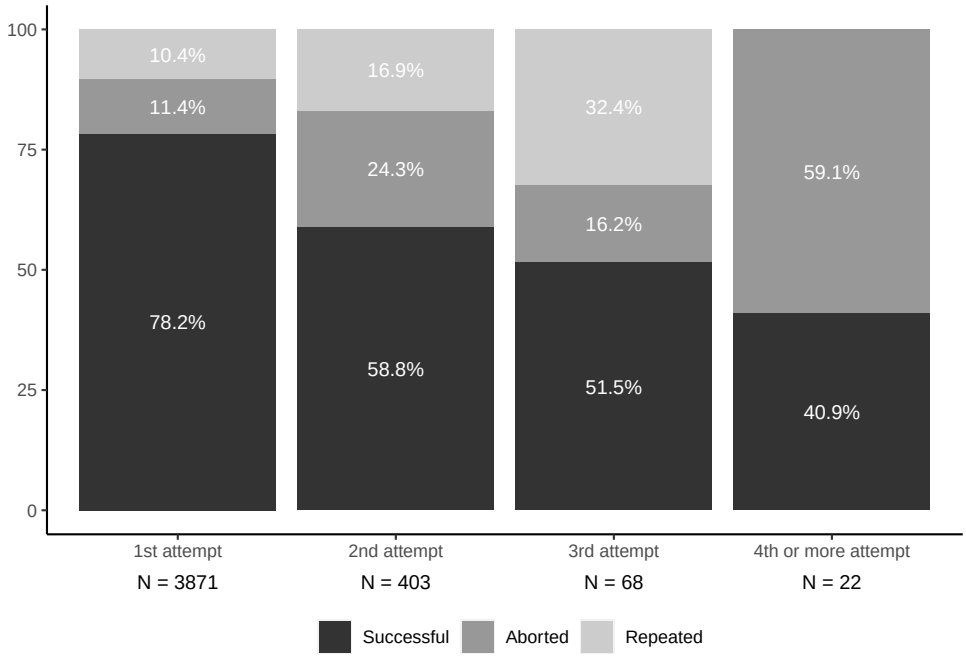

Figure 8: Distribution of the GLT status of experiments at each attempt

We also examine the success rates of Cas9 gene editing for essential and non-essential genes. Figure 9 displays the number and proportion of essential and non-essential genes that were succcessful with GLT, aborted upon no GLT and repeated upon no GLT in repeated trials. We observe that the success rate is consistently lower for essential genes than non-essential genes in each sequential attempt.

The reason for failure in obtaining GLT was available for 83 experiments, where 67 failed in the first trial. Figure 10 displays the distribution of the identified reasons of failure in these experiments. The majority of failures was due to a failure in allele QC. In 43.4% of the failed experiments, there was either no progeny or only wild-type progeny was obtained. Loosing founder before breeding was another common reason for no GLT.

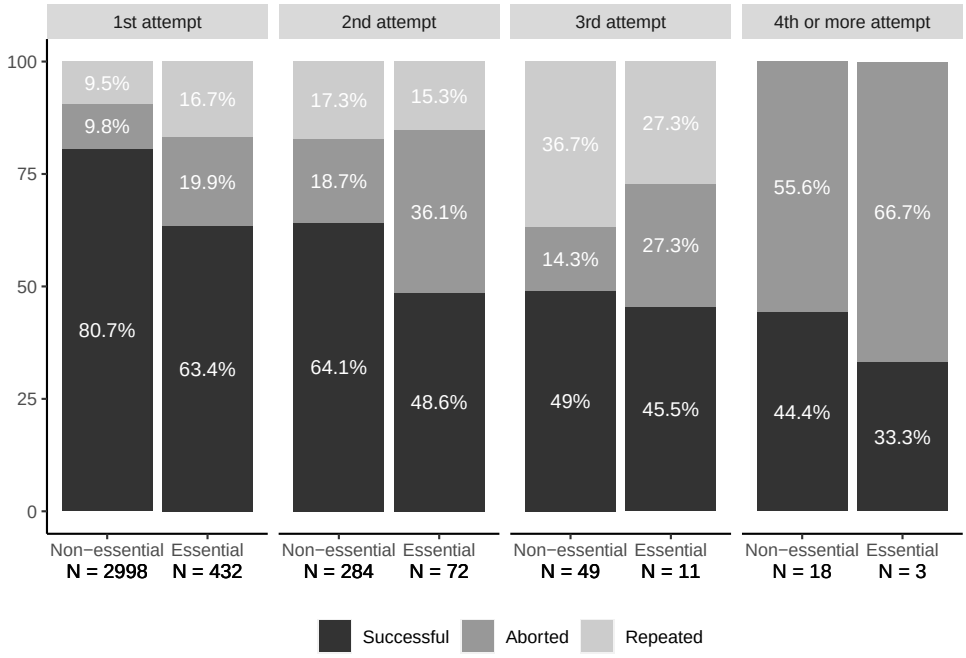

Figure 9: Distribution of the GLT status of experiments at each attempt for essential and non-essential genes

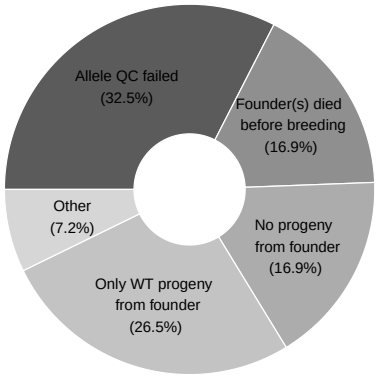

Figure 10: Distribution of reasons for GLT failure.

Model-based Assessment of Experimental Factors in Repeated Attempts of Gene Editing

Another question of interest is whether changing different experimental factors increase or decrease the success rate of the repeated attempts from the immediately previous attempt. To address this question, we consider all consecutive attempts and identify the experimental factors that differ between the two attempts. For the initial investigations, we perform the chi-square test for contingency tables and assess whether there is any association between the GLT status and changing each experimental factor. The results suggest no significant association with changing the delivery method (p-value = 0.507), number of guides (p-value = 0.086, when considering both increase and decrease rather than only change) and exon (p-value = 0.192). However, changing the guide size (p-value = 0.01) is found to be associated with the GLT status.

We further consider a logistic regression model for the GLT status conditional on the changes in the experimental factors, namely, the delivery method, number of guides, guide sequence and exon. The fitted model suggests that a borderline significant effect (p-value = 0.048) for the guide sequence on the GLT status after accounting for the other factors in the model. The odds ratio for the change in the guide sequence is 1.88, which indicates that changing the guide sequence upon an unsuccessful attempt increases the odds of obtaining GLT by a factor of 1.88 compared to the not changing the guide sequence. See below for the summary of the fitted model and effect plots. Even though the other effects were not significant, the negative slope in the effect plots indicate a decrease in the GLT rate (e.g., Cas9 type) and the positive slope indicates an increase in the GLT rate (e.g., Guide sequence).

		GLT	
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	1.18	0.95 – 1.47	0.136
Δ.Delivery.Method	0.74	0.45 – 1.21	0.222
Δ.Num.Guides [↓]	1.88	0.92 – 4.07	0.092
Δ.Num.Guides [↑]	0.86	0.37 – 2.03	0.731
Δ.Guide.Seq	1.88	1.02 – 3.58	**0.048**
Δ.Exon	1.22	0.44 – 3.58	0.709
Observations	498		

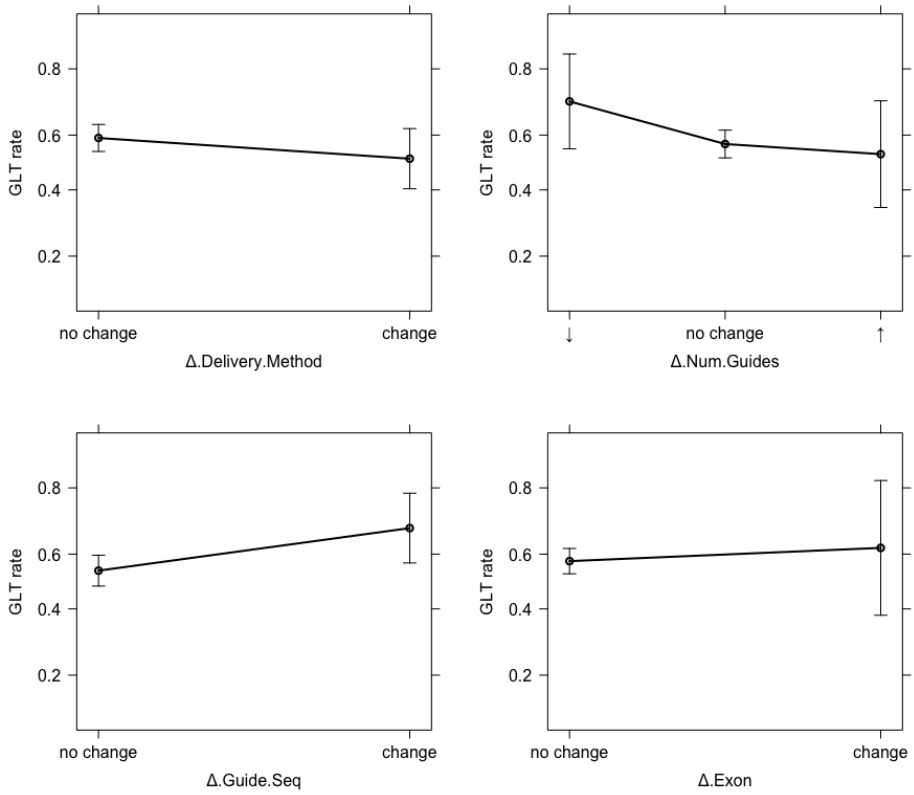

Table 3: Summary of estimated effects.

		fit	se	lower	upper
Delivery.Method	no change	0.590	0.025	0.541	0.637
	change	0.514	0.057	0.404	0.623
Num.Guides	decrease	0.712	0.073	0.550	0.833
	no change	0.568	0.026	0.517	0.618
	increase	0.531	0.100	0.339	0.714
Guide.Seq	no change	0.540	0.029	0.483	0.596
	change	0.688	0.057	0.568	0.787
Exon	no change	0.575	0.023	0.529	0.620
	change	0.622	0.120	0.378	0.817

Since the changes in experimental factors are expected to be associated with each other, we calculated the phi coefficient as well as the tetrachoric correlation coefficient as measures of association between two binary variables. We found moderate-to-strong associations between the changes in the guide sequence with exon($\phi = 0.41$, $\rho = 0.884$) and number of guides ($\phi = 0.482$, $\rho = 0.739$). Since these associations may cause collinearity in the logistic regression model, we also fit an elastic net logistic regression model which performs variable selection and accounts for collinearity through penalization. The penalty parameters (obtained using 5-fold cross-validation on 80% training data) are $\lambda = 0.07$ and $\alpha = 0.1$. The resulting

model kept all four experimental factors as relevant to explain the GLT status, suggesting that the elastic net penalty may not be necessary, and hence the fitted logistic regression model is sufficient. Note that both the logistic regression and the elastic net regression have only 57.58% classification accuracy of the GLT status, which is only slightly better than a random guess.

A Detailed Look at the Experimental Factors

We next examine whether the proportion of essential genes show any significant difference between various categories of experimental factors. This was assessed using the chi-square test for proportions. Table 4 reports the proportion of essential genes in each category of the experimental factors along with the p-values for the pairwise comparisons (adjusted for multiple comparisons based on the Holm method).

Table 4: Pairwise comparisons of the proportion of essential genes in each experimental factor.

		Proportion 1	Proportion 2	p-value
Delivery Method	Cytoplasmic Injection vs Electroporation	0.110	0.142	0.025
	Cytoplasmic Injection vs Pronuclear Injection	0.110	0.072	0.077
	Electroporation vs Pronuclear Injection	0.142	0.072	0.004
Number of Guides	2 Guides vs 4 Guides	0.128	0.127	0.986
Deletion Size	(0,360] vs (360,490]	0.138	0.111	1.000
	(0,360] vs (490,640]	0.138	0.126	1.000
	(0,360] vs (640,880]	0.138	0.136	1.000
	(0,360] vs (880, 1400]	0.138	0.086	0.095
	(0,360] vs 1400+	0.138	0.159	1.000
	(360,490] vs (490,640]	0.111	0.126	1.000
	(360,490] vs (640,880]	0.111	0.136	1.000
	(360,490] vs (880, 1400]	0.111	0.086	1.000
	(360,490] vs 1400+	0.111	0.159	0.274
	(490,640] vs (640,880]	0.126	0.136	1.000
	(490,640] vs (880, 1400]	0.126	0.086	0.395
	(490,640] vs 1400+	0.126	0.159	1.000
	(640,880] vs (880, 1400]	0.136	0.086	0.124
	(640,880] vs 1400+	0.136	0.159	1.000
	(880, 1400] vs 1400+	0.086	0.159	0.004
Number of Founders	1 vs 2	0.101	0.137	0.132
	1 vs 3	0.101	0.094	1.000
	1 vs 4+	0.101	0.096	1.000
	2 vs 3	0.137	0.094	0.132
	2 vs 4+	0.137	0.096	0.132
	3 vs 4+	0.094	0.096	1.000

The findings suggest that the proportion of essential genes do not show any significant difference

among the categories of the number of guides, the number of founders selected for breeding and the deletion size (except for the last two categories, but since the bins were user-defined this could change with a different choice). However, the same is not true for the delivery methods. The proportion of essential genes is much higher in electroporation than cytoplasmic injection and pronuclear injection. Despite this fact, electroporation achieved the highest median founder rate (see Figure 4). Hence, the high proportion of essential genes edited by this method did not result in an inferior performance in gene editing.