Supplementary Information

Programming cells by multiplex genome engineering and accelerated evolution

Harris H. Wang[1,2,3,8], Farren J. Isaacs[1,8], Peter A. Carr[4,5], Zachary Z. Sun[6], George Xu[6], Craig R. Forest[7] & George M. Church[1]

1. Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.
2. Program in Biophysics, Harvard University, Cambridge, MA 02138, USA.
3. Program in Medical Engineering Medical Physics, Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA.
4. The Center for Bits and Atoms, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
5. Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
6. Harvard College, Cambridge, MA 02138, USA.
7. George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.

8. These authors contributed equally to this work.

Correspondence and request for materials should be addressed to G.M.C. (http://arep.med.harvard.edu/gmc/email.html)

Optimized Design Criteria for MAGE Oligonucleotides

Oligonucleotide-mediated allelic replacement was achieved in the modified *E. coli* strain EcNR2 ($mutS^-$, λ-Red^+) by directing oligos to the lagging strand of the replication fork during DNA replication[1]. Targeting the lagging strand of replicating DNA with single-stranded oligonucleotides (ss-oligos) has been shown to be more efficient than targeting the leading strand in this $mutS^-$ strain (EcNR2)[2]. The replacement efficiency was characterized by using oligos either to inactivate the *lacZ* gene and screen for white colonies on Xgal/IPTG-containing agar plates or to fix a defective *cat* gene and select for chloramphenicol resistant colonies.
In rare cases during MAGE experiments, we observed small genomic sequence changes in the oligo targeted region that are not by design, i.e., 7^+ bp mutations when only 6 bp are targeted (Fig. 3). We hypothesize that these additional mutations are likely the result of allelic replacement by faulty oligos that arise from errors during oligo synthesis. Purification of oligos may reduce instances of such cases.

Supplementary Figure 2a shows that replacement efficiency was found to be dependent on oligo length, highest at 90 basepairs (bp). We hypothesize that the 90 bp oligo has the most optimal replacement efficiency for two main reasons. First, the λ-Red single-stranded DNA-binding protein β, has been shown to require at least 30 bp to complex with oligos *in vitro*[3]. *In vivo*, shorter oligos have fewer basepairs of homology to hybridize to the targeted chromosomal site, thus decreasing the likelihood of replacement. Second, while oligos longer than 90 bp may have more regions of homology to the chromosome, they are also more likely to form secondary structures. Inhibitory secondary structures (*e.g.*, hairpin loops) can lead to dramatically lower efficiencies of replacement since reducing the number of exposed bases on the oligo will decrease the frequency of hybridization to its chromosomal target. Along these lines, we observed that oligos with computationally predicted minimal folding energies[4] of less than -12.5 kcal/mol showed significantly reduced allelic replacement frequencies experimentally

(Supplementary Fig. 2b). Phosphorothioate bonds located at the terminal bases may increase replacement efficiency by preventing *in vivo* degradation of synthetic oligonucleotide molecules by endogenous exonucleases in the cell[5]. Phosphorothioated nucleotides increased replacement efficiency by more than 2-fold when placed at the 5' terminus, but showed no effect when placed at the 3' terminus (Supplementary Fig. 2c). Increasing the number of phosphorothioated bases at the 5' terminus increased the efficiency of replacement, which saturates to its highest level at four phosphorothioated bases (Supplementary Fig. 2d). Oligonucleotides, in which all bases contained phosphorothioated bonds, did not incorporate into the chromosome (data not shown). The replacement efficiency remains high across a wide range of oligo concentrations (0.05-50 μM), thus allowing for large and highly complex oligo pools (Supplementary Fig. 2e). Allelic replacement efficiency was low when low concentrations of oligos (<0.05 μM) were used, suggesting a dilution effect. In fact, at low oligo concentrations (*i.e.*, 0.005 μM), there are on average three DNA molecules per volume of a cell (~10^{-18} m^3), leading to drastically decreased likelihood of a replacement event. Therefore, increasing the amount of oligos available for allelic replacement by either increasing the oligo concentration during electroporation or by increasing the oligo half-life inside the cell (via terminal phosphorothioated nucleotides) will lead to higher efficiencies of replacement. Interestingly, we observed chromosomal deletions of up to 45 kbp with a single 90mer oligo using the EcNR2 (recA+) strain as well as a recA- EcNR2 derivative, suggesting a recA-independent β-mediated mechanisms.

Design of MAGE Oligonucleotides for DXP Pathway

The main text of this paper described the design criteria that were implemented to optimize the DXP pathway for lycopene production. Here, we provide additional details to clarify our oligo design criteria. The design of every oligo was based on optimization experiments such that oligo length, concentration, stability, secondary structure, strand bias and modification were optimal (Supplementary Table 1 and Supplementary Fig. 2). Two main oligo design strategies were implemented: 1) oligos with specified sequences produced specific changes by making targeted modification that knocked out the expression of target genes (*ytjC*, *fdhF*, *aceE*, *gdhA*) and 2) oligos with degenerate sequences produced diverse changes tailored for exploring a vast sequence space of RBS strengths. Importantly, in both oligo designs, the location of the genetic modification is precise and well-defined based on homology arms of the oligos. As described in the main text, the degenerate oligos were designed to mutate RBS sequences to be more similar to the canonical Shine-Dalgarno sequence (TAAGGAGGT)[6], giving rise to enhanced translation efficiency. More specifically, RBS optimization utilized 90mer oligo pools containing the DDRRRRRDDDD degeneracy at the 41-51 bp position of the oligos (D = G, A, T and R = G, A). This mutation region targeted the -4 through -14 positions from the start codon of each gene with an optimal RBS spacing of 5 bp for replacement by one of the oligos from the degenerate pool. We also calculated the cost and maximum level of degeneracy that can be introduced into a single oligo. For 30 USD we obtain 50 nmol yield of a 90mer oligo, giving us 3×10^{16} molecules, which can support full degeneracy of 27 bp.

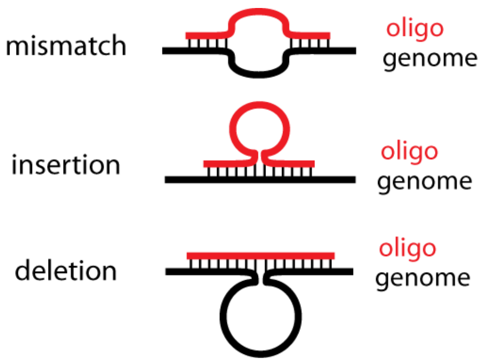
MAGE Automation

Automation instrumentation was constructed using the following major components:

Electroporator: ECM 630	BTX Technologies Inc (MA, USA)
Digital controllers: RS-232 serial modules	Superlogics Inc. (MA, USA)
Syringe pumps: Cavro XLP600 9-port	Tecan Group Ltd. (NC, USA)
Solenoid valves: Miniature Rocker Isolation Valves	Central Distribution Sales (NH, USA)
Temperature controller: CNI-3233-C24	Omega Engineering Inc. (CT, USA)
Orbital shaker: Advanced 3500 Orbital Shaker	VWR International LLC (PA, USA)
Control system software: LabView	National Instruments (TX, USA)
Growth chamber system: custom manufactured	David Breslau Design Inc. (NH, USA)

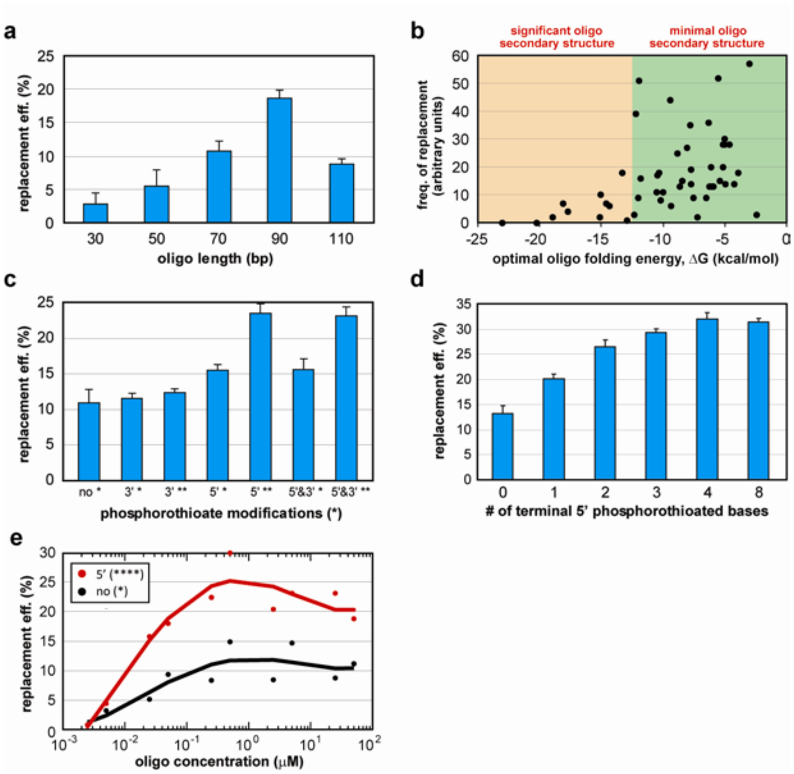
1. Costantino, N. & Court, D.L. Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. *Proc Natl Acad Sci U S A* **100**, 15748-15753 (2003).
2. Ellis, H.M., Yu, D., DiTizio, T. & Court, D.L. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci U S A* **98**, 6742-6746 (2001).
3. Mythili, E., Kumar, K.A. & Muniyappa, K. Characterization of the DNA-binding domain of beta protein, a component of phage lambda red-pathway, by UV catalyzed cross-linking. *Gene* **182**, 81-87 (1996).
4. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406-3415 (2003).
5. Wu, X.S. et al. Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2508-2513 (2005).
6. Chen, H., Bjerknes, M., Kumar, R. & Jay, E. Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia coli mRNAs. *Nucleic Acids Res* **22**, 4953-4957 (1994).

Supplementary Figure 1



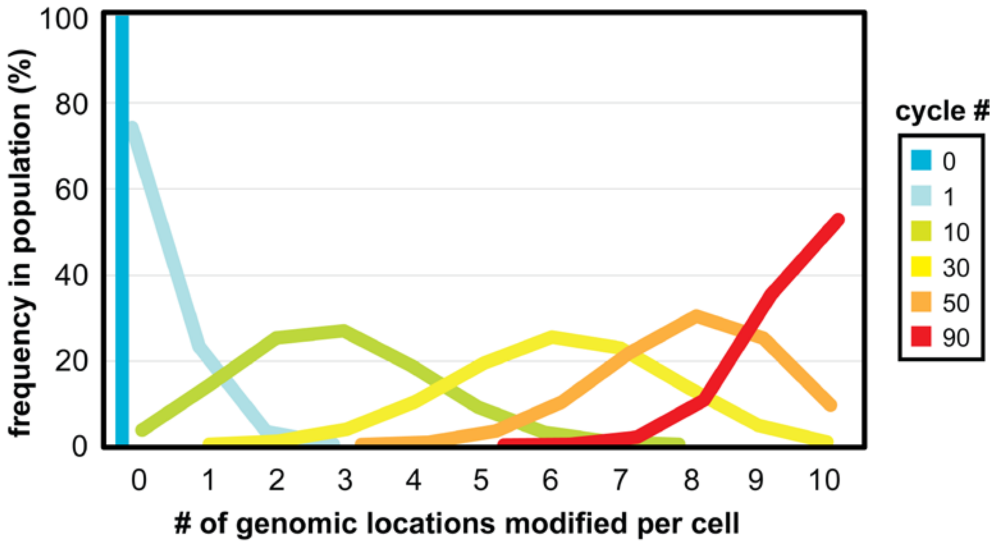
Supplementary Figure 1 | Diagram of the oligo-genome hybridization structure during mismatch, insertion, and deletion modifications.

Supplementary Figure 2



Supplementary Figure 2 | Characterization of the allelic replacement frequency in the MAGE strain (EcNR2) and its derivative (EcFI5) by screening for the introduction of a nonsense mutation in the *lacZ* gene or recovery of the *cmR* gene. a, Replacement efficiency as a function of oligonucleotide length. Oligos contain two phosphorothioated bonds at both the 3' and 5' termini. **b,** Predicted optimal folding energy ΔG of 90-mer oligos as a function of frequency of replacement. Oligos with ΔG < -12.5 kcal/mol are considered to have significant secondary structure that hinder allelic replacement. **c,** The effect of terminal phosphorothioated bonds on replacement efficiency. Oligos of 90 bp with 0, 1(*) or 2 (**) phosphorothioated bonds between bases at the 3', 5' or both 3' and 5' termini were tested. **d,** Replacement efficiency as a function of the number of consecutive terminal 5' phosphorothioated bonds of a 90mer oligo as measured by the introduction of a nonsense mutation in the *lacZ* gene. **e,** Replacement efficiency as a function of the concentration of 90-mer oligos containing no phosphorothioated bonds (in black) or 4 phosphorothioated bonds at the 5' terminus (in red). Solid lines represent data fitted using polynomial functions. Error bars, ±SD.

Supplementary Figure 3



Supplementary Figure 3 | Predicted distribution of genetic variants in a population that has undergone simultaneous allelic manipulation at 10 different genomic locations at 30% overall replacement efficiency. In this case, 10 different genes are simultaneously targeted for inactivation. Each colored solid line represents the histographic distribution of variants containing different numbers of knockouts (KO) across the population as a function of MAGE cycle number. As MAGE cycles increase, population evolves towards acquiring all 10 gene KO's. The population is binomially distributed according to the equation:

$$P(K,N) = \sum_{j=0}^{K} \binom{K}{j} \left((1-M)^{N}\right)^{(K-j)} \left(1-(1-M)^{N}\right)^{j},$$

where *K* is the number of loci simultaneously targeted, *N* is the number of MAGE cycles, and *M* is the mutation rate at any individual locus.

Supplementary Table 1

MAGE Parameters	Optimal Values
Oligo length	90 bp
Oligo concentration range	0.05 – 50 μM
Oligo stability	Four 5’ phosphorothioated bases
Oligo secondary structure	> –12.5 kcal/mol
Strand bias	Target lagging strand
Size of genetic modification	Predict efficiency using hybridization energy
Cycle time	2 – 2.5 hours

Supplementary Table 1 | Optimized parameters for maximal allelic replacement efficiency.

Supplementary Table 2

EcHW2a (KO's: none)		
idi	wild-type RBS	acatgtgagaaattatg
	optimized RBS	ggaaggggatgattatg
EcHW2b (KO's: Δ*gdhA*, Δ*ytjC*)		
dxs	wild-type RBS	ttaataggcccctgatg
	optimized RBS	taggaaatggtctgatg
EcHW2c (KO's: none)		
dxs	wild-type RBS	ttaataggcccctgatg
	optimized RBS	aaaaggaagaactgatg
ispA	wild-type RBS	ccggacaatgagtaatg
	optimized RBS	ggagaagggaagtaatg
EcHW2d (KO's: Δ*fdhF*)		
dxs	wild-type RBS	ttaataggcccctgatg
	optimized RBS	gtaaggagaagctgatg
EcHW2e (KO's: none)		
dxs	wild-type RBS	ttaataggcccctgatg
	optimized RBS	ggagaaggaaactgatg
idi	wild-type RBS	acatgtgagaaattatg
	optimized RBS	tgaggaataaaattatg
EcHW2f (KO's: Δ*ytjC*)		
dxs	wild-type RBS	ttaataggcccctgatg
	optimized RBS	tagagaagagactgatg
rpoS	wild-type RBS	gtaggagccaccttatg
	optimized RBS	gagaggatggacttatg
idi	wild-type RBS	acatgtgagaaattatg
	optimized RBS	aaaagaggttgattatg
dxr	wild-type RBS	actctggatgtttcatg
	optimized RBS	ttaagggtgtattcatg

Supplementary Table 2 | Optimized RBS sequences of strains EcHW2a-f

Programming cells by multiplex genome engineering and accelerated evolution

Supplementary Table 3 - Sequences of all oligos used
List generated by H.H.Wang, 2009

(note: * indicates phosphorothiolated bond)

Fig 2a	**introduce bp mismatches**	
	lacZ_mut_1	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGAGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATC*C*C
	lacZ_mut_2	C*A*TGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTAAGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACAT*C*C
	lacZ_mut_3	G*A*CCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTTAATGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA*C*A
	lacZ_mut_4	A*T*GACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTTAACACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAG*C*A
	lacZ_mut_6	A*T*TCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTCTAGATTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGC*C*A
	lacZ_mut_10	T*G*ATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGAGCTCTTCTAACCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCC*C*C
	lacZ_mut_12	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTCAATAACGGTTGCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATC*C*C
	lacZ_mut_18	G*A*AAACCCTGGCGTTACCCAACTTAATCGCCTTGCATAGGGATAACAGGGTAATAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC*G*C
	lacZ_mut_30	A*T*GATTACGGATTCACTGGCCGTCGTTTTATAAGCTTGGCATTCTAGATATCGTTAAGCCACCCAACTTAATCGCCTTGCAGCACATC*C*C
Fig 2b	**introduce bp insertions**	
	lacZ_ins_1	C*A*TGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGAGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACA*T*C
	lacZ_ins_2	C*A*TGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTAAGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCAC*A*T
	lacZ_ins_3	G*A*CCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTTAAGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCA*G*C
	lacZ_ins_4	A*T*GACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTTAACCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTT*G*C
	lacZ_ins_6	C*A*CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGTCTAGAGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTT*C*G
	lacZ_ins_8	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGAGCTCTTCTCTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGC*A*G
	lacZ_ins_10	T*G*ATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGAGCTCTTCTAACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTG*C*A
	lacZ_ins_12	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTCAATAACGGTTGGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCC*T*T
	lacZ_ins_16	G*A*TTACGGATTCACTGGCCGTCGTTTTACAACGTCGTTCTAGAGCTCTTCCTAGACTGGGAAAACCCTGGCGTTACCCAACTTAATCG*C*C
	lacZ_ins_18	G*G*CGTTACCCAACTTAATCGCCTTGCAGCACATCCCTAGGGATAACAGGGTAATCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCC*G*C
	lacZ_ins_30	A*T*GATTACGGATTCACTGGCCGTCGTTTTATAAGCTTGGCATTCTAGATATCGTTAAGCCCAACGTCGTGACTGGGAAAACCCTGGCG*T*T
Fig 2c	**introduce bp deletions**	
	lacZ_del_1	A*G*CGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGC*G*G
	lacZ_del_2	A*T*GACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA*C*A
	lacZ_del_3	C*C*ATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCC*C*C
	lacZ_del_4	T*G*ATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGCCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT*T*T
	lacZ_del_7	C*A*TGACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATC*C*C
	lacZ_del_10	A*T*GACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCC*T*T
	lacZ_del_17	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGG*C*G
	lacZ_del_26	A*T*GACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT*A*A
	lacZ_del_100	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCC*G*G
	lacZ_del_1000	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGG*T*C
	lacZ_del_3000	G*T*GAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTTGTCAAAAATAATAATAACCGGGCAGGCCATGTCTGCCCGTA*T*T
	lacZ_del_10000	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGCCGTCCTGCAGATGAATCTGCCACAGCCCCTCTTCGCCTGGTA*A*C
	lacZ_del_45000	T*T*ACCACATCAAATCATCAGGCACTTTGAAGTCGGCATACGGGTCGTATAACATCCCACATCACTATCAAGAAGATTATCTTTACCTC*T*G

			hyb dG (kcal/mol)
Fig 2d	**introduce bp mismatches**		
	lacZ_oligo_m1_v1(**)	G*G*AAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAA*T*C	-117.7
	lacZ_oligo_m1_v2(**)	G*G*AAACAGCTatgACCATGATTACGGATTCACTGGCAGTCGTTTGACAACGTAGTGACTGG GAAAACCCTGGCGTTACCCAACTTAA*T*C	-109.5
	lacZ_oligo_m1_v3(**)	G*G*AAACAGCTatgACCATGATTACGGATGCACTGGCCGTCGTTTTACAACGTCGTGACTGAGAAAACCCTGGCGTTACCCAACTTAA*T*C	-112.3
	lacZ_oligo_m1_v4(**)	G*G*AAACAGCTatgACCATGATTACGGATGCACTGGCCGTCGTTTGACAACGTCGTGACTGAGAAAACCCTGGCGTTACCCAACTTAA*T*C	-109.6
	lacZ_oligo_m1_v5(**)	G*G*AAACAGCTatgACCATGATTACGGATGCACTGGCAGTCGTTTGACAACGTAGTGACTGAGAAAACCCTGGCGTTACCCAACTTAA*T*C	-101.4
	lacZ_oligo_m1_v6(**)	G*G*AAACAGCTatgACCATGATTACGGATGCACGGGCAGTCCTTTGACACCGTAGTGCCTGAGAAAACCCTGGCGTTACCCAACTTAA*T*C	-84.5
	lacZ_oligo_m2_v1(**)	G*G*AAACAGCTatgACCATGATTACGGATTCACTGGTAATCGTTGGCCAACGGACTGACTGGGAAAACCCTGGCGTTACCCAACTTAA*T*C	-94.4
	lacZ_oligo_m2_v2*(**)	G*G*AAACAGCTatgACCATGATTACGGAGGAACTGGCCGTCGTTGGCCAACGTCGTGACTAATAAAACCCTGGCGTTACCCAACTTAA*T*C	-95.4
	lacZ_oligo_m2_v3(**)	G*G*AAACAGCTatgACCATGATTACGGAGGAACTGGTAATCGTTGGCCAACGGACTGACTTAAAAAACCCTGGCGTTACCCAACTTAA*T*C	-76.6
	lacZ_oligo_m2_v4(**)	G*G*AAACAGCTatgACCATGATTACGGAGGAATAAGTAATTAATGGCCTGAGGACTTAGTTAAAAAACCCTGGCGTTACCCAACTTAA*T*C	-64.0
	lacZ_oligo_m3_v1(**)	G*G*AAACAGATatgACCAGGATTACGGCTTCACTGGACGTCGTTTGACAACGTCATGACTGGGCAAACCCTGACGTTACCCCACTTAA*T*C	-87.2
	lacZ_oligo_m3_v2(**)	G*G*AAACAAAGatgACCCGAATTACGACGTCACTGTAAGTCGTTGGCCAACGTTAAGACTGGACCAACCCTTAAGTTACCTCCCTTAA*T*C	-41.2
	lacZ_oligo_m3_v3(**)	A*G*AAACAGCTaGgACCATGATTCCGGATTCACTAGCCGTCGTTTGACAACGTCGTAACTGGGAAAAACCTGGCGTTAACCAACTTAA*T*A	-94.3
	lacZ_oligo_m3_v4(**)	A*T*CAACAGCTCGAACCATGATGTAGGATTCACGAACCGTCGTTGGCCAACGTCGGACCTGGGAAACAACTGGCGTTTAACAACTTAC*G*A	-55.8

Fig 3	**introduce bp mismatches (degenerate oligos)**	
	lacZ_30mer_degen (**)	A*T*GATTACGGATTCACTGGCCGTCGTTTTACA NNNNNNNNNNNNNNNNNNNNNNNNNNNNNN CAACTTAATCGCCTTGCAGCACATC*C*C
	lacZ_6inters_degen (**)	A*T*GATTACGGATTCACTGGCCGTCGTTTTACANCGTCGNGACTGNGAAAANCCTGGNGTTACNCAACTTAATCGCCTTGCAGCACATC*C*C
	lacZ_6mer_degen(**)	A*T*GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGAC NNNNNNAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATC*C*C
Fig 3	**PCR primers for sequencing**	
	lacZ_pcr_seq_fprimer	gtaaaacgacggcaGGCAGTGAGCGCAACG
	lacZ_pcr_seq_rprimer	TTCTCCGTGGGAACAAACG
Fig 5	**introduce bp mismatches (degenerate oligos)**	
	dxs_oligo	g*c*ggactacatcatccagcgtaataaataaacaataagtaDDRRRRRDDDDctgatgagttttgatattgccaaatacccgaccctgg*c*a
	appY_oligo	a*t*tgacagatgaaaactacggagcaaacataatccatcttHHHHYYYYYHHaaaattaaaacaagttgatagtagtcaaataacaacc*a*a
	rpoS_oligo	c*a*tttaaatcatgaactttcagcgtattctgactcataagHHHHYYYYYHHccgtgatcccttgacggaacattcaagcaaaagcctg*g*t
	crl_oligo	t*c*aatctgctcttcgggtgtccactcggtaacgtcattgcHHHHYYYYYHHgtgatgcaactgttttaccaaattggcaaaatttagc*a*a
	elbA_oligo	a*g*aaataaacatcctaatgaaccatattaaataccgtgggDDRRRRRDDDDcaaatgaagtggatagtaattgacacggtaattcaac*c*t
	elbB_oligo	c*g*ccgcatccgctcagaattacgccaatttcttcattgtHHHHYYYYYHHcgcaattgactgaaacacattaattttttaataaaa*a*t
	yjiD_oligo	t*t*tcaggtaaaacagcctgaagtgattgtcgcatcatttHHHHYYYYYHHacgtttagcaacatatgcggcttgccctgaacaccgg*g*c
	idi_oligo	t*t*cactcttcaattatctataatgatgagtgatcagaattDDRRRRRDDDDattatgcaaacggaacacgtcattttattgaatgcac*a*g
	ispC_oligo	t*c*gagccggtcgagcccagaatggtgagttgcttcatgaaHHHHYYYYYHHtgagacagaataaaaagcaaaacgccgccagccgatc*c*g
	ispD/F_oligo	g*a*accacggcgcaaacatccaaatgagtggttgccatgttHHHHYYYYYHHtgatttatcgattgttttgccccgcagactgtgcgcg*c*t
	purHD_oligo	t*c*gcgaaaaattcagctaacgctctctgtaatagtcaaatDDRRRRRDDDDaccatgcaacaacgtcgtccagtccgccgcgctctgc*t*c
	rnlA_oligo	t*t*caggggggtattattgtagagtttccccatatgtttctDDRRRRRDDDDgaaatgacaatcaggagttacaaaaacttaaatctgg*t*c
	yggT_oligo	t*g*caattttggtgcgcgtgattactctaaaaaataaggaDDRRRRRDDDDgccatgaatacgttgactttcctgctttcaacggtca*t*t
	ispE_oligo	c*a*ccgacggtggtcaacgcatcaagttaaaaatggataacDDRRRRRDDDDataatgcggacacagtggccctctccggcaaaactta*a*t
	ispG_oligo	t*t*gattttctacgttgaattggagcctggttatgcatgaaHHHHYYYYYHHgttacccgtctgttactgcgccggtgattgttcggca*t*t
	ispH_oligo	c*a*caaaaaccacgcgggttggccaacaggatctgcatgttHHHHYYYYYHHgccggatcgatttccagcacttcaatatcaaaatgaa*c*g
	ispA_oligo	g*c*tgtctgacaatgaagacgcctctctaaccccttttacaDDRRRRRDDDDgtaatggactttccgcagcaactcgaagcctgcgtta*a*g
	ycgZ_oligo	c*t*cccgcagaatctaaagtcactgaattttgatgcatgctHHHHYYYYYHHaaaaatgttaatcaatgttaagtgaagtaacaatagc*a*a
	ymgA_oligo	t*t*tcatattttatacgttcattatcagatgtcttcataacHHHHYYYYYHHtaacctgttaaataatcaggctgttgcattattcaaa*a*a
	ariR_oligo	c*a*gttattgcattatgaattgtagtatcttcaagcatggtHHHHYYYYYHHttacttatgaaattttaatgtattctgtttattttct*t*a
	gdhA_oligo	C*A*TATTCTCTGGAGTCATTCCTCAACCATGTC TAAAAGCGCGACCCGAATTAAACCGAGTTCGCGCAAGCCGTTCGTGAAGTAATGAC*C*A
	aceE_oligo	G*A*GCACGCTCAACACCTTCTTCACGGATGACCGA TTAGATCGCCTGGAGTCAGTCGCGAGTTTCGATCGGATCCACGTCATTTGGGAA*A*C
	fdhF_oligo	G*C*AAAATCGTCCGGGCGGAGGCAGCGCAGGGG TAAACCAACCAGGGTACCCTGTGACTGAAGGGTTATTATGGCTGGGACTTCATTAA*C*G
	ytjC_oligo	G*C*TCACCTTTGGCGGTCAGCGGGCTGTCAGACTGGCC CTAAATACGTCACTCGGCGTTCCACTGCGTTTCACCGTGGCGGACTAGGTA*T*A
Fig 5	**forward PCR primer for**	**Sequence**
	dxs_f	gtaaaacgacggcaACCAGCAACTTGGTAAAAGTACC
	appY_f	gtaaaacgacggcaAAAAATGTAACTTTGAAATAAGTTAGAATAAAAAACAA
	rpoS_f	gtaaaacgacggcaCGGGAACAACAAGAAGTTAAGG
	crl_f	gtaaaacgacggcaTCAGCCCGGAAGAGGAC
	elbA_f	gtaaaacgacggcaGAAGTCAATAATTTATTGCTTTCACAAAATCT
	elbB_f	gtaaaacgacggcaGCGCCAACACCAGGG
	yjiD_f	gtaaaacgacggcaCATGACTTTTTTCCTGAAAAGTCAGT
	idi_f	gtaaaacgacggcaCTCTCTATTCCTGTCATTTCTGACTG
	ispC_f	gtaaaacgacggcaAGACGACGATCGCCGT
	ispD/F_f	gtaaaacgacggcaCGAAACTTAAAGCGCGAAACG
	purHD_f	gtaaaacgacggcaGTAATTTCTGTATTTTGCCCACGG
	rnlA_f	gtaaaacgacggcaGAGTTGCTGCTTGGTGAGAT
	yggT_f	gtaaaacgacggcaAGGCAGTTTGAAGTTGCAC
	ispE_f	gtaaaacgacggcaTTACCGGGTGATGCAACC
	ispG_f	gtaaaacgacggcaCGTAAAGACGGTAATTTGAACTTAACC
	ispH_f	gtaaaacgacggcaCCGTGAATTTATGGATGCAGG
	ispA_f	gtaaaacgacggcaCTGAGCGAGCTGGAACAG
	ycgZ_f	gtaaaacgacggcaAAAAGTTGAACACTAGTTGGCGA

	ymgA_f	gtaaaacgacggcaAGGAAACTCTCGGGGAAATTG
	ariR_f	gtaaaacgacggcaTAAGCAATTACTCGTAATGAAATCGG
	gdhA_f	gtaaaacgacggcaGCTTACCGTCACATTCTTGATG
	aceE_f	gtaaaacgacggcaGCAACTAAACGTAGAACCTGTCT
	fdhF_f	gtaaaacgacggcaCCGTAATATCAGGGAATGACCC
	ytjC_f	gtaaaacgacggcaTGGCCTGAATTTTAGCGGG
Fig 5	**reverse PCR primer for sequencing**	
	dxs_r	CGATTTTGTCGCGGCG
	appY_r	CAAACAGAGAAAAAACAAGCGATAATG
	rpoS_r	CGATCATCCGGCGGC
	crl_r	GTCGTTCAACCACTTCAGTGT
	elbA_r	TATTATGTAAACATTTATTTGTTATTTTGCTTTCTCC
	elbB_r	CAAACCCCCCCGGC
	yjiD_r	CCAGAGTGTGGCAGTACG
	idi_r	CAGGAGGCGTAATTTCCACG
	ispC_r	GATCAACATCCTCAAGCGCT
	ispD/F_r	CGGCTTTCAGACCTGCC
	purHD_r	GGATCTGATGTTCTTCCATAATGGC
	rnlA_r	CAGCAGGATTGATGGTTTCAAAAA
	yggT_r	GTAAACCGGCAATCCAGATGA
	ispE_r	TGTCAATGCTGATATTCGCACC
	ispG_r	TCAGCGCAATGCGATAGTC
	ispH_r	AGGTGGCATCAAACACCG
	ispA_r	GACGCAGATCGTCATCATCC
	ycgZ_r	TCTTCATAACGACTTCATAGCTAACC
	ymgA_r	TACTTATGAAATTTTAATGTATTCTGTTTATTTTCTTACC
	ariR_r	TATACTTAATAATTAGAAGTTACATATCATCAGCTG
	gdhA_r	CCTTTACCACCGCCCATC
	aceE_r	CGTGCACGGAAGAAGTGG
	fdhF_r	CCAATAACGGCGCGC
	ytjC_r	ACGCGATCGCTGAGC
Suppl Fig 2a	**introduce bp mismatches**	
	cat_fwd_stop	GCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGAACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAA
	cat_fwd_restore	GCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAA
	cat_fwd_re110*	T*A*TCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAA*A*G
	cat_fwd_re90*	G*C*ATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTA*A*A
	cat_fwd_re70*	G*A*ACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTAC*G*G
	cat_fwd_re50*	A*G*GCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGC*T*G
	cat_fwd_re30*	G*T*CAGTTGCTCAATGTACCTATAACCAG*A*C
Suppl Fig 2b	**introduce 1 bp mismatch**	
	argO_oligo	a*g*gcaagcttagcgcctctgttttattttccatcagatagcgcTtaactgaacaaggcttgtgcatgagcaataccgtctctcgcca*g*c
	atpE_oligo	t*a*gttaacgttctgatattgctctttaaataaaagcaacgcttaTtacgcgacagcgaacatcacgtacagacccagacctacagcga*t*c
	b1146_oligo	c*g*attagcctcatcgttctgtggttaaaaattgaaagtgttctgTtaatctttcggatagatatccggtcttaagtcagatttcgtaa*t*t
	b1228_oligo	c*a*atcaaatattctgcccatgcggggaaggatgagaagcttcgaTtaaggttcggctcgagcgtcagcgagagagcgttgccgcaggc*a*a
	b1578_oligo	c*g*ggcttgtaatgggttaagtgataacagatgtctggaaatataAgggcaaatccacaggtgaactacgctcctctgtttgttacgca*a*t
	b4273_oligo	g*t*ataaaatatccgtattcatatcagcacaaggtggatttgcccTtatatttccagacatctgttatcacttaacccattacaagccc*g*c
	coaD_oligo	c*t*gccggagaatgtccatcaggcgctgatggcgaagttagcgtaAcgtttatgccggatggtatgccatccggcgcgcatgaattact*t*c
	fabH_oligo	c*c*agggaacacaaatgcaaattgcgtcatgttttaatccttatcTtagaaacgaaccagcgcggagccccaggtgaatccaccgccaa*a*g
	fliN_oligo	g*a*tatcattactccgtctgagcgaatgcgccgcctgagccgttaAtgatgaataaccacgctactgtgcaatcttccgcgccggtttc*t*g
	frlD_oligo	a*c*ggcgtgcgcggcgaaaaccattcagtaccacggtgcctggtaAgtataacgttggcgtgagcatcttcacgccaacgtgctgttac*t*t
	frmR_oligo	t*c*cgttgacgacactattgaactggttcgtgcctatcttaaataActgaatctattaccatattgaggaagagcgagagatgaaatca*c*g
	hda_oligo	t*t*cggataaggcgttcgcgccgcatccgacaataaacaccttatTtacaacttcagaatttctttcacaaacggaatggtcagcttac*g*t
	hycI_oligo	t*c*accccatcaagaacatccctgtcctgattccttaatgaaaaaTtactcttcttccaccgctaactgcgcgaagccgccatttcctt*c*c
	hypB_oligo	a*t*ggaccagtggctgaactggctggagacacagcgatgtgcataAgcgttcccggccagatccgcaccattgacggcaaccaggcgaa*a*g
	ilvA_oligo	g*c*gcgggtaggcctgataagcgaagcgctatcaggcatttttccTtaacccgccaaaaagaacctgaacgccgggttattggtttcgt*c*g
	kgtP_oligo	g*c*cggtgcaagcaccggctataccgtctggcaactgacccgtcaTtaaagacgcatccccttccctttgcgatgtagcatcaaagaaa*c*c
	lolA_oligo	t*a*ttatccgaaaaatcgagcgacagattgctcactcaggtgcctTtacttacgttgatcatctaccgtgacgccttgcggcggggtga*a*g
	lon_oligo	a*g*cctgccagccctgtttttattagtgcattttgcgcgaggtcaTtattttgcagtcacaacctgcataccagacggttcattttgca*g*c
	lplA_oligo	g*a*gctacgggagttatcggcatggatggcgggggctgtaaggtaAttacccgcccatgcgggcaactttctcttcgatttgccggatt*t*t
	lpxK_oligo	a*g*cgacattcatgactccatcaatcgaacgctgccgcggcgtaaTtagttgccagaagccagcaaggttagttgcgtaagcagtttcg*c*t
	marA_oligo	g*g*cgaatcgcgctttttacatccattaaatcattacaacagctaAttgaaaacgtgacaacgtcactgaggcaatcatgaaaccactt*t*c
	mreC_oligo	a*g*agagccagattacccagcgtccctggctacgatagctcgccaTtattgccctcccggcgcacgcgcaggcggttgagcaccacttt*g*c
	murF_oligo	g*t*taaagccggaataatatttgaccaaatgttcggccagccaaaTtaacatgtcccattctcctgtaaagcgcgtactacctcttcca*t*g
	paaX_oligo	c*g*gatggacaaaactctcttctggcacaaccggagtcagaccgtTtatctgataaattggcataacgcctcctgttcaatattcaagc*c*g
	pbl_oligo	t*a*tgccccaaaatatatattgtatataccaggcttaatgaactaAacaatcgtaaggcttcagcaaaatgagcctaaagcctcttttt*t*t
	pgpA_oligo	c*c*tgatgagacgtgacaagcgtcacatcaggcatcggtgcacaaTtacgacagaatacccagcggccagtgatgaccgataaaataca*g*g
	plsX_oligo	a*t*ataccgtcacttgcaaactgcgagttcgctggcagcgtcctgTtaccgcagagttccgcttttgccaccgtccagcagctcaaaac*c*a
	priB_oligo	t*g*aaacggcagaacttgcgacgacggaaataacgtgccatatggTtagtctccagaatctatcaattcaatctgctcggcatgcaaaa*c*c
	priC_oligo	g*c*gctggaaaaaatcgaaaacaggttagcgcgtttaacccgctaAcaatggagagaatatgtcactggaaaatgcccctgacgatgtc*a*a
	ptsN_oligo	t*a*tcaaatcattacggataccgaaggtactccggatgaagcgtaAttattcggtaatgtctcttttagacgttgtgaggagaaacagt*a*c
	relA_oligo	a*c*agtatatatcaatctacattgtagatacgagcaaatttcggcTtaactcccgtgcaaccgacgcgcgtcgataacatccggcacct*g*g
	speG_oligo	c*a*caaaacaccgggtcagactctcctgaagccgaccgcacaataAcattaataatatcgatcgtatttttgatggtgtaaacccgtt*c*g
	sucB_oligo	g*a*taatgccttatccggtctacagtgcaggtgaaacttaaactaTtacacgtccagcagcagacgcgtcggatcttccagcaactctt*t*g
	sufA_oligo	t*t*tgacatcgtcagttgcttcagtattacgagacatagtaccgcTtataccccaaagctttcgccacagccacattcattctgggctt*t*a
	yaiS_oligo	t*a*tttcacagtcaaacttgctgcattaaaaaaacataaaagctaAgaacgacgcgattatatgcgtcatgatcgcctgcgtgcagttg*c*a
	ybaA_oligo	t*g*aattttgggctgacggcacgccgcacagcccgtcacgctacTtattcgtcgatgattgactcgaatccgccatagatcattcgct*t*g
	ybcF_oligo	a*c*gcgtcaagcgtcgcatcaggcacaaatgtctaatgcctacgaTtacagcgaaatacaggtccccgcttcgcccgccagcgtctctt*c*a
	ybjK_oligo	t*c*ggttcaaggttgatgggttttttgttatctaaaacttatctaTtaccctgcaacccctctcaaccatcctcaaaatctcctcgcgcg*a*t
	ybjR_oligo	g*a*gaaaactgccgcctggctacgcgaccacggaaaactgccgcgTtaatcctgcccgtatttctccagcaatgcttcggcaatcgcct*g*a
	ycbK_oligo	a*a*taatacgatagttcatactgcccctgtttcgttaagcaattaTtaccagtgccgtgctggcccggtatcaatatgcacaaagttac*t*a
	ycgX_oligo	c*t*tttatatattgaagatgtagaaaaataataaaccattaatctaActaaagttggatacttaagaaatgcttcataattcagtaaggc*a*t
	yejE_oligo	g*g*tgaagccgtccgcgacgcatttgatcctaataaggcggtgtaAcatgacgcaaactctgttagcgattgaaaatttgtcggtgggt*t*t
	yffB_oligo	g*g*tttcagtgattccagttatcagcaatttttccatgaggtgtaAtctatgtcgtgcccggttattgagctgacacaacagcttattc*g*c
	yfiA_oligo	t*c*ggtgaaagacgccaacttcgtcgaagaagttgaagaagagtaAtcctttatattgagtgtatcgccaacgcgccttcgggcgcgtt*t*t
	ygaU_oligo	t*g*acgccaattcccattatccagcaggcgatggctggcaattaaTtactcttccggaatacgcaacacttgccccggataaatttttat*c*c
	ygbA_oligo	t*g*atttaactgcaaattgccggacagatctgcctgtccggcataTtattcatgaggtttttttcggacgatatttttccggcagttctg*g*c
	ygiZ_oligo	t*t*ctctgtctatgagagccgttaaaacgactctcatagattttaTtaatagcaaaatataaaccgtccccaaaaaagccaccaaccac*a*a
	yhcO_oligo	c*t*tgtccgacgatattgcccccggttcgggggcttttttttgcgTtaatgacgaacattaaaacgcaaatgcccttccagctcttcct*c*t
	yibA_oligo	a*a*tgaaattataacttccgctattgataagctgaagcgtttcataAcgtgagttgcctatgcacagtgggggattcccgccggcacggt*g*c
	yjjV_oligo	a*t*tgtgcgccaactgccggatgcggcgtgaacgccttatccggcTtacggcacgttaaacaacgtatacgtgttattaagcaacgctt*g*c
	ynfA_oligo	g*c*tgacacttcctcttatcttattgataaaatggatttatgttcTtacgtgcgcccccagcccgcaacaatgatcaacatgccgcaaa*g*c
	ypjC_oligo	a*a*accaacatgaaaatcactttttcttatattatcgaaaaggttTtattcatttcttttagcgcattcaaaaaactgatcggcattat*t*t
	yqiB_oligo	a*g*ggttaacaggctttccaaatggtgtccttaggtttcacgacgTtaataaaccggaatcgccatcgctccatgtgctaaacagtatc*g*c
Suppl Fig 2c&d	**introduce bp mismatches**	
	lacZ_oligo_m1_v1(no *)	GGAAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
	lacZ_oligo_m1_v1(3'*)	GGAAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAAT*C
	lacZ_oligo_m1_v1(5'*)	G*GAAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
	lacZ_oligo_m1_v1(*)	G*GAAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAAT*C
	lacZ_oligo_m1_v1(3'**)	GGAAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAA*T*C
	lacZ_oligo_m1_v1(5'**)	G*G*AAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
	lacZ_oligo_m1_v1(**)	G*G*AAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAA*T*C
	lacZ_oligo_m1_v1(5'***)	G*G*A*AACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
	lacZ_oligo_m1_v1(5'****)	G*G*A*A*ACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
	lacZ_oligo_m1_v1(5'*[8])	G*G*A*A*A*C*A*G*CTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
Suppl Fig 2e	**introduce bp mismatches**	
	lacZ_oligo_m1_v1(no *)	GGAAACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
	lacZ_oligo_m1_v1(5'****)	G*G*A*A*ACAGCTatgACCATGATTACGGATTCACTGGCCGTCGTTTGACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC