Supplementary material for:

Design and experimental evaluation of a minimal, innocuous watermarking strategy to distinguish near-identical DNA and RNA sequences

Francine J. Boonekamp[1#], Sofia Dashko[1#], Donna Duiker[1], Thies Gehrmann[2], Marcel van den Broek[1], Maxime den Ridder[1], Martin Pabst[1], Vincent Robert[2], Thomas Abeel[1], Eline D. Postma[1], Jean-Marc Daran[1], Pascale Daran-Lapujade[1,#]

Table S1 – Codon usage in *S. cerevisiae*

Codons used for codon optimization and their abundance (data from https://www.yeastgenome.org/)

Amino acid	codon	Total nb[a]	/1000[b]	Fraction[c]
Ala	GCG	17988	6.16	0.11
	GCA	47538	16.27	0.3
	GCT	59300	20.29	0.37
	GCC	35410	12.12	0.22
Arg	AGG	27561	9.43	0.21
	AGA	61537	21.06	0.48
	CGG	5299	1.81	0.04
	CGA	9050	3.1	0.07
	CGT	18272	6.25	0.14
	CGC	7644	2.62	0.06
Gly	GGG	17673	6.05	0.12
	GGA	32723	11.2	0.23
	GGT	66198	22.66	0.46
	GGC	28522	9.76	0.2
Leu	TTG	77261	26.44	0.28
	TTA	77615	26.56	0.28
	CTG	31054	10.63	0.11
	CTA	39440	13.5	0.14
	CTT	35753	12.24	0.13
	CTC	16086	5.51	0.06
Pro	CCG	15778	5.4	0.12
	CCA	51993	17.79	0.41
	CCT	39685	13.58	0.31
	CCC	20139	6.89	0.16
Ser	AGT	42657	14.6	0.16
	AGC	29003	9.93	0.11
	TCG	25381	8.69	0.1
	TCA	55725	19.07	0.21
	TCT	68207	23.34	0.26
	TCC	40972	14.02	0.16
Thr	ACG	23766	8.13	0.14
	ACA	53147	18.19	0.31
	ACT	59096	20.23	0.34
	ACC	36395	12.46	0.21
Val	GTG	31266	10.7	0.19
	GTA	35397	12.11	0.22
	GTT	62735	21.47	0.39
	GTC	32738	11.2	0.2

[a] Total number of occurrences of this codon in the yeast genome

[b] Number of occurrences of this codon per 1000 codons in the yeast genome

[c] Fraction of occurrence of this codon usage from the set of codons representing the same amino acid

Table S2 – Watermarked glycolytic genes

List of watermarks introduced in the glycolytic and fermentative genes and of the resulting change in codon usage

Gene	Position (bp)	Native codon	WM codon	Fraction of native codon[a]	Fraction of WM codon[b]	Change in fraction (a-b)
TPI1	375	TTG	TTA	0.28	0.28	0
	408	GCC	GCA	0.22	0.3	0.08
	420	TTG	TTA	0.28	0.28	0
	441	TTG	TTA	0.28	0.28	0
	453	TTG	TTA	0.28	0.28	0
	645	AGC	AGT	0.11	0.16	0.05
	651	GCC	GCA	0.22	0.3	0.08
	657	ACC	ACA	0.21	0.31	0.1
	678	GTC	GTA	0.2	0.22	0.02
	693	GTC	GTA	0.2	0.22	0.02
	708	TTG	TTA	0.28	0.28	0
	Sum of all substitutions[c]					**0.35**
FBA1	558	GTT	GTC	0.39	0.2	-0.19
	573	GCT	GCC	0.37	0.22	-0.15
	588	TTG	TTA	0.28	0.28	0
	618	GTC	GTT	0.2	0.39	0.19
	630	TTG	TTA	0.28	0.28	0
	1005	GTT	GTC	0.39	0.2	-0.19
	999	GTC	GTT	0.2	0.39	0.19
	1023	ACC	ACA	0.21	0.31	0.1
	1041	ACC	ACA	0.21	0.31	0.1
	1065	ACC	ACA	0.21	0.31	0.1
	Sum of all substitutions					**1.21**
GPM1	381	TCT	TCA	0.26	0.21	-0.05
	393	TCT	TCA	0.26	0.21	-0.05
	435	GTC	GTA	0.2	0.22	0.02
	447	ACT	ACA	0.34	0.31	-0.03
	456	TTG	TTA	0.28	0.28	0
	462	TTG	TTA	0.28	0.28	0
	639	GTC	GTA	0.2	0.22	0.02
	660	TTG	TTA	0.28	0.28	0
	699	GCT	GCC	0.37	0.22	-0.15
	702	GCC	GCT	0.22	0.37	0.15
	726	GCC	GCA	0.22	0.3	0.08
	Sum of all substitutions					**0.55**
HXK2	735	TCC	TCA	0.16	0.21	0.05

	750	CTA	CTT	0.14	0.13	-0.01
	762	CTA	CTT	0.14	0.13	-0.01
	795	GCC	GCA	0.22	0.3	0.08
	816	TCC	TCA	0.16	0.21	0.05
	1365	CCT	CCC	0.31	0.16	-0.15
	1380	TCC	TCA	0.16	0.21	0.05
	1392	GCC	GCA	0.22	0.3	0.08
	1413	GCC	GCA	0.22	0.3	0.08
	1440	TCC	TCA	0.16	0.21	0.05
	Sum of all substitutions					**0.61**
PDC1	885	GTC	GTA	0.2	0.22	0.02
	867	TTG	TTA	0.28	0.28	0
	882	ACC	ACA	0.21	0.31	0.1
	909	ACC	ACA	0.21	0.31	0.1
	933	TCC	TCA	0.16	0.21	0.05
	1593	TCT	TCA	0.26	0.21	-0.05
	1626	GTC	GTA	0.2	0.22	0.02
	1647	TTG	TTA	0.28	0.28	0
	1665	TTG	TTA	0.28	0.28	0
	1677	ACC	ACA	0.21	0.31	0.1
	Sum of all substitutions					**0.44**
PFK1	1506	TTA	TTG	0.28	0.28	0
	1512	ACT	ACA	0.34	0.31	-0.03
	1515	CTA	CTT	0.14	0.13	-0.01
	1545	CTG	CTT	0.11	0.13	0.02
	1554	ACC	ACA	0.21	0.31	0.1
	1563	ACT	ACA	0.34	0.31	-0.03
	1575	TTA	TTG	0.28	0.28	0
	2880	GCT	GCC	0.37	0.22	-0.15
	2907	CTG	CTT	0.11	0.13	0.02
	2910	TCC	TCA	0.16	0.21	0.05
	2937	GTA	GTG	0.22	0.19	-0.03
	2940	GCC	GCA	0.22	0.3	0.08
	2946	TTA	TTG	0.28	0.28	0
	2949	GCC	GCA	0.22	0.3	0.08
	Sum of all substitutions					**0.6**
PFK2	1461	ACT	ACA	0.34	0.31	-0.03
	1479	CGT	CGA	0.14	0.07	-0.07
	1494	TTA	TTG	0.28	0.28	0
	1503	CTT	CTA	0.13	0.14	0.01
	1530	TCC	TCA	0.16	0.21	0.05
	2775	TTG	TTA	0.28	0.28	0
	2799	TCC	TCA	0.16	0.21	0.05
	2817	GTC	GTA	0.2	0.22	0.02
	2841	CTC	CTG	0.06	0.11	0.05

	2862	GGA	GGC	0.23	0.2	-0.03
		Sum of all substitutions				**0.31**
PGI1	837	GTC	GTA	0.2	0.22	0.02
	855	GTC	GTA	0.2	0.22	0.02
	861	TCG	TCC	0.1	0.16	0.06
	882	GCC	GCA	0.22	0.3	0.08
	894	GGC	GGA	0.2	0.23	0.03
	1563	GTC	GTA	0.2	0.22	0.02
	1581	GGC	GGA	0.2	0.23	0.03
	1599	TCC	TCA	0.16	0.21	0.05
	1611	TCT	TCA	0.26	0.21	-0.05
	1629	ACC	ACA	0.21	0.31	0.1
		Sum of all substitutions				**0.46**
PGK1	630	TTA	TTG	0.28	0.28	0
	639	GCC	GCA	0.22	0.3	0.08
	663	TTG	TTA	0.28	0.28	0
	687	GTC	GTA	0.2	0.22	0.02
	693	TCT	TCA	0.26	0.21	-0.05
	1164	TCC	TCA	0.16	0.21	0.05
	1176	ACT	ACA	0.34	0.31	-0.03
	1200	TTA	TTG	0.28	0.28	0
	1218	TTG	TTA	0.28	0.28	0
	1239	TCC	TCA	0.16	0.21	0.05
		Sum of all substitutions				**0.28**
PYK1	762	TTG	TTA	0.28	0.28	0
	789	GCC	GCA	0.22	0.3	0.08
	819	GCC	GCA	0.22	0.3	0.08
	837	GTC	GTA	0.2	0.22	0.02
	849	TTG	TTA	0.28	0.28	0
	1419	TTG	TTA	0.28	0.28	0
	1443	TCC	TCA	0.16	0.21	0.05
	1461	GCC	GCA	0.22	0.3	0.08
	1476	TCC	TCA	0.16	0.21	0.05
	1491	GTC	GTA	0.2	0.22	0.02
		Sum of all substitutions				**0.38**
TDH3	516	TTG	TTA	0.28	0.28	0
	522	ACC	ACA	0.21	0.31	0.1
	555	ACT	ACA	0.34	0.31	-0.03
	570	TCC	TCA	0.16	0.21	0.05
	597	ACC	ACA	0.21	0.31	0.1
	903	TTG	TTA	0.28	0.28	0
	918	GTC	GTA	0.2	0.22	0.02
	930	TCC	TCA	0.16	0.21	0.05
	960	ACC	ACA	0.21	0.31	0.1
	975	TTG	TTA	0.28	0.28	0

		Sum of all substitutions				**0.45**
ADH1	531	TCC	TCA	0.16	0.21	0.05
	549	CTA	CTT	0.14	0.13	-0.01
	558	TTG	TTA	0.28	0.28	0
	573	GCC	GCA	0.22	0.3	0.08
	597	TTG	TTA	0.28	0.28	0
	939	TTG	TTA	0.28	0.28	0
	942	GTC	GTA	0.2	0.22	0.02
	966	GGC	GGA	0.2	0.23	0.03
	978	TTG	TTA	0.28	0.28	0
	1029	GTT	GTC	0.39	0.2	-0.19
	1032	GTT	GTC	0.39	0.2	-0.19
		Sum of all substitutions				**0.57**
ENO2	663	ACC	ACA	0.21	0.31	0.1
	678	TTG	TTA	0.28	0.28	0
	684	TTG	TTA	0.28	0.28	0
	726	GTC	GTA	0.2	0.22	0.02
	738	TTG	TTA	0.28	0.28	0
	750	TCC	TCA	0.16	0.21	0.05
	1212	TCC	TCA	0.16	0.21	0.05
	1242	TTG	TTA	0.28	0.28	0
	1260	TTG	TTA	0.28	0.28	0
	1272	GCT	GCC	0.37	0.22	-0.15
	1281	GCC	GCA	0.22	0.3	0.08
				Sum of all substitutions		**0.45**

[a] and [b], fraction of occurrence of this codon is used from the set of codons representing the same amino acid according to Table S1

[c], sum of all substitutions for each watermarked gene, calculated as $\sum_{i=1}^{n}(|a_i - b_i|)$

Table S3 – Comparing Alignment and markerQuant for differential quantification of watermarked and native glycolytic transcripts

Here, we show, for simulated data, the ability of markerQuant and STAR[1] to retrieve the real count of reads originating from either the native sequence, or the watermarked sequence. As for evaluating the watermarking method, we simulated RNA-Sequencing reads using the polyester R package[2]. We compared the % retrieval between STAR and markerQuant. For STAR, we removed reads that aligned to more than one transcript using samtools[3]. This results in a lower retrieval fraction of the generated reads using STAR than when using markerQuant. It shows that while STAR may also be able to discern between the native and watermarked transcripts, markerQuant is able to retrieve a higher fraction of the reads.

	Generated counts				Star Quant				MarkerQuant				Star % Retrieved					MarkerQuant % Retrieved				
	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4	Mean	S1	S2	S3	S4	Mean
ADH1_coding	158	195	751	838	74	79	317	392	85	95	377	452	46,8	40,5	42,2	46,8	44,1	53,8	48,7	50,2	53,9	51,7
ADH1_w	725	828	167	191	315	368	70	89	375	433	84	108	43,4	44,4	41,9	46,6	44,1	51,7	52,3	50,3	56,5	52,7
ADH3_coding	246	238	957	994	102	101	383	385	121	121	456	475	41,5	42,4	40,0	38,7	40,7	49,2	50,8	47,6	47,8	48,9
ADH3_w	977	917	288	248	401	352	120	99	458	430	141	116	41,0	38,4	41,7	39,9	40,3	46,9	46,9	49,0	46,8	47,4
ENO2_coding	237	238	935	971	90	70	318	339	106	93	381	416	38,0	29,4	34,0	34,9	34,1	44,7	39,1	40,7	42,8	41,8
ENO2_w	1055	1018	273	211	327	361	98	64	408	414	116	80	31,0	35,5	35,9	30,3	33,2	38,7	40,7	42,5	37,9	39,9
FBA1_coding	195	223	854	868	77	98	339	344	103	115	395	402	39,5	43,9	39,7	39,6	40,7	52,8	51,6	46,3	46,3	49,2
FBA1_w	793	822	271	219	321	336	99	87	369	402	118	102	40,5	40,9	36,5	39,7	39,4	46,5	48,9	43,5	46,6	46,4
GPM1_coding	127	130	633	635	74	81	370	384	91	93	446	450	58,3	62,3	58,5	60,5	59,9	71,7	71,5	70,5	70,9	71,1
GPM1_w	599	596	134	139	344	369	86	78	428	435	100	104	57,4	61,9	64,2	56,1	59,9	71,5	73,0	74,6	74,8	73,5
HXK2_coding	240	288	1084	1200	66	75	339	351	77	96	409	412	27,5	26,0	31,3	29,3	28,5	32,1	33,3	37,7	34,3	34,4
HXK2_w	1170	1162	277	218	342	337	82	61	400	408	95	71	29,2	29,0	29,6	28,0	29,0	34,2	35,1	34,3	32,6	34,0
PDC1_coding	360	368	1331	1451	80	84	318	346	93	99	392	411	22,2	22,8	23,9	23,8	23,2	25,8	26,9	29,5	28,3	27,6
PDC1_w	1340	1343	307	323	336	323	88	86	397	390	104	101	25,1	24,1	28,7	26,6	26,1	29,6	29,0	33,9	31,3	31,0
PFK1_coding	633	663	2523	2413	92	60	315	332	100	86	370	391	14,5	9,0	12,5	13,8	12,5	15,8	13,0	14,7	16,2	14,9

PFK1_w	2087	2388	609	598	286	286	88	61	341	337	101	75	13,7	12,0	14,4	10,2	12,6	16,3	14,1	16,6	12,5	14,9
PFK2_coding	514	561	2367	2342	58	84	305	298	77	103	377	365	11,3	15,0	12,9	12,7	13,0	15,0	18,4	15,9	15,6	16,2
PFK2_w	2174	2316	608	542	273	307	78	92	316	358	87	104	12,6	13,3	12,8	17,0	13,9	14,5	15,5	14,3	19,2	15,9
PGI1_coding	381	275	1302	1202	94	59	300	287	123	72	365	332	24,7	21,5	23,0	23,9	23,3	32,3	26,2	28,0	27,6	28,5
PGI1_w	1228	1371	311	288	276	332	68	77	330	394	77	89	22,5	24,2	21,9	26,7	23,8	26,9	28,7	24,8	30,9	27,8
PGK1_coding	264	246	945	1015	100	84	327	352	112	105	381	410	37,9	34,1	34,6	34,7	35,3	42,4	42,7	40,3	40,4	41,5
PGK1_w	941	1019	245	173	329	329	78	51	382	379	95	64	35,0	32,3	31,8	29,5	32,1	40,6	37,2	38,8	37,0	38,4
PYK1_coding	284	317	1408	1314	81	94	389	370	100	107	473	431	28,5	29,7	27,6	28,2	28,5	35,2	33,8	33,6	32,8	33,8
PYK1_w	1283	1115	285	291	379	318	77	74	443	381	88	94	29,5	28,5	27,0	25,4	27,6	34,5	34,2	30,9	32,3	33,0
TDH3_coding	249	164	797	745	108	78	379	335	134	91	444	413	43,4	47,6	47,6	45,0	45,9	53,8	55,5	55,7	55,4	55,1
TDH3_w	817	861	214	219	374	401	85	108	444	473	108	126	45,8	46,6	39,7	49,3	45,3	54,3	54,9	50,5	57,5	54,3
TPI1_coding	144	138	643	523	94	82	400	339	112	99	477	399	65,3	59,4	62,2	64,8	62,9	77,8	71,7	74,2	76,3	75,0
TPI1_w	605	605	150	159	377	367	94	93	446	443	117	105	62,3	60,7	62,7	58,5	61,0	73,7	73,2	78,0	66,0	72,7
													Mean				**35,0**	**Mean**				**41,8**

Table S4 – Plasmids used in this study

Plasmid name	Relevant genotype	Source
pMEL10	*2 µm Amp^R URA3 SNR52p::gRNA-CAN1.Y::SUP4t*	[4]
pROS12	*2 µm Amp^R hphNT1SNR52p::gRNA-CAN1.Y gRNA-ADE2.Y::SUP4t*	[4]
pROS13	*2 µm Amp^R KanMX SNR52p::gRNA-CAN1.Y gRNA-ADE2.Y::SUP4t*	[4]
p426-SNR52pgRNA. CAN1.Y-SUP4t	*2 µm Amp^R URA3 SNR52p::gRNA-CAN1.Y::SUP4t*	Addgene
pUDR 529	*2 µm Amp^R hphNT1 SNR52p::gRNA-KlURA3::SUP4t*	This study
pUDR531	*2 µm Amp^R hphNT1 SNR52p::gRNA-TPI1::SUP4t*	This study
pUDR532	*2 µm Amp^R hphNT1 SNR52p::gRNA-PYK1::SUP4t*	This study
pUDR413	*2 µm Amp^R kanMX SNR52p::gRNA-CAN1.Y gRNA- RECYCLE SinLoG::SUP4t*	This study
pUDR314	*2 µm Amp^R AmdSYM SNR52p::gRNA-SGA1::SUP4t*	This study
pUDC073	*CEN6/ARS4, URA3, GAL1p-I -SCEI-CYC1t*	[4]
pYTK002	*Cam^{R}ConLS*	[5]
pYTK047	*Cam^{R}GFP*	[5]
pYTK072	*Cam^{R}ConRE*	[5]
pYTK078	*Cam^{R}cloNAT^R*	[5]
pYTK081	*Cam^{R}CEN6/ARS4*	[5]
pYTK083	*Cam^{R}Amp^R-ColE1*	[5]
pGGKd012	*CEN6/ARS4 cloNAT^R GFP*	This study, kindly provided by Arthur Gorter de Vries
Promoter and terminator part plasmids used in this study		
Plasmid name	Relevant genotype	Source
pUD565	*Cam^{R}GFP*	GeneArt
pGGKp096	*Cam^R pHXK2 Sc*	GeneArt
pGGKp097	*Cam^{R}tHXK2Sc*	GeneArt
pGGKp098	*Cam^R pPGI1Sc*	GeneArt
pGGKp099	*Cam^R tPGI1Sc*	GeneArt
pGGKp100	*Cam^R pPFK1Sc*	GeneArt
pGGKp101	*Cam^R tPFK1Sc*	GeneArt
pGGKp102	*Cam^R pPFK2Sc*	GeneArt

pGGKp103	Cam^R tPFK2Sc	GeneArt
pGGKp104	*Cam^R pFBA1Sc*	GeneArt
pGGKp105	*Cam^R tFBA1Sc*	GeneArt
pGGKp106	*Cam^R tTDH3Sc*	GeneArt
pGGKp107	*Cam^R tGPM1Sc*	GeneArt
pGGKp108	*Cam^R tPYK1Sc*	GeneArt
pGGKp109	*Cam^R pPDC1Sc*	GeneArt
pGGKp110	*Cam^R tPDC1Sc*	GeneArt
pGGKp111	*Cam^R pADH1Sc*	GeneArt
pGGKp114	*Cam^R pTPI1Sc*	GeneArt
pGGKp115	*Cam^R tTPI1Sc*	GeneArt
pGGKp116	*Cam^R pPGM1Sc*	GeneArt
pGGKp117	*Cam^R pPYK1Sc*	GeneArt
Watermarked coding sequence part plasmids used in this study		
Plasmid name	Relevant genotype	Source
pGGKp137	*Cam^R ADH1_*Sc*	GeneArt
pGGKp139	*Cam^R ENO2_*Sc*	GeneArt
pGGKp140	*Cam^R FBA1_*Sc*	GeneArt
pGGKp141	*Cam^R PGM1_*Sc*	GeneArt
pGGKp142	*Cam^R HXK2_*Sc*	GeneArt
pGGKp143	*Cam^R PDC1_*Sc*	GeneArt
pGGKp144	*Cam^R PFK1_*Sc*	GeneArt
pGGKp145	*Cam^R PFK2_*Sc*	GeneArt
pGGKp146	*Cam^R PGI1_*Sc*	GeneArt
pGGKp147	*Cam^R PGK1_*Sc*	GeneArt
pGGKp148	*Cam^R PYK1_*Sc*	GeneArt
pGGKp149	*Cam^R TDH3_*Sc*	GeneArt
pGGKp150	*Cam^R TPI1_*Sc*	GeneArt
Watermarked expression plasmids used in this study		
pUDC212	*CEN6/ARS6 AmpR pFBA1Sc-FBA1_*Sc-tFBA1Sc*	This study
pUDC213	*CEN6/ARS6 AmpR pPGM1Sc-PGM1_*Sc-tPGM1Sc*	This study
pUDC214	*CEN6/ARS6 AmpR pHXK2Sc-HXK2_*Sc-tHXK2Sc*	This study
pUDC215	*CEN6/ARS6 AmpR pPDC1Sc-PDC1_*Sc-tPDC1Sc*	This study
pUDC216	*CEN6/ARS6 AmpR pPFK1Sc-PFK1_*Sc-tPDC1Sc*	This study
pUDC217	*CEN6/ARS6 AmpR pPFK2Sc-PFK2_*Sc-tPFK2Sc*	This study
pUDC219	*CEN6/ARS6 AmpR pPGK1Sc-PGK1_*Sc-tPGK1Sc*	This study
pUDC220	*CEN6/ARS6 AmpR pPYK1Sc-PYK1_*Sc-tPYK1Sc*	This study
pUDC222	*CEN6/ARS6 AmpR pTPI1Sc-TPI1_*Sc-tTDH3Sc*	This study
pUDC229	*CEN6/ARS6 AmpR pADH1Sc-ADH1_*Sc-tADH1Sc*	This study
pUDC230	*CEN6/ARS6 AmpR*	This study

	pTDH3Sc-TDH3_*Sc-tTDH3Sc	
pUDC231	*CEN6/ARS6 AmpR pENO2Sc-ENO2_*Sc-tENO2Sc*	This study
pUDC232	*CEN6/ARS6 AmpR pPGI1Sc-PGI1_*Sc-tPGI1Sc*	This study

Table S5 – Whole genome sequencing of the constructed strains

List of mutations in glycolytic expression cassettes and helper elements in IMX1770, IMX1771 and IMX2028 as compared to the parental strain IMX589.

	Fragment	Location	Mutation
IMX1770	*PFK1* promoter	-581	A > AT
	PGK1 terminator	211	G > GA
	ENO2 terminator	280	C > A
	HIS3 promoter	-132	C > CA
IMX1771	*GPM1* promoter	-459	CTATATATATA > C
	HIS3 terminator	114	C > CA
IMX2028	*GPM1* promoter	-459	CTATATATATA > C
	HIS3 terminator	114	C > CA
	FBA1 terminator*	294	CA>C
	PFK2 promoter*	-22	A>G

* In watermarked glycolysis cassette.

Table S6 – gRNA sequence for selective gene editing

Gene	Native gRNA sequence	Watermarked gRNA sequence
PGK1	TTGAATCTTGTCAGCAACCTT**G**G	TTGAATCTTGTCAGCAACCTT**T**G
PYK1	ACTTAGCAAT**C**AATTTCTTTTG**G**	ACTTAGCAAT**T**AATTTCTTTTG**T**
PDC1	GTGGAATTCGACAATGTTCTT**G**G	GTGGAATTCGACAATGTTCT**T**G
TPI1	AACAACATC**C**AAAGTCTTACC**G**G	AACAACATC**T**AAAGTCTTACC**T**G

For each native glycolytic gene targeted for DNA editing, the PAM is underlined. In The corresponding watermarked gene, the bases modified to prevent editing by Cas9 are shown in red bold font. Only gRNA's for *PYK1* and *TPI1* were used in this study.

Table S7 – List of primers used in this study

A. Primers used for construction of IMX1338

Fragment	Primer name	Primer number	Primer sequence
gRNA *S. pombe His5*	10904_Sphis5 _targetRNA FW	10904	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA GATAAATGATCATAAACAACCATGGGTAGGAGTTTTAGA GCTAGAAATAGCAAGTTAAAATAAG
Repair with *GAL1p - -I SceI – ADH1t*	10708_glk1oh _prGAL1_FW	10708	aaaccacaacaccaccactaatacaactctatcatacac aagatgGCAGTGAGCGCAACGCAATTAATG
Repair with *GAL1p - -I SceI – ADH1t*	10709_glk1oh _terCYC1_RV	10709	gtacggtgggatacgtacacaaaccaaaaaaatgtaaaa agatcaCGACTCACTATAGGGCGAATTGG
S. pombe His5	His5 ORF fwd	6190	CCATGCGCGCGGCTACTAG
glk1 flank	GLK1RV1	1461	CCCGTTTCCGATGATATTG
glk1 flank	GLK1FW3	1553	AAAAACGGGAAATAACAATAACGAC
S. pombe His5	His5 ORF rv	6189	CTACAAAAGCCCTCCTACCCATGG

B. Primers used to add SHR sequences to the SinLoG fragments for assembly

Fragment	Primer name	Primer number	Primer Sequence
ARS418	ARS418 Fw + Tag CAN1	12953	GGTGTATGACTTATGAGGGTGAGAATGCGAAATGGCGTGGGAATGTGATTAA AGGTAATATGAAAGTTTATGTTTTTTCACTGGA
	ARS418 Rv + AH	12954	CTCAGCCTTAGCCAATATGATCATGTCGTTGCGTCTCGGACCATCTAGTCTA CTCTGAAGTTGAAAAGCAATATTAAAATATAGA
	ARS418-SGA1	14624	TTTTTCTCATCTCTTGGCTCTGGATCCGTTATCTGTTCTGTTACACAAG AAATCGTACATTGAAAGTTTATGTTTTTTCA

FBA1	FBA1 Fw + H	12955	GTCACGGGTTCTCAGCAATTCGAGCTATTACCGATGATGGCTGAGGCGTTAG AGTAATCTAACGTGAACAACAATACCAGCCTTC
	FBA1 Rv + Tag CAN1	12956	CTTCAGAGTAGACTAGATGGTCCGAGACGCAACGACATGATCATATTGGCTA AGGCTGAGAATGAGCTATCAAAAACGATAGATC
TPI1	TPI Fw + H	12957	AGATTACTCTAACGCCTCAGCCATCATCGGTAATAGCTCGAATTGCTGAGAA CCCGTGACAACGAAGACCCAGAGATGTTGTTGT
	TPI Rv + P	12958	CTGATAGTGCTGTAAGTCGCCTCCATCTTAGCAGAGCTGTCCCTGAATGCGT ACTCGTGATGAGTAACCCATATAGAGATCGTAC
PGK1	PGK1 Fw + Q	13223	GAGCTGAATGTATATGCTGCGGGATCATTGCACAGCTCTGAGAGCCCTGCAA CGCGATATTCTTTTTATTAACCTTAATTTTTAT
	PGK1 Rv + P	13224	TCACGAGTACGCATTCAGGGACAGCTCTGCTAAGATGGAGGCGACTTACAGC ACTATCAGAAATAATATCCTTCTCGAAAGCTTT
ADH1	ADH1 Fw + N	12961	TTCTAGGCTTTGATGCAAGGTCCACATATCTTCGTTAGGACTCAATCGTGGC TGCTGATCAACGAAGTCCAATGCTAGTAGAGAA
	ADH1 Rv + Q	13225	ATATCGCGTTGCAGGGCTCTCAGAGCTGTGCAATGATCCCGCAGCATATACA TTCAGCTCTTGTCCTCTGAGGACATAAAATACA
PYK1	PYK1 Fw + N	12963	GATCAGCAGCCACGATTGAGTCCTAACGAAGATATGTGGACCTTGCATCAAA GCCTAGAAAACGTGGTCAAACTTCAGAACTAAG
	PYK1 Rv + O	12964	ATACTCCCTGCACAGATGAGTCAAGCTATTGAACACCGAGAACGCGCTGAAC GATCATTCATAATCATGATAACCTTGAGGGAAG
TDH3	TDH3 Fw + A	13226	GTGCCTATTGATGATCTGGCGGAATGTCTGCCGTGCCATAGCCATGCCTTCA CATATAGTATACTAGCGTTGAATGTTAGCGTCA

	TDH3 Rv + O	12966	GAATGATCGTTCAGCGCGTTCTCGGTGTTCAATAGCTTGACTCATCTGTGCA GGGAGTATATCCTGGCGGAAAAAATTCATTTGT
ENO2	ENO2 Fw + A	12967	ACTATATGTGAAGGCATGGCTATGGCACGGCAGACATTCCGCCAGATCATCA ATAGGCACAACGGATGATGAAAACACTAAACGA
	ENO2 Rv + B	13227	GTTGAACATTCTTAGGCTGGTCGAATCATTTAGACACGGGCATCGTCCTCTC GAAAGGTGTAACGAAGACGTTACCAGCTGATTG
HXK2	HXK2 Fw + B	12969	CACCTTTCGAGAGGACGATGCCCGTGTCTAAATGATTCGACCAGCCTAAGAA TGTTCAACAACGACGCTGGTAAAGTACAGCTAC
	HXK2 Rv + C	12970	CTAGCGTGTCCTCGCATAGTTCTTAGATTGTCGCTACGGCATATACGATCCG TGAGACGTACTTGAACAATAAATACGAAATCCT
PGI1	PGI1 Fw + D	12971	AATCACTCTCCATACAGGGTTTCATACATTTCTCCACGGGACCCACAGTCGT AGATGCGTAACGTATTCTTAGTGGATAACATGC
	PGI1 Rv + C	12972	ACGTCTCACGGATCGTATATGCCGTAGCGACAATCTAAGAACTATGCGAGGA CACGCTAGTTTTAAACAGTTGATGAGAACCTTT
PFK1	PFK1 Fw + D	12973	ACGCATCTACGACTGTGGGTCCCGTGGAGAAATGTATGAAACCCTGTATGGA GAGTGATTAACGCGGCTAGTAAAAAAGAAAATT
	PFK1 Rv + J	12974	CGACGAGATGCTCAGACTATGTGTTCTACCTGCTTGGACATCTTCGCGTATA TGACGGCCATTCCATAGCTTAGTTTAATCAAGG
PFK2	PFK2 Fw + BP	12975	GAGATGACTGGGTCCACTCTTTCGTGTATTTCGAGAGAGCGATACGCATGTC TCCATCGTAACGATTCTCTGCTGCTTTGTTGCA
	PFK2 Rv + J	12976	GGCCGTCATATACGCGAAGATGTCCAAGCAGGTAGAACACATAGTCTGAGCA TCTCGTCGAAATCGTCTATATCACATATTCCAG

HIS3	HIS3 Fw + BP	13268	ACGATGGAGACATGCGTATCGCTCTCTCGAAATACACGAAAGAGTGGACCCA GTCATCTCTGCGGCATCAGAGCAGATTGTACTG
	HIS3 Rv + L	12978	GCCGTAGCTTCCGCAAGTATGCCGTAGTTGAAGAGCATTTGCCGTCGGTTCA GGTCATATGCATCTGTGCGGTATTTCACACCGC
URA3	URA3+BP	13271	ACGATGGAGACATGCGTATCGCTCTCTCGAAATACACGAAAGAGTGGACC CAGTCATCTCAACTGTCATCCTGCGTGAAGATTAA
	URA3 – Tag L	14623	GCCGTAGCTTCCGCAAGTATGCCGTAGTTGAAGAGCATTTGCCGTCGGTT CAGGTCATATAGTGTTGCACCGTGCCAATG
GPM1	GPM1 Fw + M	12979	ACGAGAGATGAAGGCTCACCGATGGACTTAGTATGATGCCATGCTGGAAGCT CCGGTCATAACGGTGATACTTTGACAGGAGCTA
	GPM1 Rv + L	12980	ATATGACCTGAACCGACGGCAAATGCTCTTCAACTACGGCATACTTGCGGAA GCTACGGCTATTGCTATAACATGTCATGTCACC
PDC1	PDC1 Fw + M	12981	ATGACCGGAGCTTCCAGCATGGCATCATACTAAGTCCATCGGTGAGCCTTCA TCTCTCGTAACGCATGCGACTGGGTGAGCATAT
	PDC1 Rv + AR	12982	TGACGAGATTTGAGAAGTCCCCAATATCGACTCGTGATGTGCCATGCGTGCT GTCAGTATACAGTGTTCCTTAATCAAGGATACC
ARS1211	ARS1211 Fw +AR	12983	ATACTGACAGCACGCATGGCACATCACGAGTCGATATTGGGGACTTCTCAAA TCTCGTCAGACATAGTATTTCGCAACCTTTCAG
	ARS1211 Rv + Tag CAN1	12984	GATGAGAAAAGTAAAGAATTGTATCCATTGCGCTCTTTCCCGACGAGAGTAA ATGGCGAGGACAGGCGTTTCTGTACCGCTGTTA
	ARS1211 SHR SGA1	13270	TCTCGCTTTTCTTTTATTTTTTTTTGTCTACAAACTCTGTAAAACTTCTT GTCTTATTTGGACAGGCGTTTCTGTACCGCTGTTA

C. Diagnostic primers used to confirm correct assembly of the SinLoG

Fragment	Primer name	Primer number	Primer Sequence
CAN1-ARS418	CAN1 PAGE rv	5822	TGAAGGAGTTTCAAATGCTTCTAC
	ARS418 Rv + AH	12954	CTCAGCCTTAGCCAATATGATCATGTCGTTGCGTCTCGGACCATCTAGTCTACTCTGAAGTTGAA AAGCAATATTAAAATATAGA
SGA1 – ARS418	11898_SeqFW_SGA	11898	CGCGGAAACGGGTATTAGGG
	ARS418 Rv + AH	12954	CTCAGCCTTAGCCAATATGATCATGTCGTTGCGTCTCGGACCATCTAGTCTACTCTGAAGTTGAA AAGCAATATTAAAATATAGA
CAN1-FBA1	CAN1 PAGE rv	5822	TGAAGGAGTTTCAAATGCTTCTAC
	FBA1 Flank left	5232	ATTTTACTCACGCTTGAAATTAACGGC
SGA1 – FBA1	11898_SeqFW_SGA	11898	CGCGGAAACGGGTATTAGGG
	FBA1 Flank left	5232	ATTTTACTCACGCTTGAAATTAACGGC
FBA1-TPI1	FBA1 Flank right FW	5026	CGTATTACGATAATCCTGCTGTC
	11213_gntK_RV_seq4	11213	ACCGAGGTGGTATCCGAGAG
TPI1-PGK1	FK115	2554	GTGGCATGTGAGATTCTCC
	pPGK1_fw 1	8739	GCTGCTACTGTTGCAAAGGC
PGK1-ADH1	10131_AL_ctrl_rv	10131	TGAGGGTAACATCAATTCAAGAAG

	ADH1 Flank left RV	5043	CGAGCAAATGCCTGCAAATCG
ADH1-PYK1	YAT2_2	7496	CAGCTCTGGAACAACGACATCTG
	11927_pPYK1_seq_5	11927	CATTTAGCTTGCCGGAAAAA
PYK1-TDH3	J_FW (PDH con- struct ctrl)	2914	GTCGTCATAACGATGAGGTGTTGC
	Sc_TDH3_T_RV	6493	GTGAATTTACTTTAAATCTTGCATTTAAATAAATT
TDH3-ENO2	Ptdh3 Ctrl Rv2	1523	GGGCATGTACGGGTTACAG
	ENO2 Flank left RV	5013	GGCGTGCGGTGTGTAGATGTATC
ENO2-HXK2	FK244	4034	CACCTTTCGAGAGGACGATG
	B-reverse	2364	ACGGAATAGAACACGATATTTGC
HXK2-PGI1	Probe HXK2 fw	2818	CACAAGCCAGAAAGGGTTCC
	G_RV_EF	5643	ACATTAAAGCTATATTATAAGCAAAG
PGI1-PFK1	m-PCR-HR2-RV	2669	GCGCGTGGCTTCCTATAATC
	Sc PUT1 SDC fw	3268	GTGGCGCTATTTCCCTTATC
PFK1-PFK2	11621_FW_PFK1_seq2	11621	CAATTGCGTGACGGTAGAAC
	Q-PFK2-FW	739	TTGGTGGTTTCGAAGCTTTTG
PFK2-HIS3	FG-forward	2371	TGTCTTACCCTGGACGGTATC

	FK074	1978	CGCTTTACTAGGGCTTTCTG
PFK2-URA3	FG-forward	2371	TGTCTTACCCTGGACGGTATC
	B-forward	2363	TTACCACCATCCAATGCAGAC
HIS3-GPM1	diag primer 3fw	5898	TCCCTCCACCAAAGGTGTTC
	GPM1 FW	3321	TTAGTTAGACACGGTCAATCC
URA3-GPM1	KlURA3 C	170	TTGGCTAATCATGACCCC
	GPM1 FW	3321	TTAGTTAGACACGGTCAATCC
GPM1-PDC1	GPM1_P_FW	6344	TATTGTAATATGTGTGTTTGTTTGGATTATTAAG
	PDC1_P_RV	6351	TTTGATTGATTTGACTGTGTTATTTTGCG
PDC1-CAN1	Q-PDC1-FW	757	GCTTTCGTCACCCCAATGG
	CAT2promDiagfw	7317	GCCGGTCACAACCCTCTTTC
PDC1-SGA1	Q-PDC1-FW	757	GCTTTCGTCACCCCAATGG
	Q_OutsideSGA_RV	7479	GGACGTTCCGACATAGTATC
ARS1211-CAN1	ARS1211 Fw +AR	12983	ATACTGACAGCACGCATGGCACATCACGAGTCGATATTGGGGACTTCTCAAATCTCGTCAGACAT AGTATTTCGCAACCTTTCAG
	CAT2promDiagfw	7317	GCCGGTCACAACCCTCTTTC
ARS1211-SGA1	ARS1211 Fw +AR	12983	ATACTGACAGCACGCATGGCACATCACGAGTCGATATTGGGGACTTCTCAAATCTCGTCAGACAT AGTATTTCGCAACCTTTCAG
	Q_OutsideSGA_RV	7479	GGACGTTCCGACATAGTATC

D. Primers used for confirmation of integration of the *KlURA3* repair fragment

Fragment	Primer name	Primer number	Primer Sequence
SGA1-SGA1	SeqFW_SGA	11898	CGCGGAAACGGGTATTAGGG
	Q_OutsideSGA_RV	7479	GGACGTTCCGACATAGTATC
SGA1-URA3	SeqFW_SGA	11898	CGCGGAAACGGGTATTAGGG
	B-forward	2363	TTACCACCATCCAATGCAGAC
URA3-SGA	KlURA3 C	170	TTGGCTAATCATGACCCC
	Q_OutsideSGA_RV	7479	GGACGTTCCGACATAGTATC

E. Primers used to amplify promoter and terminator sequences of the SinLoG

Fragment	Primer name	Primer number	Sequence
ENO2 promoter	tENO2 Fw	13228	GCATCGTCTCATCGGTCTCAATCCAGTGCTTTTAACTAAGAATTATTAG
	tENO2 Rv	13229	ATGCCGTCTCAGGTCTCACAGCTAACGAAGACGTTACCAGCTGATTG
TDH3 promoter	pTDH3 Fw	13230	AAGCATCGTCTCATCGGTCTCAAACGATACTAGCGTTGAATGTTAGCGTCA
	10754_TDH3 sc prom rv Ytoolkit	10754	TTATGCCGTCTCAGGTCTCACATA

PGK1 promoter	pPGK1 Fw	13231	AAGCATCGTCTCATCGGTCTCAAACGTCTTTTTATTAACCTTAATTTTTAT
	9422 PGK1 sc prom rev Ytoolkit	9422	TTATGCCGTCTCAGGTCTCACATA
ADH1 terminator	ADH1 sc term Ytoolkit	10769	AAGCATCGTCTCATCGGTCTCAATCCGCGAATTTCTTATGATTTATGATTTT
	10770_ADH1 term rv Ytoolkit	10770	TTATGCCGTCTCAGGTCTCACAGCCAACAGGTGTTGTCCTCTG
PGK1 terminator	10763_PGK1 term fw Ytoolkit	10763	AAGCATCGTCTCATCGGTCTCAATCCATTGAATTGAATTGAAATCGATAG
	10764_PGK1 term rv Ytoolkit	10764	TTATGCCGTCTCAGGTCTCACAGCCGAAATAATATCCTTCTCGAAAG

F. Primers used for construction of pUDR413

Primer name	Number	Primer Sequence
CRISPR RNA Recycl fw	6131	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCTTACAATATAGTGATAATCGGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
2mu inside rv	5975	AACGAGCTACTAAAATATTGCGAA
2mu inside fw-2	6296	GTTCTACAAAAATGCATCCCGAG
LP crRNA rv	5941	GCTGGCCTTTTGCTCACATG
AmdS CRISPR fw PAGE	6070	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCCATTTACAATATAGTGATAATCGGTTTTAGAGCTAG AAATAGCAAGTTAAAATAAGGCTAGTCCGTTATC
for f3h	3841	CACCTTTCGAGAGGACGATG

G. **Primers used for the construction and confirmation of the gRNA plasmid for selective removal** of the native glycolytic gene copy and their corresponding repair fragments

Primer name	Number	Primer Sequence
Construction of the gRNA plasmids		
TPI1_targetRNA fw	14517	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCAACAACATCCAAAGTCTTACGTTTT AGAGCTAGAAATAGCAAGTTAAAATAAG
TPI1_ targetRNA_dg rev	14518	GTAAGACTTTGGATGTTGTT
CDC19_targetRNA fw	14519	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCACTTAGCAATCAATTTCTTTGTTTT AGAGCTAGAAATAGCAAGTTAAAATAAG
CDC19_ targetRNA_dg rev	14520	AAAGAAATTGATTGCTAAGT
Repair DNA		

TPI1_repair oligo fw	14525	TGTTTGTATTCTTTTCTTGCTTAAATCTATAACTACAAAAAACACATACATAAACTAAAAGATTAATATAATTAT ATAAAAATATTATCTTCTTTTCTTTATATCTAGTGTTATGTAAAA
TPI1_repair oligo rv	14526	TTTTACATAACACTAGATATAAAGAAAAGAAGATAATATTTTTATATAATTATATTAATCTTTTAGTTTATGTAT GTGTTTTTTGTAGTTATAGATTTAAGCAAGAAAAGAATACAAACA
CDC19_repair oligo fw	14527	ATTATTCTCTCTTGTTTCTATTTACAAGACACCAATCAAAACAAATAAAACATCATCACAAAAAAGAATCATGAT TGAATGAAGATATTATTTTTTTGAATTATATTTTTTAAATTTTAT
CDC19_repair oligo rv	14528	ATAAAATTTAAAAAATATAATTCAAAAAAATAATATCTTCATTCAATCATGATTCTTTTTTGTGATGATGTTTTA TTTGTTTTGATTGGTGTCTTGTAAATAGAAACAAGAGAGAATAAT
Diagnostic PCR for selective editing		
TPI1_diagnostic_FW	3514	CTGACAGGTGGTTTGTTACG
TPI1_diagnostic_FW	6406	CATTCACGAAATTTAAGTGGCTCAG
PYK1_diagnostic_FW	4667	CCTTGAGGGAAGATTATCTTGCG
PYK1_diagnostic_FW	11915	GAGTGAGTGCTTTGTTCAATGG

Table S8 – RNA-Sequencing Depth

Strain	Replicate	Depth
Native	1	22089278
Native	2	22216419
Native	3	24257740
Watermarked	1	24070180
Watermarked	2	24780729
Watermarked	3	23334043
Double	1	23480289
Double	2	25448312
Double	3	23022264

Table S9 – Sampling OD for RNA seq

Optical density of IMX1770, IMX1771 and IMX2028 batch cultures analysed by RNA sequencing

Strain	Reactor	OD_{660}
IMX1770	27	7,8
IMX1770	28	9,95
IMX1770	29	8,25
IMX1771	27	7,75
IMX1771	28	8,05
IMX1771	29	7,95
IMX2028	27	4,5
IMX2028	29	4,51
IMX2028	30	5,79

Figure S1 - *In silico* comparison of two watermarking approaches for *ADH1*.

A

B

Alignment of the native and watermarked *ADH1* sequence using watermarking strategy 1 (two watermarking areas, A) or 2 (watermarks evenly spread over the coding region, B). Mismatches corresponding to the watermarks are highlighted in orange

Figure S2 – step-by-step watermarking

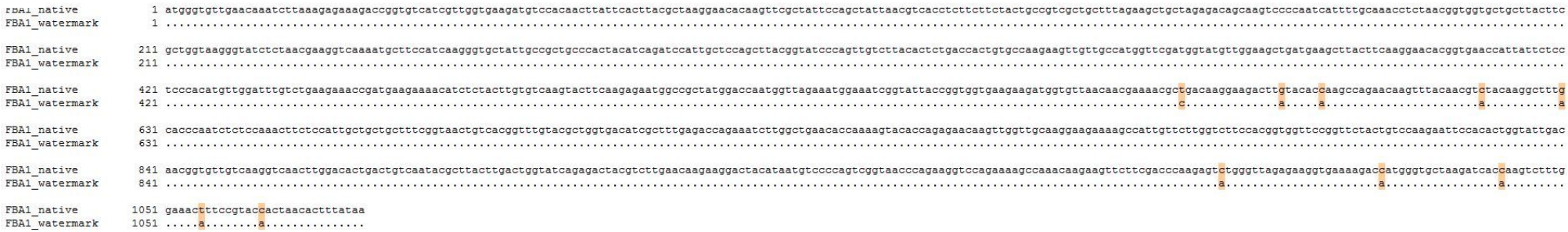
Introduction of watermark in CDS. Workflow

1. Open CDS se in Clone Manager.
2. Select first 101bp of sequence and annotate them as feature Region “buffer”. This fragment should be preserved to avoid alteration on the transcription level.
3. Go to the middle of CDS and proceed with the first watermark. Highlight central 100bp area and annotate it as a feature Region “watermark”.
4. Find 5-10 AA with 4 or 6 triplets (**AGPTVLRS**). Select those, so they are scattered along the whole 100bp and not grouped together (to make possible to detect watermark by program after).
5. Perform codon replacement on selected AA selecting an alternative codon (of the same AA) which has the closest percentage of similarity to the native variant in the sequence Codon usage table of *S. cerevisiae* attached). Annotate every change done as a Feature “Marker”. In the description mention how much % gain or loss did you introduce with this change.
   Note: for the following codons changes should be avoided and only done if less than 5 changes will be done otherwise:
   G (GGT) 61%
   R (AGA) 54%
   For triplet’s replacement when fraction % will drop ~20% try to find alternatives to fulfil 5 bp change.
6. Introduce 1 more watermark using the same algoritm at the end of the CDS (10 bp before the stop codon).
7. Add corresponding flanks to make the sequence compatible with the toolkit. Make sure there are no BsaI, BsmBI or NotI recognition sites in CDS (apart from the one present in the flank).

5’: GCATCGTCTCATCGGTCTCATATG
3’: ATCCTGAGACCTGAGACGGCAT

Figure S3 – Watermarking of *FBA1* and *ENO2*

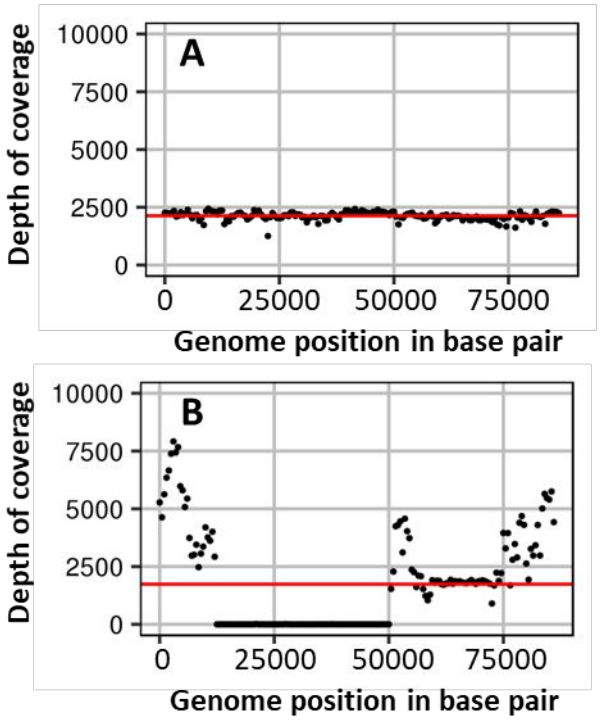
A



B

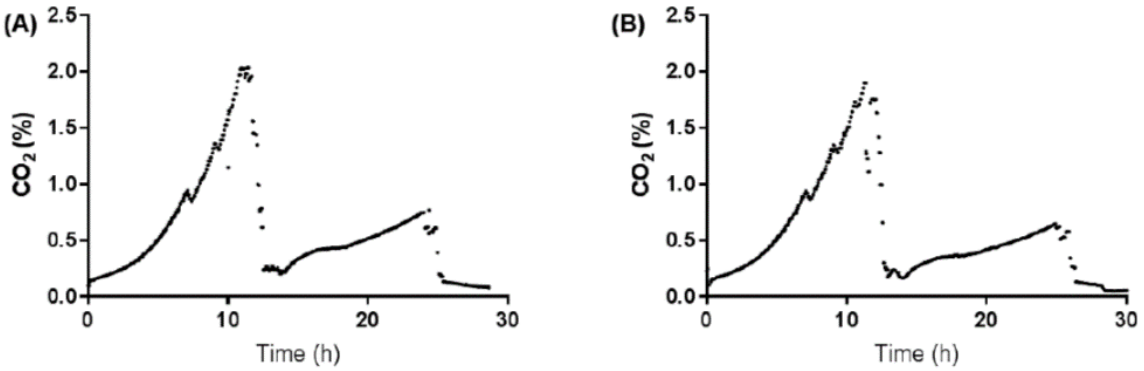
Alignment of the sequence of the coding region of native versus watermarked *FBA1* (A) and *ENO2* (B). Watermarks are highlighted in orange.

Figure S4 – Genetic characterization of IMX2028



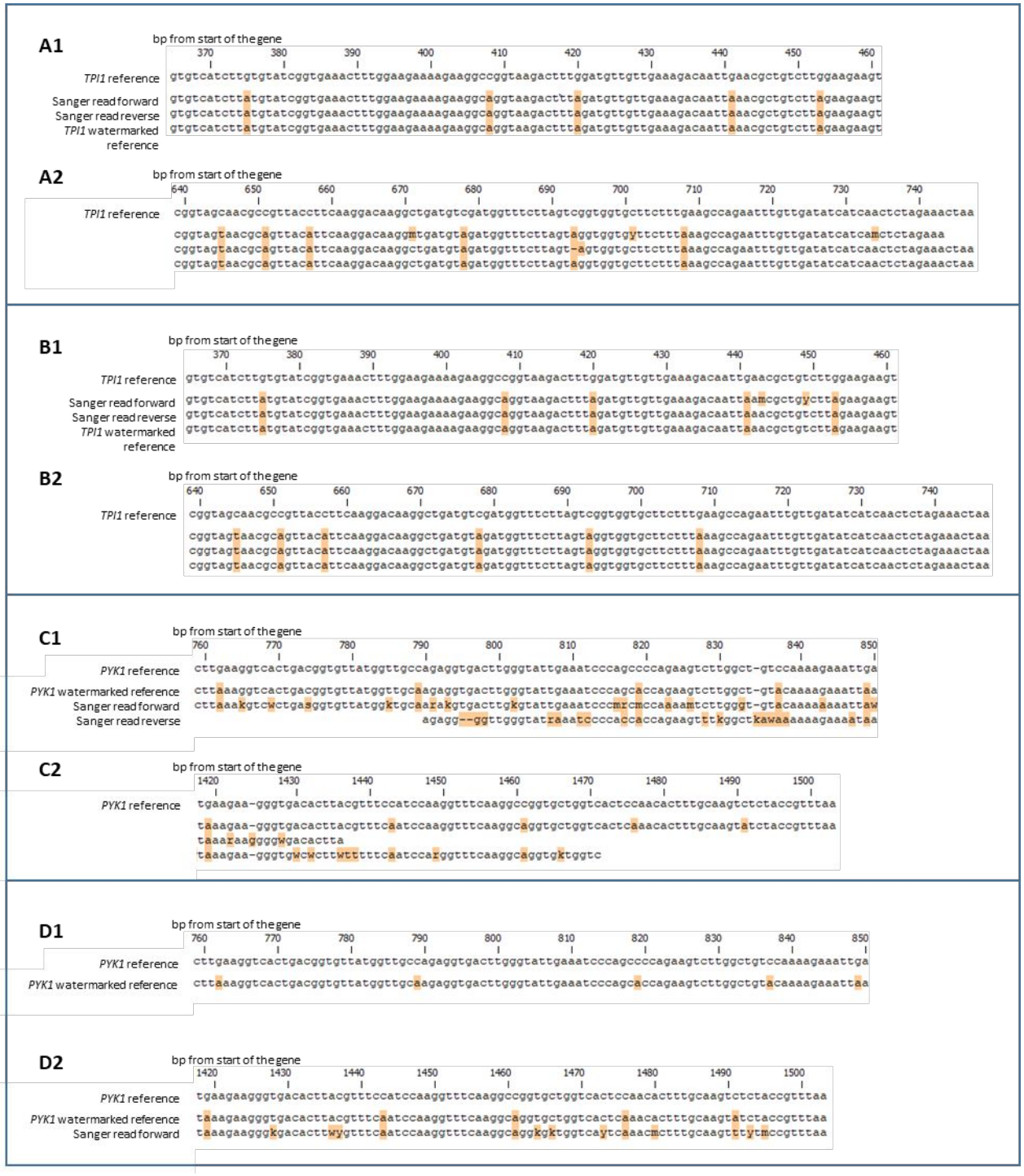
A. Coverage plot of mapped whole genome sequence reads of IMX1771 against the reference strain CEN.PK113-7D, only representing the mitochondrial genome. The uniform coverage around 2500 shows that IMX1771 mitochondrial genome is intact. **B.** Same plot as **A.** showing the coverage plot of the mitochondrial DNA of IMX2028 mapped against the reference strain CEN.PK113-7D. A large fraction of IMX2028 mitochondrial genome is missing. Black dots are averages over 500 bp non-overlapping windows and the red line depicts the median of the 500 bp non-overlapping windows.

Figure S5 – CO_2 profiles of batch cultures with IMX1770 and IMX1771



Growth profiles based on the CO_2 off-gas data of IMX1770 (**A**) and IMX1771 (**B**) when grown in aerobic batch cultures in bioreactors with glucose as sole carbon source. During the first 12-13 hours of culture, glucose was consumed and fermented into ethanol and other minor products. After glucose exhaustion (as seen by the fast drop in CO_2 in the off-gas) the strains used the fermentation products as carbon and energy sources to resume metabolic activity and growth.

Figure S6 – Confirmation by Sanger sequencing of selective DNA editing



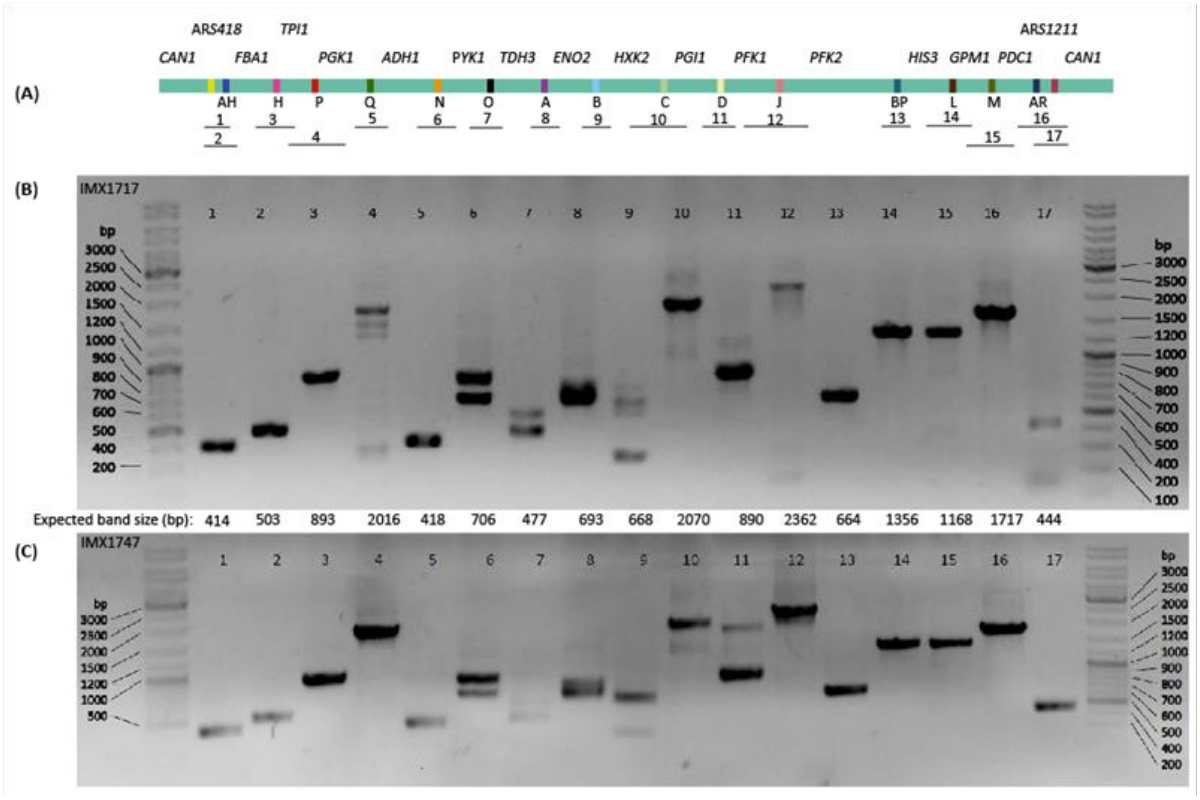
Sequencing of the non-edited *TPI1* and *PYK1* genes after selective CRISPR-Cas selective editing.

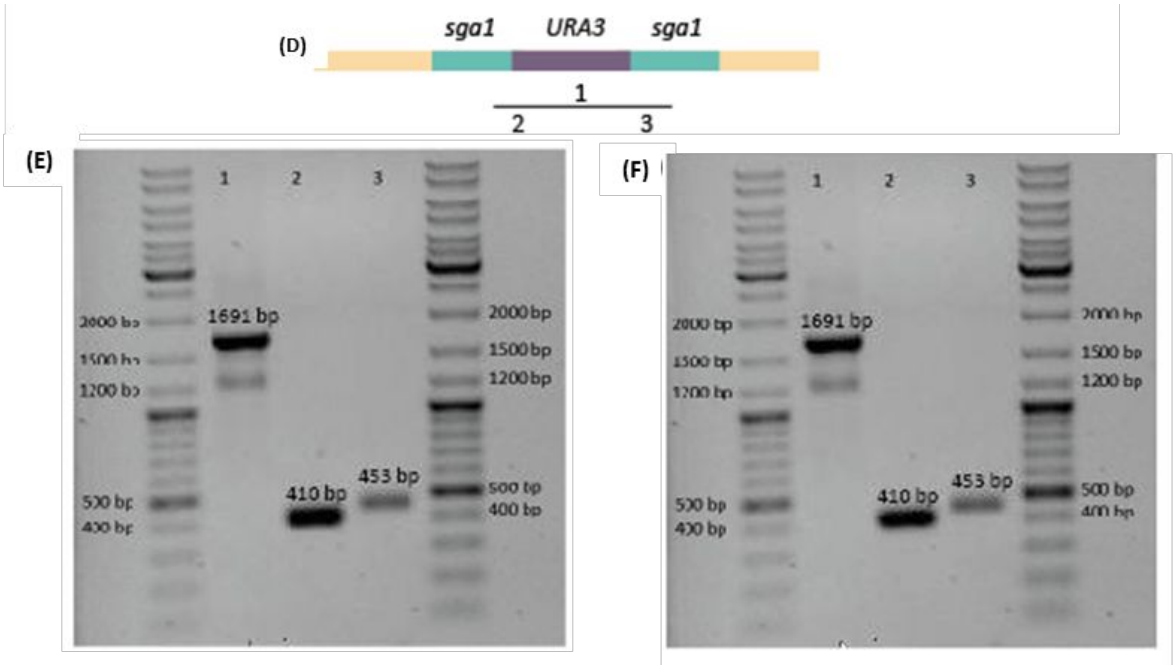
For *TPI1*, the band of 1125 bp for colonies number 9 and 10 showing two bands in the gel shown in Fig. 7A, was excised and purified using the Zymoclean Gel DNA Recovery kit. Sanger sequencing was performed according to BaseClear protocol with 15 ng/ 100bp and 25 pmol primer in a volume of 20 µl. Primers 3514 and 6406 were supplied for sequencing. **A1** and **B1,** sequence in the middle of *TPI1*

for colony 9 and 10, respectively. **A2** and **B2**, sequence at the end of *TPI1* until the stop codon for colony 9 and 10 respectively.

For *PYK1*, the band of 2177 bp for colonies number 1 and 5 showing two bands in the gel shown in Fig. 7A, was excised and purified using the Zymoclean Gel DNA Recovery kit. Sanger sequencing was performed according to BaseClear protocol with 15 ng/100bp and 25 pmol primer in a volume of 20 µl. Primers 745 and 4667 were supplied for sequencing. **C1** and **D1,** sequence in the middle of *PYK1* for colony 1 and 5, respectively. **C2** and **D2**, sequence at the end of *PYK1* until the stop codon for colony 1 and 5 respectively.

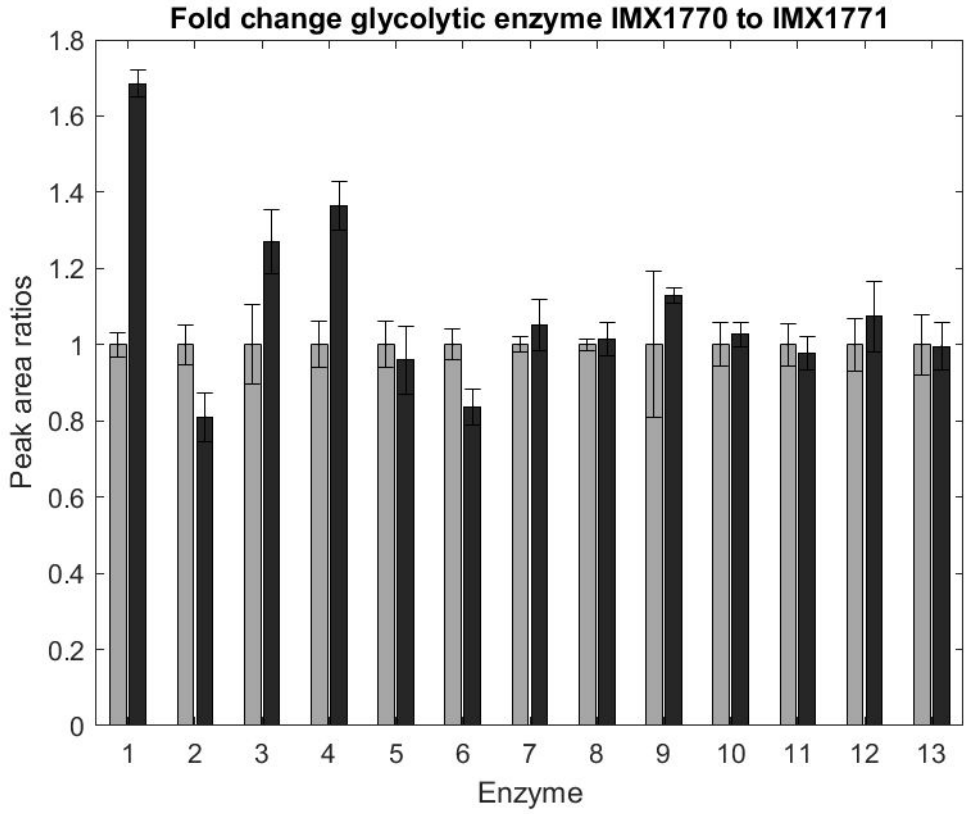
Figure S7 – PCR confirmation IMX1770 and IMX1771





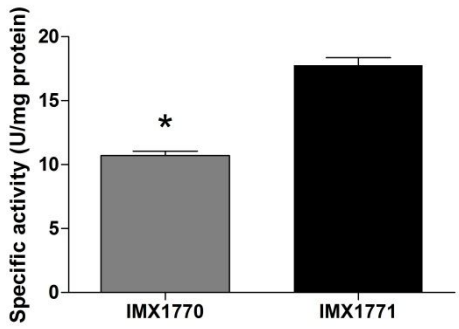
PCR confirmation of the SinLoG strains. (**A**) Schematic representation of the PCR setup to confirm integration of all fragments of the SinLoG. (**B**) Confirmation of integration of all SinLoG fragments in IMX1770. (**C**) Confirmation of integration of all SinLoG fragments in IMX1771. Confirmation of integration of the *KlURA3* repair fragment for the construction of IMX1770 and IMX1771. (**D**) Schematic representation of the PCR setup to confirm integration of the *KlURA3* gene in the *SGA1* locus. (**E**) Confirmation of KlURA3 integration in strain IMX1770. (**F**) Confirmation of *KlURA3* integration in strain IMX1771.

Figure S8 - Label free quantification (LFQ) of glycolytic protein abundance



Label free quantification (LFQ) of glycolytic protein abundance of watermarked (IMX1770, grey bars) and non-watermarked strain (IMX1771, black bars). Data represent the average and standard deviation for three biological replicates. 1=Gpm1, 2=Tdh3, 3=Eno2, 4=Pyk1, 5=Tpi1, 6=Pfk2, 7=Hxk2, 8=Pfk1, 9=Fba1, 10=Pdc1, 11=Pgk1, 12=Pgi1, 13=Adh1. The fold change for Gpm1 is significantly different between the strains (Student t-test, p-value threshold 0.05, two-tailed test, homoscedastic).

Figure S9 – Specific activity of Gpm1 (phosphoglucomutase) in shake flask cultures



Specific enzyme activity of the Gpm1 enzyme from IMX1770 (watermarked) and IMX1770 (native) strains. Samples were taken in mid-exponential phase. Bars represent the average and standard deviation of measurements from three independent shake flasks for each strain. Stars indicate enzyme activities that are significantly different between the two strains (Student t-test, p-value threshold 0.05, two-tailed test, homoscedastic).

Reference list

1. Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013,** *29* (1), 15-21.
2. Frazee, A. C.; Jaffe, A. E.; Langmead, B.; Leek, J. T., Polyester: simulating RNA-seq datasets with differential transcript expression. *Bioinformatics* **2015,** *31* (17), 2778-84.
3. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; Genome Project Data Processing, S., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009,** *25* (16), 2078-9.
4. Kuijpers, N. G.; Chroumpi, S.; Vos, T.; Solis-Escalante, D.; Bosman, L.; Pronk, J. T.; Daran, J. M.; Daran-Lapujade, P., One-step assembly and targeted integration of multigene constructs assisted by the I-SceI meganuclease in Saccharomyces cerevisiae. *FEMS Yeast Res* **2013,** *13* (8), 769-81.
5. Lee, M. E.; DeLoache, W. C.; Cervantes, B.; Dueber, J. E., A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth. Biol* **2015,** *4* (9), 975-986.