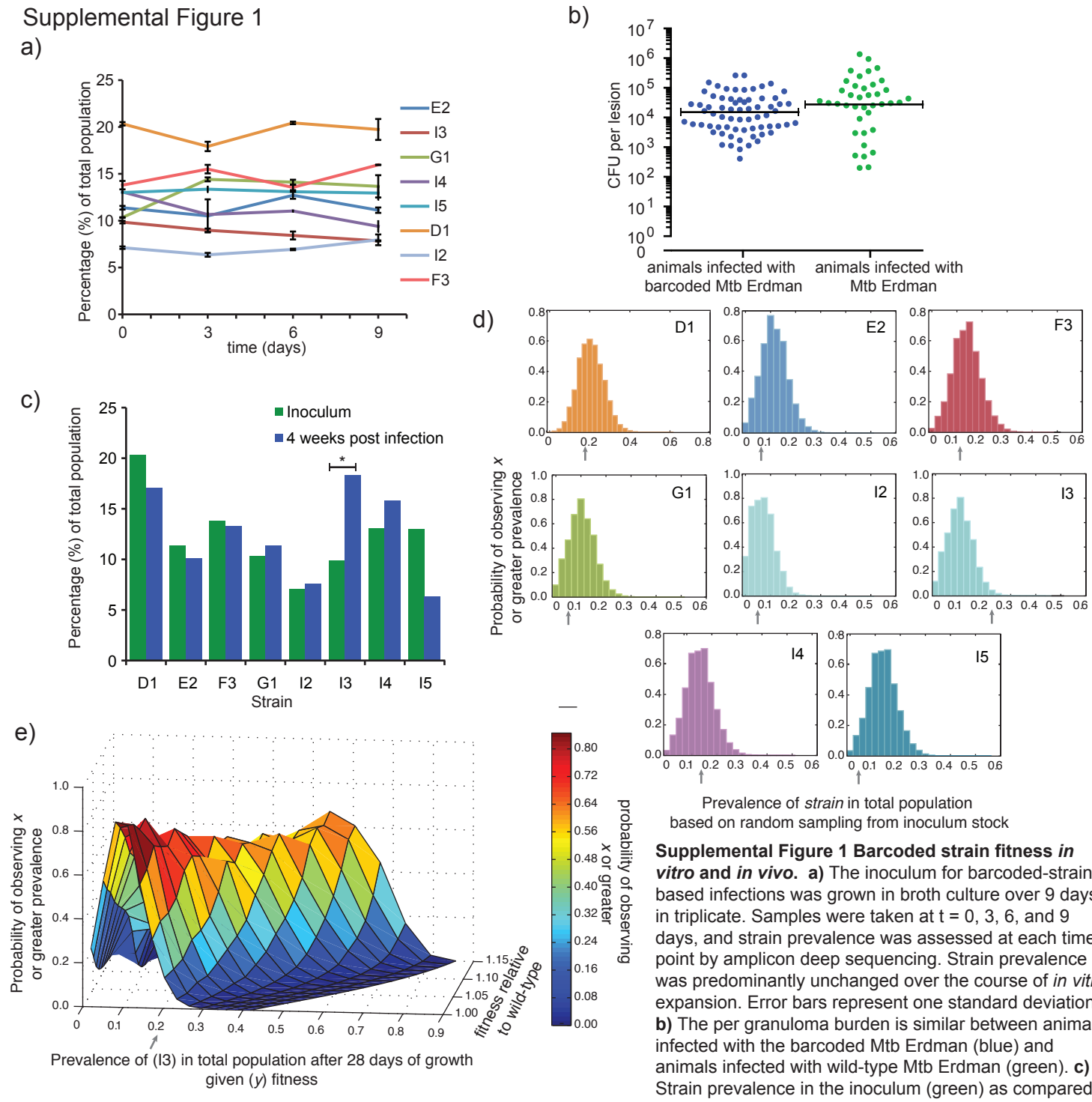
Sterilization of granulomas is common in both active and latent tuberculosis despite extensive within-host variability in bacterial killing

Philana Ling Lin*, Christopher B. Ford*, M. Teresa Coleman, Amy J. Myers, Richa Gawande, Thomas Ioerger, James Sacchettini, Sarah M. Fortune§, JoAnne L. Flynn§

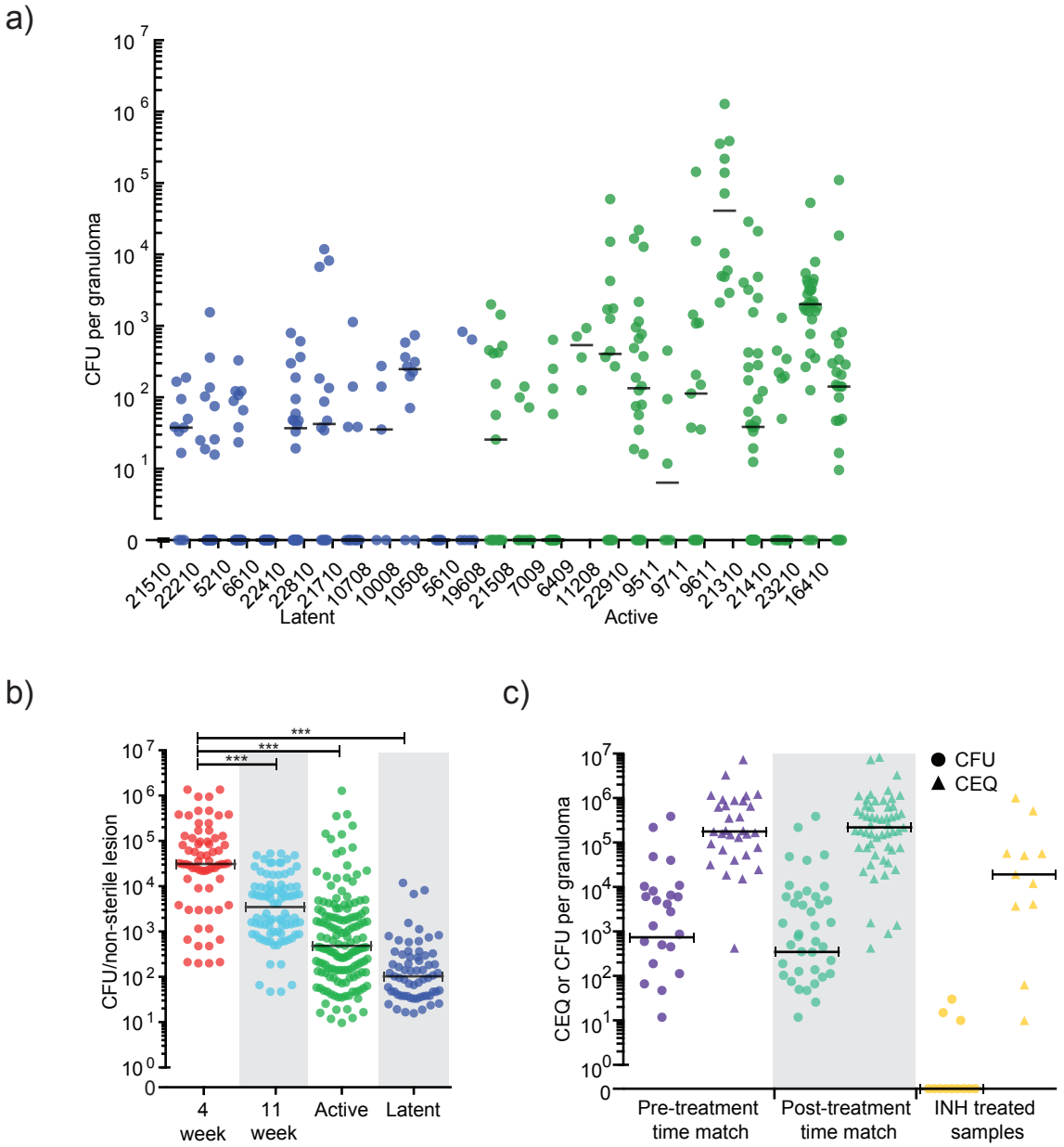
*Authors contributed equally; §Corresponding author

Supplemental Figure 1



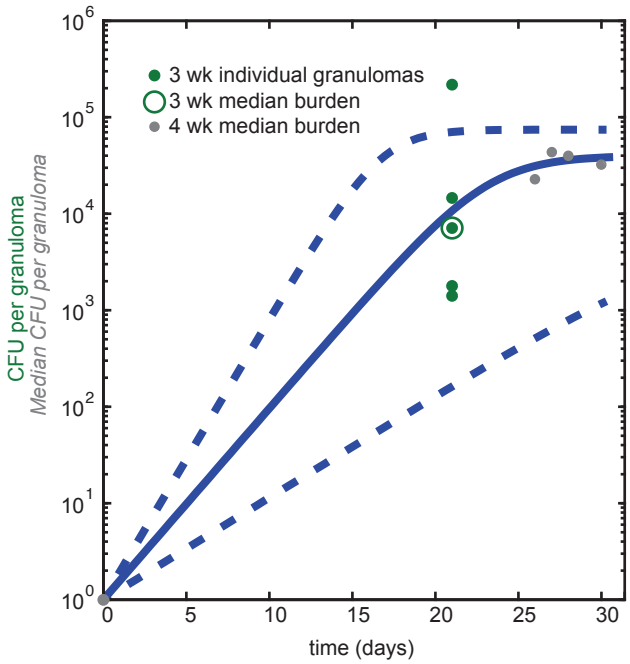
Supplemental Figure 1 Barcoded strain fitness *in vitro* and *in vivo*. **a)** The inoculum for barcoded-strain based infections was grown in broth culture over 9 days, in triplicate. Samples were taken at t = 0, 3, 6, and 9 days, and strain prevalence was assessed at each time point by amplicon deep sequencing. Strain prevalence was predominantly unchanged over the course of *in vitro* expansion. Error bars represent one standard deviation. **b)** The per granuloma burden is similar between animals infected with the barcoded Mtb Erdman (blue) and animals infected with wild-type Mtb Erdman (green). **c)** Strain prevalence in the inoculum (green) as compared to the average prevalence of each strain across all samples after 4 weeks of *in vivo* growth (blue). The majority of strains are not significantly changed from inoculum to 28 days post infection. Strain I3 saw a significant increase (~8%, p = 0.023). **d)** To determine if differences after 28 days of *in vivo* growth were due to stochastic sampling of the inoculum, sampling was simulated based on the prevalence of each strain in the inoculum stock (S_i). Using a binomial distribution with 34 trials, and the probability of success equal to S_i, 1e6 random numbers were generated for each strain, i, to build the expected sampling distribution. For strain I3, the probability of the observed prevalence after infection was less than 0.05, suggesting this prevalence may have been due to changes in strain fitness. Arrows represent observed prevalence. **e)** To estimate the potential change in fitness, bacterial growth *in vivo* was simulated for strain I3 according to equation (1), with fitness ranging from 1.00 (no change) to 1.15 (a 15% increase). With a 1% increase in fitness, the prevalence of change in I3 is no longer signficant, and with a 5% increase, the probability of the observed prevalence is maximized. Together, these data and simulations suggest that strain I3 may have a slight increase in fitness, though, as stated in text, this does not change our final conclusions. The grey arrow represents the observed prevalence.

Supplemental Figure 2



Supplemental Figure 2 CFU per granuloma in active and latent disease and CEQ controls. Wide and overlapping bacterial burden in individual lesions from monkeys with active TB and clinically latent infection. At necropsy, individual scan-identified granulomas are plated for bacterial burden. Shown are CFU/granuloma from 13 monkeys with active TB and 11 clinically latent monkeys. Bars represent medians, each circle is a single granuloma from that animal, animal numbers are on the X axis. **(b)** CFU from nonsterile lesions from monkeys at 4 weeks p.i. (n=4) is significantly higher than at 11 weeks (n=3), in active disease (n=13) and in clinically latent infection (n=11). All groups are significantly different from each other by pairwise comparison ($p<0.001$, Mann-Whitney). Circles represent individual lesions. **(c)** Lesions from animals treated with INH (treatment began on average 13 weeks post-infection, yellow) were analyzed for CFU and CEQ to determine if CEQ are stable in the absence of bacterial replication. While the median for CFU is 0, CEQ are significantly higher after two months of treatment, suggesting that CEQ persist in the absence of bacterial replication. Control groups were time matched such that the average time to necropsy was equal to either (1) the time elapsed until the first day of INH treatment (blue) or until the end of INH treatment (green). In both cases, CEQ was ~20 fold lower, suggesting a maximum rate of loss of 4% per day degradation of genomes in lesions treated by drug.

Supplemental Figure 3



Supplemental Figure 3 Phase I CFU model with 3 week CFU data. The ODE based model of bacterial growth between zero and 30 days fitted to median bacterial burden (CFU) per granuloma at each time point (grey symbols) is displayed in blue (dashed lines represent 95% confidence intervals). CFU per granuloma from an additional animal sacrificed at 3 weeks post infection is displayed. Each granuloma is represented by a single green symbol, the median value is circled.

Supplemental Table 1- Barcoded isolates

Strain[1]	Gene	Coordinate	WT	Mutant	Gene name or function	codon	mutation
E2	Rv0591	690767	G	C	*mce2C*	ttg	leu -> phe
I3	Rv0668	766229	G	A	*rpoC*	ggc	gly -> ala
G1*	Rv0876c	975906	T	C	unknown	ttt	*phe -> phe*
I4	Rv1644	1854208	G	T	*tsnR*	gct	ala->ser
I5	Rv1650	1857515	G	A	*pheT*	ggg	gly->arg
D1*	Rv2092c	2350697	G	T	*helY*	acg	*thr->thr*
I2	Rv2187	2448250	G	A	*fadD15*	cga	arg->glu
F3*	Rv3273	3655598	G	T	carbonic anhydrase	ctg	*leu->leu*

[1]Each strain was previously whole genome sequenced to identify polymorphisms relative to the parent strain, *M. tuberculosis* Erdman.

* Strains are marked by synonymous SNP.

Supplemental Table 2 – Percentage of lesions with *x* strains

Strain count (*x*)	Granulomas	Clustered granulomas	2 week granulomas	3 week granuloms	Lymph nodes
1	78.9	63.6	75.0	83.3	45.5
2	14.0	18.2	63.0	16.7	-
3	35.1	4.55	63.0	-	-
4	35.1	9.09	12.5	-	-
5	-	4.55	-	-	27.3
6	-	-	-	-	27.3

Supplemental Table 3 – Instantaneous and average doubling time in individual lesions

Time (days)	**1**	**10**	**20**	**30**	**40**	**80**[3]	**160 (and after)** [3]
Instanteous [1]	1.83 (2.18 – 1.59)	1.83 (2.18 – 1.59)	1.87 (2.20 – 1.76)	6.50 (2.82 –41.61)	483.57 (28.50 – 1.33e4)	*Inf* (*Inf – Inf*)	*Inf* (*Inf – Inf*)
Average[2]	1.83 (2.18 – 1.59)	1.83 (2.18 – 1.59)	1.83 (2.19 – 1. 60)	2.17 (2.24 – 4.11)	28.18 (3.99 –554.59)	*Inf* (*Inf – Inf*)	*Inf* (*Inf* – Inf)

[1] Instantaneous doubling times were estimated using equation (4) based on population size at time(t) (inferred from median CEQ per lesion) and represent the estimated doubling time at time (t). These values do not compensate for potential sources of loss of CEQ over time. In parentheses are the estimated doubling times for the 95% confidence intervals, as seen in figure 4a (lower and upper bound).
[2] Average doubling times based on median CEQ population size at time (t) were estimated using equation (4) and represent the estimated doubling time from time(0) to time(t). These values do not compensate for potential sources of loss of CEQ over time. In parentheses are the estimated doubling times for the 95% confidence intervals, as seen in figure 4a (lower and upper bound, respectively).
[3] *Inf* denotes values where the estimated rate of change in CEQ was effectively zero, giving an infinitely large doubling time.

Supplemental Table 4 – Animals used in this study

Experiment	Monkey ID	Mtb Strain	Days post infection
3 week	22110	Erdman	20
4 week	7110	Erdman	30
4 week	7709	Erdman	28
4 week	8109	Erdman	27
4 week	8709	Erdman	26
11 week	2412	Erdman	83
11 week	2512	Erdman	85
11 week	17111	Erdman	84
Latent	5210	Erdman	316
Latent	21510	Erdman	370
Latent	5610	Erdman	209
Latent	22210	Erdman	267
Latent	10508	Erdman	353
Latent	21710	Erdman	370
Latent	10508	Erdman	353
Latent	22810	Erdman	386
Latent	10708	Erdman	446
Latent	10008	Erdman	481
Latent	19808	Erdman	502
Latent	22410	Erdman	580
Active	9611	Erdman	93
Active	6409	Erdman	140
Active	9711	Erdman	175
Active	7009	Erdman	194
Active	9511	Erdman	198
Active	23210	Erdman	211
Active	21410	Erdman	267
Active	19608	Erdman	209
Active	21310	Erdman	230
Active	21410	Erdman	267
Active	21508	Erdman	271
Active	22910	Erdman	288
Active	11208	Erdman	355
Active	16410	Erdman	400
Active- INH treated	4008	Erdman	169
Active- INH treated	5808	Erdman	168
Active- INH treated	8309	Erdman	110
Active- INH treated	9009	Erdman	208
Active- INH treated	19308	Erdman	129
4 week Barcode	7010	Erdman barcoded	27
4 week Barcode	7210	Erdman barcoded	29

Supplemental Table 5 – Primers used in this study

Name	Target	Project	Sequence (5' – 3')
CBF216	sigF	CEQ	GCG GGT CGG GCT GGT CAA C
CBF217	sigF	CEQ	CCT CGC CCA TGA TGG TAG GAA C
CBF218[1]	sigF	CEQ	FAM - TCG GAC TTC GTC TCC TTC – IB
CBF175[2]	Rv0591	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CACGGCCCGACACCCCACTAC
CBF176[2]	Rv3273	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GCAGGGCAGCGCGAACGTG
CBF177[2]	Rv0876c	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GTCGTCGGGAGGGTAGTTGGC
CBF178[2]	Rv0668	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GTCGAGCGTGGTCAAGACCTGG
CBF179[2]	Rv1644	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CAGTGGACGGCGAGATGGCT
CBF180[2]	Rv1650	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GACCTGATCTGCGGCAACCC
CBF181[2]	Rv2187	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GTCGCGGCCATGGTGTTCGA
CBF182[2]	Rv2092c	Barcode	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CGCCTGTTGCGGACCCATC
CBF183[2]	Rv0591	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**CTCAAGCGGTTGGGTGCGTG
CBF184[2]	Rv3273	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**CCAGCTTGACCAGCTGGGCA
CBF185[2]	Rv0876c	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**TCGCCACGGCAGACTCCTC
CBF186[2]	Rv0668	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**GGTGGAAGGTGCGCATGGTCAG
CBF187[2]	Rv1644	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**TGCGTAGGGCCGGACAATCA
CBF188[2]	Rv1650	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**GAGCACCCGCGTGGTCATG
CBF189[2]	Rv2187	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**CCGCACCCACAGCCAGAATC
CBF190[2]	Rv2092c	Barcode	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**GCTGGAGCGGCTGGTGAAGTT
CBF191[3]	Read 2, Index 1	Barcode	**CAAGCAGAAGACGGCATACGAGAT***ATCACGGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF192[3]	Read 2, Index 2	Barcode	**CAAGCAGAAGACGGCATACGAGAT***CGATGTGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF193[3]	Read 2, Index 3	Barcode	**CAAGCAGAAGACGGCATACGAGAT***TTAGGCGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF194[3]	Read 2, Index 4	Barcode	**CAAGCAGAAGACGGCATACGAGAT***TGACCAGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF195[3]	Read 2, Index 5	Barcode	**CAAGCAGAAGACGGCATACGAGAT***ACAGTGGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF196[3]	Read 2, Index 6	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GCCAATGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF197[3]	Read 2, Index 7	Barcode	**CAAGCAGAAGACGGCATACGAGAT***CAGATCGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF198[3]	Read 2, Index 8	Barcode	**CAAGCAGAAGACGGCATACGAGAT***ACTTGAGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF199[3]	Read 2, Index 9	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GATCAGGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF200[3]	Read 2, Index 10	Barcode	**CAAGCAGAAGACGGCATACGAGAT***TAGCTTGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF201[3]	Read 2, Index 11	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GGCTACGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF202[3]	Read 2, Index 12	Barcode	**CAAGCAGAAGACGGCATACGAGAT***CTTGTAGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF203[3]	Read 2, Index 13	Barcode	**CAAGCAGAAGACGGCATACGAGAT***AGTCAAGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF204[3]	Read 2, Index 14	Barcode	**CAAGCAGAAGACGGCATACGAGAT***AGTTCCGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF205[3]	Read 2, Index 15	Barcode	**CAAGCAGAAGACGGCATACGAGAT***ATGTCAGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
CBF206[3]	Read 2, Index 16	Barcode	**CAAGCAGAAGACGGCATACGAGAT***CCGTCCGT*GACTGGAGTTCAGACGTGTGCTCTTCCGATCT
(cont.)	**(cont.)**	**(cont.)**	**(cont.)**

Name	Target	Project	Sequence (5' – 3')
CBF207[3]	Read 2, Index 17	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GTCCGCGT*GACTGGAGTTCAGACGTGTGCT CTTCCGATCT
CBF208[3]	Read 2, Index 18	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GTGAAAGT*GACTGGAGTTCAGACGTGTGCT CTTCCGATCT
CBF209[3]	Read 2, Index 19	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GTGGCCGT*GACTGGAGTTCAGACGTGTGCT CTTCCGATCT
CBF210[3]	Read 2, Index 20	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GTTTCGGT*GACTGGAGTTCAGACGTGTGCT CTTCCGATCT
CBF211[3]	Read 2, Index 21	Barcode	**CAAGCAGAAGACGGCATACGAGAT***CGTACGGT*GACTGGAGTTCAGACGTGTGCT CTTCCGATCT
CBF212[3]	Read 2, Index 22	Barcode	**CAAGCAGAAGACGGCATACGAGAT***GAGTGGGT*GACTGGAGTTCAGACGTGTGCT CTTCCGATCT
CBF213[3]	Read 2, Index 23	Barcode	**CAAGCAGAAGACGGCATACGAGAT***ACTGATGT*GACTGGAGTTCAGACGTGTGCT CTTCCGATCT
CBF214[3]	Read 2, Index 24	Barcode	**CAAGCAGAAGACGGCATACGAGAT***ATTCCTGT*GACTGGAGTTCAGACGTGTGCTC TTCCGATCT
CBF215[4]	Read 1	Barcode	**AATGATACGGCGACCACCGAGATCT**ACACTCTTTCCCTACACGACGCTCTTCCG ATCT

[1]Probe structure included a 5' FAM attachment, a 3'Iowa Black Quencher, and an internal Zen Quencher, not noted in the above sequence.

[2] Primer structure included gene sequence homology and Illumina Read 1 or Read 2 primer homology (bold)

[3] Primer structure included Illumina attachment homology (bold), Illumina TruSeq index (italics), and Illumina Read 2 primer homology designed to amplify the primer homology included in primers CBF175-190

[4] Primer structure included Illumina attachment homology (bold) and Illumina Read 1 primer homology designed to amplify the primer homology included in primers CBF175-190