Local genetic context shapes the function of a gene regulatory network

Anna Nagy-Staroń[1], Kathrin Tomasek[1], Caroline Caruso Carter[1], Elisabeth Sonnleitner[2], Bor Kavčič[1], Tiago Paixão[1, 3], Călin C. Guet[1]

[1] Institute of Science and Technology Austria, Klosterneuburg, Austria

[2] Department of Microbiology, Immunobiology and Genetics, Max F. Perutz Laboratories, Center Of Molecular Biology, University of Vienna, Vienna, Austria

[3] Present address: Instituto Gulbenkian de Ciência, Oeiras, Portugal

Corresponding author: Călin C. Guet email: calin@ist.ac.at

Gene expression levels are influenced by multiple coexisting molecular mechanisms. Some of these interactions, such as those of transcription factors and promoters have been studied extensively. However, predicting phenotypes of gene regulatory networks remains a major challenge. Here, we use a well-defined synthetic gene regulatory network to study how network phenotypes depend on local genetic context, i.e. the genetic neighborhood of a transcription factor and its relative position. We show that one gene regulatory network with fixed topology can display not only quantitatively but also qualitatively different phenotypes, depending solely on the local genetic context of its components. Our results demonstrate that changes in local genetic context can place a single transcriptional unit within two separate regulons without the need for complex regulatory sequences. We propose that relative order of individual transcriptional units, with its potential for combinatorial complexity, plays an important role in shaping phenotypes of gene regulatory networks.

Introduction

Changes in regulatory connections between individual transcriptional units (TUs) or, in other words, the rewiring of gene regulatory networks (GRNs), is a major genetic mechanism underlying phenotypic diversity (Shubin et al., 2009; Wagner and Lynch, 2010; Wray, 2007). A lot of effort has been put into understanding how mutations in transcription factors and their DNA binding sites within promoter regions influence GRN behavior, plasticity and evolution (Babu et al., 2004; Balaji et al., 2007; Chen et al., 2012; Igler et al., 2018; Isalan et al., 2008; Nocedal et al., 2017). However, we are still unable to predict GRN phenotypes from first principles (Browning and Busby, 2016).

Genes and the genetic elements that regulate them, promoters, are arranged in a linear manner on chromosomes. Thus they are embedded into a larger genetic context, represented by the changing genetic background of the rest of the genomic sequence or by their specific physical location on the chromosome. The genetic context in which GRNs find themselves in, influences and modulates the way these respond to the environment and interact with other GRNs and it also shapes the interactions within the GRN itself (Cardinale and Arkin, 2012; Chan et al., 2005; Steinrueck and Guet, 2017; Wu and Rao, 2010). In bacteria, gene expression levels are thought to be determined by RNA polymerase recognizing promoter sequences and subsequently initiating transcription, which is the key step at which a large part of transcriptional regulation takes place (Browning and Busby, 2004). However, context effects resulting from occupying a particular location within the genome can significantly alter expression levels (Junier, 2014; Lagomarsino et al., 2015; Scholz et al., 2019). Distance to the

46 origin of replication influences transcription levels due to gene dosage effects, the presence of
47 transcriptionally active and silent regions, as well as spatial and temporal variation in DNA
48 superhelicity; while collisions between DNA replication and transcription influence gene
49 expression levels differently on leading and lagging strands (Beckwith et al., 1966; Bryant et al.,
50 2014; Mirkin et al., 2006; Sobetzko et al., 2012; Vora et al., 2009). At a local scale,
51 transcriptional interference, transcription-coupled DNA supercoiling, presence of *cis*-antisense
52 RNA, as well as transcriptional read-through, all link together the expression of neighboring TUs
53 (Cambray et al., 2013; Georg and Hess, 2011; Liu and Wang, 1987; Reynolds et al., 1992;
54 Shearwin et al., 2005; Wu and Fang, 2003). Within operons, number, length and order of genes
55 can all affect gene expression (Jacob and Monod, 1961; Lim et al., 2011; Zipser, 1969). All of
56 these factors, that can individually modulate gene expression, vary simultaneously across the
57 genome, with potential for significant combinatorial effects (Meyer et al., 2018; Scholz et al.,
58 2019).
59
60 While these multiple local context-dependent mechanisms are known to modulate gene
61 expression, the qualitative phenotype of a GRN is usually thought to be defined solely by the
62 network topology, with the connections within GRN considered independent of the physical
63 location of the network components (Babu et al., 2004; Mangan et al., 2003), and determined
64 simply by the promoter sequences. One of the reasons for this assumption is the belief that *cis*-
65 regulatory changes are less pleiotropic than changes to the protein itself (Prud'homme et al.,
66 2006), although some have questioned this (Stern and Orgogozo, 2008). However, other non-
67 coding genetic factors such as transcriptional read-through or supercoiling have the potential to

change gene expression with the same pleiotropic freedom as *cis*-regulatory changes. Here, we ask how the immediate local genetic context outside of individual TUs of a GRN can alter both, the qualitative and quantitative phenotype of a network, and how many phenotypes are accessible for this particular GRN, while the network topology *per se* remains unchanged. In order to keep the number of genetic interactions to a minimum, we chose to study a synthetic GRN. This tractable system allows for a simplified description of more complex naturally occurring GRNs, where a large number of inherently complex interactions make such a question very difficult to answer experimentally (Mukherji and van Oudenaarden, 2009; Wolf and Arkin, 2003). Here, we shuffle individual TUs (understood here as the unit formed by: the mRNA coding sequence, the promoter driving its expression and the transcriptional terminator marking the end of the transcribed sequence) of a GRN. In doing so, we alter solely the *local genetic context,* while keeping the actual interactions (topology) within the GRN unchanged and thus the number of interactions to a tractable minimum. We then define the *phenotype of the GRN* as the levels of gene expression measured across four different environments, defined by the presence or absence of two different chemical inducers that alter the binding state of two different well characterized transcription factors. Qualitative phenotypes are here based on a set of binary output values for each input state, therefore defining different logical operators (e.g. NOR, ON, OFF) (for details see *Threshold for assigning a phenotype to individual GRNs*). Quantitative phenotypes are defined as a set of four expression values varying continuously within one particular behavior. In this way, we systematically explore the space of possible phenotypes of the GRN and thereby we can disentangle the effects of local genetic context from multiple other factors that can affect gene expression levels.

Results

Our GRN (Fig. 1A) is composed of the genes coding for three of the best characterized repressors: LacI, TetR and lambda CI (abbreviations used throughout the text: L, T and C, respectively), and the promoters they control, P_{lac}, P_{tet} and P_R. The three repressor genes are transcriptionally interconnected into a GRN, with LacI repressing both *tetR* and its own expression, and TetR repressing expression of *cI*. The controlled promoters are synthetic variants of the P_L promoter of phage lambda with two *tet* or *lac* operator sites located in the direct vicinity of the -35 and -10 promoter elements (Lutz and Bujard, 1997). The binding state of TetR and LacI changes in the presence of inducers: anhydrotetracycline (aTc) and isopropyl β-D-thiogalactopyranoside (IPTG), respectively. A *yfp* gene, expressed from a CI-controlled P_R promoter, serves as output. The *yfp* gene is located separately from the rest of the genetic circuit at a transcriptionally insulated locus on the chromosome (*attB* site of phage P21). In our synthetic system, each individual promoter and transcription factor gene it controls are separated from the neighboring one by a strong transcriptional terminator T1 of the *rrnB* gene (Orosz et al., 1991), forming an individual TU (Fig. 1B). We chose T1 as one of the strongest transcriptional terminators in *E. coli* to transcriptionally insulate individual TUs from one another (Cambray et al., 2013).

Phenotype of GRN depends on local genetic context despite identical topology

We asked whether the local genetic context can influence the phenotype of this GRN. First, we developed a simple mathematical model of the mechanistic basis of gene expression for this specific network topology (Fig. S1). This model tracks the concentrations of all three repressors

as their respective promoter activities are influenced by their known specific network interactions. The predicted phenotype of our GRN will depend on the presence of aTc, but not on the presence of IPTG (Fig. 1A). We wanted to test if this phenotype is independent of the relative TU order and orientation, and we aimed to build plasmids with all possible 48 relative TU order permutations with fixed positions (Riordan, J., 2003) such that: (i) every TU can occupy any of the three positions, (ii) every TU is present only once, (iii) both forward and reverse orientations are possible, and (iv) network topology stays the same (Fig. 1B). To facilitate comparisons among all GRNs we used a threshold on the expression of the YFP output for assigning a binary output value to each environment and so defined a phenotype the GRN can achieve (for thresholds used to assign a particular phenotype to individual networks, see *Material and Methods*).

The phenotype of strains carrying the resulting 37 plasmids (multiple attempts to clone eleven of the TU order permutations failed, see *Supplemental material*) varied widely both quantitatively and qualitatively (Fig. 1CD, S2 and S3). More than half (20) of the tested GRN permutations showed a phenotype which was qualitatively different than what was predicted *ab initio*. We also observed multiple quantitative differences in expression levels within one class of logical phenotypes (e.g. permutations CLT, CTL_r, T_rCL and LT_rC_r, Fig. 1C).

GRN phenotype is influenced by local genetic context independently of the replicon

We then asked how and why the changes in relative order of individual TUs affect the phenotype of the GRN, although network topology and individual genetic components of the

GRN (i.e. individual TUs) remain unchanged. In order to disentangle the specific interactions between network components we focused on six GRNs which differ in relative TU order but not in gene orientation. These six GRNs show two qualitatively different phenotypes (NOR and NOT (aTc), Fig. 2A, upper panel). In four out of six strains (LCT, LTC, TCL and TLC), induction with IPTG shifted *yfp* expression levels to the OFF state. These population-level findings are also observed at the single cell level (Fig. 2A, lower panel). We tested to what extent our observation from plasmid-based TUs apply to chromosomally located GRNs by integrating three networks with varying TU order (CTL, LCT and TLC) at a transcriptionally insulated chromosomal locus (*attB* site of phage HK022). In line with the plasmid-based GRNs, these strains also showed a dependency of phenotype on relative TU order, demonstrating that this is not an effect related to plasmid localization of our GRN (Fig. 2B). When *lacI*, *tetR*, and *cI* are integrated at separate, transcriptionally insulated loci on the bacterial chromosome, the network phenotype is identical with the one predicted *ab initio* from its topology, confirming that it is the transcription of neighboring genes that changes the network's phenotype (Fig. 2C).

Differences in cI expression lead to phenotypes that depend on relative TU order

To elucidate the molecular basis of the observed phenotypic variability, we first asked whether the relative TU order-dependent differences in phenotypes can be traced back to changes in levels of *cI* gene expression. We isolated total RNA from strains differing in relative position of the *cI* gene (CLT and TLC) grown in the absence or presence of IPTG, and quantified *cI* transcript levels using RT-qPCR. *cI* expression after IPTG induction in strain TLC was over 10-fold higher than in strain CLT (Fig. 3A), suggesting that the differences in *yfp* fluorescence were indeed due

to differences in *cI* expression. In order to corroborate our findings at the mRNA level with protein expression levels, we replaced the *cI* gene in strains CLT and TLC with *yfp,* and confirmed relative gene order effects on *yfp* expression directly (Fig. 3B).

Expression of the *cI* gene in our GRN is controlled by TetR (Fig. 1A). Thus differences in *cI* expression levels between the different relative TU order variants can be due to (i) global changes in gene expression of P_{tet} controlled genes, which propagate to changes in expression of *cI*; or (ii) local effects, such as transcriptional read-through or changes in supercoiling levels. To distinguish between these two possibilities, we measured the activity of the promoter driving *cI* expression by supplying P_{tet}-*cfp* in *trans* on a second plasmid. In all six strains with different relative TU order, P_{tet} activity was strongly induced with aTc, while no P_{tet} activation was observed after IPTG induction (Fig. 3C). This indicates that *cI* expression observed after IPTG induction is not due to global removal of TetR repression, but rather due to a local effect on *cI* gene expression. Furthermore, levels of TetR-dependent repression do not depend on relative TU order. This local effect depends on gene expression from the *lac* promoter in an IPTG-dependent manner (Fig. 3D).

Transcriptional read-through is the molecular mechanism underlying context-dependent GRN phenotype

We hypothesized that transcriptional read-through would be consistent with the context-dependent effects we measured. We observed that in the cases when the *cI* gene is at the second or third position of the GRN (Fig. 1B), transcriptional read-through from upstream TUs (*tetR* and/or *lacI*) may be enough to transcribe *cI* despite TetR-dependent repression, and in

178 turn shut down P_R activity in response to IPTG (see also *Supplemental material* for effects of
179 transcriptional read-through into *tetR* and *lacI*). We asked whether the potential for
180 transcriptional read-through, together with the knowledge about the individual genetic
181 components of this network, is enough to unambiguously predict the phenotype of the GRN
182 permutation variants we built. For this purpose, we added the effects of transcriptional read-
183 through to our mathematical model such that promoter activities were now influenced not only
184 by network interactions but also by the activity of neighboring genes (Fig. S1).
185 To facilitate the analysis of numerous GRNs we divided the 48 possible TU permutations into 24
186 pairs that differ only in orientation with respect to the plasmid backbone (Fig. S2 and S3). 26
187 networks/13 pairs showed the same phenotype in both orientations (Fig. S2), 8 networks/4
188 pairs showed different phenotype in each orientation (Fig. S3A), for 3 networks no
189 corresponding pair was cloned (Fig. S3B) and 8 networks/4 pairs were not cloned.
190 For networks showing the same phenotype in both orientations we assumed there is no
191 significant influence of plasmid backbone elements (Fig. S2). Here, the model including
192 transcriptional read-through agreed for 20 networks/10 pairs (Fig. S2A), and did not agree with
193 6 networks/3 pairs (Fig. S2B). A null-model that did not account for transcriptional read-through
194 failed to predict the observed differences in phenotypes.
195 Our experimental approach to test the transcriptional read-through hypothesis is based on the
196 premise that transcriptional read-through does not depend on a functional promoter of a
197 downstream gene, in contrast to supercoiling- or RNAP concentration-dependent effects.
198 Therefore, if GRN behavior is due to transcriptional read-through, mutating the P_{tet} promoter
199 should not affect the responsiveness to IPTG. We tested this prediction by introducing two

point mutations into the -10 element of P_{tet} in a number of different strains (Fig. 4A and Fig. S4). These two point mutations render P_{tet} inactive and therefore prevent transcription from this promoter (Fig. S5). The phenotype in the absence of any inducers and with only IPTG was identical to the phenotype in the original strains, thus confirming that *cI* expression is initiated at an upstream promoter (Fig. 4B and Fig. S4).

If the mechanism behind different phenotypes is transcriptional read-through, change of the terminator strength should lead to a change in phenotype. We have chosen network TLC in which our model predicted that a change of terminator strength will lead to alter the phenotype. In this GRN we exchanged the T1 terminator preceding *cI* to either a stronger double T1T2 terminator of the *rrnB* locus, or to the weaker T*crp* and T*tonB* terminators (Cambray et al., 2013). Change from T1 to T1T2 changes the phenotype of the network to the one predicted by the model and observed when single transcription factors are incorporated in separate loci on the chromosome (Fig. 5A). Change from T1 to either T*crp* or T*tonB* leads to a completely OFF phenotype, as expected when transcriptional read-through through weaker terminators leads to expression of *cI* in all four conditions.

Read-through transcripts, i.e. transcripts of more than one gene, can be detected by Northern blotting. We isolated total RNA from strains with T1, T*crp* and T*tonB* grown in the absence or presence of IPTG, and visualized transcripts on Northern blot (Fig. S6). We expected to detect *lacI* transcripts in all strains, and in case of transcriptional read-through, longer transcripts encompassing both *lacI* and *cI*. No read-through transcript starting in *tetR* was expected, since the T1 terminator separating *tetR* from *lacI* carries a RNaseE recognition site in its stem (Apirion and Miczak, 1993; Szeberényi et al., 1984), which makes it impossible to distinguish between

222 read-through transcripts and transcripts originating from individual promoters. The same is true
223 for the strain harboring T1 terminators only, in which we expected to see only single gene
224 transcripts.
225 We detected read-through transcripts encompassing both *lacI* and *cI* in strains carrying T*crp*
226 and T*tonB*, thus directly demonstrating transcriptional read-through in these two strains. We
227 also detected fragments encompassing only the *cI* gene. This may suggest RNA processing at a
228 cryptic RNase site, or additional effects, such as dislodgement of repressor by RNA polymerase
229 passing through the terminator (Palmer et al., 2011). To rule out the emergence of unpredicted
230 promoters, we fused the junctions between *lacI* and *cI*, encompassing T1, T*crp* and T*tonB*
231 terminators in front of YFP, but did not detect any significant increase in fluorescence (Fig. S7).
232 In order to check whether a combination of two terminators would act additively on stopping
233 transcriptional read-through, we inserted the weaker *crp* terminator in front of T1 terminator in
234 a different GRN, LCT, resulting in network $L^{crpT1}CT$. This network showed an intermediate
235 phenotype, consistent with predicted decrease in *cI* expression (Fig. 5B).
236 Taken together, these results strongly support that transcriptional read-through is the
237 molecular mechanism underlying the relative TU dependent phenotypes we observe, and thus
238 the different logic phenotypes our GRN can achieve.
239
240 *Interplay of several molecular mechanisms shapes GRN phenotype*
241 Eight networks (four pairs) showed different phenotypes in each orientation and in all cases
242 one phenotype from the pair was supported by our model while the other was not. This
243 suggests that there was a significant influence of the plasmid backbone elements. Just like the

244 individual TUs constituting our GRN, TUs located on the plasmid backbone (namely the
245 kanamycin resistance gene *kanR* and *repA* in the plasmid origin of replication) also have the
246 potential to influence expression of neighboring TUs, and hence the GRN phenotype. To rule
247 out transcriptional read-through from the *repA* gene located at the plasmid origin of
248 replication, we cloned a promoterless *yfp* gene downstream of *repA*, however we did not
249 detect any significant change in fluorescence (Fig. S8A). We noticed that there is a slight change
250 in repression of P_{tet} depending on the relative orientation to the plasmid backbone (Fig. S8B).
251 Since supercoiling can influence gene expression, especially of plasmid-located genes, we also
252 expect supercoiling–mediated effects to modulate expression and thus influence phenotypes of
253 our GRN (Sobetzko, 2016; Yeung et al., 2017).
254
255 *Local genetic context modulates regulation of lac promoter*
256 Our simplified synthetic system allowed us to observe the effects of local genetic context
257 created in a systematic manner by arranging transcription factor genes next to one another on
258 a plasmid (Fig. 2A) or the chromosome (Fig. 2B). We also decided to test our findings in a native
259 regulatory network of *E. coli* composed of one transcription factor – LacI, and the promoter it
260 controls, P_{lac}, and asked whether native local genetic context has the potential to modulate its
261 phenotype. Using a Δ*lacI* Δ*lacZYA* genetic background we inserted the *lac* promoter driving *yfp*
262 expression in a transcriptionally insulated locus (phage λ attachment site, *attB*) and the *lac*
263 repressor, *lacI*, into one of three loci on the right replichore (Fig. 6A). Subsequently, we
264 measured gene expression from the *lac* promoter, P_{lac}, in these three strains by monitoring *yfp*
265 fluorescence levels after treatment with increasing IPTG concentrations. The shape of the

266 induction curve was considered to be the network's phenotype. It should be stressed that the
267 *lacI* gene (under control of its native promoter) was inserted only into non-coding chromosomal
268 regions, shortly after an endogenous terminator. The loci for insertion were chosen such that
269 genes downstream of the terminator were non-essential and in the same orientation as the
270 genes upstream. Chosen terminators were of different strengths: strong, middle, and weak (D.
271 Toledo Aparicio, M. Lagator and A. Nagy-Staron, personal communication, September 2019)
272 and were located in close vicinity (39' - 43' on MG1655 chromosome) to avoid gene dosage and
273 transcription factor – promoter distance effects (Block et al., 2012; Kuhlman and Cox, 2012).
274 Moreover, the growth medium (and hence cell growth rate) was chosen such as to further
275 minimize gene dosage effects (Block et al., 2012). To assess whether genomic location affected
276 response to IPTG, we measured *yfp* fluorescence reporting on P_{lac} expression levels at several
277 points along the IPTG concentration gradient. If *lacI* is inserted after a weaker terminator,
278 expression from *lac* promoter is lower for a range of IPTG concentrations as compared to strain
279 where *lacI* is inserted after a strong terminator (Fig. 6B). We conducted an analysis of variance
280 (ANOVA) to compare the effect of genomic localization on P_{lac} activity. There was a significant
281 effect for five IPTG concentrations tested, and post hoc comparisons using the Tukey test were
282 performed (Fig. S9). We also directly assessed the amount of *lacI* transcript in these three
283 strains using RT-qPCR and saw differences in expression levels consistent with the observed P_{lac}
284 induction curves (Fig. 6C).
285 In this minimal network consisting of a transcription factor and the promoter it controls, the
286 network phenotype is indeed modulated only by the endogenous local genetic context, likely
287 due to varying levels of transcriptional read-through into the *lacI* gene. To verify this, we

performed PCR on cDNA from the three strains, using primers spanning the intergenic regions upstream of *lacI* (Fig. 6D). In all three cases, a DNA band corresponding to the amplification of the cDNA spanning the intergenic region was obtained, confirming read-through transcription into *lacI* from the upstream gene (no band was obtained with RNA as template). This corroborates our findings that genetic context of network elements can modulate network phenotype and that any kind of chromosomal rearrangement has the potential to alter network output.

Discussion

By comprehensively shuffling the relative TU order in a synthetic GRN, we show that local genetic context can significantly influence the phenotype of GRNs both quantitatively and qualitatively, and thus can qualitatively and quantitatively change the function a GRN performs. Hence, the phenotype of a GRN cannot be fully understood without reference to the local genetic context of its individual network components. We find that our GRN can access multiple phenotypes by simply shuffling the relative order of TUs without any changes in promoters and coding sequence. Thus, changes in regulatory connections between single regulons can be achieved solely by changes in local genetic context, which represent a category of mutations that is to be contrasted from single base pair mutations. The local genetic context is not limited to proximal transcription factors which are part of the same GRN, like our synthetic plasmid system that represents the most direct example. In fact, the local context effects and in particular transcriptional read-through can occur anywhere on the chromosome (Fig. 6),

309 indicating that any chromosomal rearrangement has the potential to alter not only the
310 expression levels of a gene but importantly also the function of a GRN.
311
312 Local genetic context of a TU can change after a deletion, duplication, insertion, inversion or
313 translocation event (Periwal and Scaria, 2015). These mutational events are often mediated by
314 mobile genetic elements, and their rates depend on the type of mobile element, the precise
315 genomic location as well as the organism (Díaz-Maldonado et al., 2015; Periwal and Scaria,
316 2015; Steinrueck and Guet, 2017). Reported rates span a wide range (from 10^{-3} to 10^{-8} per cell
317 per generation), but are typically orders of magnitude higher than rates of point mutations
318 (Hudson et al., 2002; Saito et al., 2010; Sousa et al., 2013; Tomanek et al., 2020). Given this
319 elevated frequency of small- and large-scale genomic rearrangements in various bacterial
320 species, changes in local genetic context have the potential to shape bacterial phenotypes even
321 in absence of sequence changes. Specific local genetic contexts could have arisen in response to
322 selection for changes in gene expression levels. Indeed, genomic rearrangements were found to
323 significantly change expression patterns in numerous organisms, including *E. coli*, *Bordetella*
324 *pertussis* and *Lactobacillus rhamnosus* (Brinig et al., 2006; Douillard et al., 2016; Raeside et al.,
325 2014; Weigand et al., 2017). It needs to be stressed that the impact of local genetic context of
326 GRN elements on fitness will strongly depend on the network's output. Examples of how even a
327 small effect can be strongly amplified further downstream in a regulatory network are the
328 regulatory circuit governing lysogenic and lytic states of phage lambda, or the processes behind
329 entry into sporulation or genetic competence in *Bacillus subtilis* (Dubnau and Losick, 2006;
330 Narula et al., 2012; Ptashne, 2004; Smits et al., 2005).

We find that transcriptional read-through is an important molecular mechanism behind the effects we observe in our synthetic GRN system. Changes in the strength of transcriptional termination generally require just a small number of mutations, often only individual point mutations (San Millan et al., 2009; Schuster et al., 1994; Weigand et al., 2017). Importantly, despite our synthetic GRN having been designed to restrict transcriptional read-through by using a single very strong transcriptional terminator, we observed a variety of phenotypes our GRN can access (Cambray et al., 2013). Naturally occurring transcriptional terminators cover a wide range of efficiencies and hence have potential to create a large number of regulatory connections between neighboring TUs (Cambray et al., 2013; Reynolds et al., 1992).

Most studies on chromosomal position effects focus explicitly on molecular mechanisms other than transcriptional read-through by insulating a reporter system with strong terminators (Berger et al., 2016; Block et al., 2012; Sousa et al., 1997). Given the fact that endogenous terminators vary widely in their efficiency, we argue that transcriptional read-through from neighboring genes is an inherent component of chromosomal position effects. It can significantly add to other genetic context-dependent effects resulting from gene dosage or DNA supercoiling. Such complex interplay of mechanisms can be seen in our synthetic genetic system, as not all phenotypes we see can be explained by transcriptional read-through alone and there are likely other molecular mechanisms of a more global nature influencing the phenotypes of our GRN. This observation highlights how challenging it is to disentangle all of the complex genetic interactions even in a very simplified synthetic GRN built out of the best-characterized transcription factors. Understanding simple synthetic systems helps to dissect

353 and explain the dynamics of more intricate and complex cellular interactions, following the
354 tradition of simple model systems that have been powerful throughout the history of molecular
355 biology.
356
357 Gene expression and its regulation is influenced by multiple coexisting molecular mechanisms,
358 through the concerted action of DNA binding proteins, including RNA polymerase, transcription
359 factors, topoisomerases, and nucleoid-associated proteins acting at different levels of
360 organization: from short promoter sequences to mega-base large DNA macro-domains (Junier,
361 2014; Lagomarsino et al., 2015). Here, we show that the local genetic context created by the
362 relative TU order can act as one of the genetic mechanisms shaping regulatory connections in
363 regulons (Fig. 7). Changes in local genetic context have the potential to place an individual TU
364 into two independent regulons without the need to evolve complex regulatory elements.
365 Transcriptional read-through, by enabling a diversity of gene expression profiles to be accessed
366 by shuffling of individual TUs, may be one of the mechanisms shaping the evolutionary
367 dynamics of bacterial genomes. Indeed, the fact that gene expression levels of one gene can be
368 influenced by the gene expression levels of its immediate neighbor has important
369 consequences for the evolution of operons. For a long time it has been debated whether any
370 selective advantage is gained from the physical proximity of two TUs and how this physical
371 proximity can be maintained before common transcription factor-based transcriptional
372 regulation can evolve (Lawrence and Roth, 1996). We suggest that physical proximity alone can
373 result in increased co-expression due to transcriptional read-through and thus can be
374 advantageous by changing gene expression patterns without the need for any changes in

promoter sequences or to the specificity of transcription factors. Indeed, correlated expression of genes reaching beyond the level of an operon has been recently observed (Junier et al., 2016; Junier and Rivoire, 2016). Our results also have important implications for comparative genomics, as sequence conservation does not necessarily equal functional conservation. Finally, there is a lesson for engineering living systems, as our results underscore the importance of understanding how nature itself can compute with GRNs (Guet et al., 2002; Kwok, 2010).

The simple synthetic and endogenous examples of GRNs we studied here, show how local genetic context can be a source of phenotypic diversity in GRNs, as the expression of a single gene or operon can be linked to levels and patterns of gene expression of its immediate chromosomal neighborhood. Systematic studies that utilize simple synthetic systems offer the promise of understanding how the genetic elements interact and result in the diversity of phenotypes we observe.

Supplemental material

Predicted effects of transcriptional read-through into each of the TFs

In our synthetic GRN, repressor-encoding genes, are separated by identical T1 terminators. Transcriptional read-through can in principle happen at any terminator. This scenario is also incorporated in our model. However, in our experimental approach we focused only on read-through into *cI*, which is the network element that directly regulates the level of the measured output *yfp* by binding to its promoter P_R (Fig. 1A). Transcriptional read-through into *tetR* or *lacI* from *cI* would happen only after induction with aTc, when output is insensitive to presence of LacI and TetR. Transcriptional read-through from *lacI* into *tetR* can only increase expression of *tetR*, which even without this effect fully represses P_{tet}. Transcriptional read-through from *tetR* into *lacI* could potentially make a difference when strains are grown without IPTG, but due to the negative feedback loop and tight repression of P_{tet} these effects are likely too subtle to be visible in our set-up. Thus, for the particular network topology of this study, any changes in levels of LacI and TetR due to transcriptional read-through cannot impact in any way the level of YFP.

Threshold for assigning a phenotype to individual GRNs

Figure S1 shows fluorescence levels of all the strains carrying different TU order permutations of our network. To define a phenotype that each of the GRNs can achieve, we assign a binary output value for each input state. For each strain, fluorescence was normalized to the highest expression level at the given time point. The OFF state was defined as at least three-fold repression compared to the highest ON state. Moreover, the minimal ON value in each GRN

was required to be at least three-fold greater than the maximal OFF value. A distribution of logical phenotypes for a varying threshold can be seen in Figure S6. In the mathematical model, the same procedure is followed. Expression values coming from the mathematical model are normalized by the highest expression value and then a threshold is applied to determine ON and OFF states. The threshold value is constant across all gene orders and orientations.

Cloning of 48 TU order permutations

Despite repeated attempts, cloning of eleven of the 48 TU order permutations failed. Eight of them were not clonable in either orientation with respect to the plasmid backbone. For three networks, we were able to clone only one orientation (Fig. S3B). Generally, we experienced increased difficulties in cloning GRNs where TUs were not oriented in one direction, which may at least in part be traced back to a number of highly homologous sequences in our plasmids.

A model for the impact of transcriptional read-through on gene expression

In order to test the mechanistic basis of the changes in gene expression caused by changes in TU order and orientation, we developed a mathematical model that takes into account transcriptional read-through between adjacent TUs. The basic scheme is depicted in Figure S4. We model the rate of transcript production by a constant term (k_i, the constitutive expression rate of the promoter), and an input dependent term that models the repression by other components ($-(1-I)\frac{X}{K_X+X}$) and a degradation term (-δ X).

$$\frac{dL}{dt} = k_L - (1 - I_1)\frac{L}{K_L + L} - \delta L + r_\chi^L$$

$$\frac{dT}{dt} = k_L - (1 - I_1)\frac{L}{K_L + L} - \delta T + r_\chi^T$$

$$\frac{dC}{dt} = k_C - (1 - I_2)\frac{T}{K_T + T} - \delta C + r_\chi^C$$

L, T and C correspond to *lacI*, *tetR* and *cI*, respectively, I_1 and I_2 correspond to the presence or absence of IPTG and aTc, respectively. I_1 and I_2 can take only one of two values, 0 or 1. It should be noted that *lacI* and *tetR* share the same promoter and are therefore controlled by the same rates of production.

Importantly, we include a term r_χ that models transcriptional read-through. This term takes into account the order and orientation of the specific gene network. When two genes, A and B, are adjacent to each other and share the same orientation this term will take the values $r_{\vec{AB}}^A = 0$ for gene A and $r_{\vec{AB}}^B = \mu A'$ for gene B, where A' corresponds to the rate of transcription of gene A ($k_A - (1 - Ix)\frac{A}{K_A + A}$). The output of the network is an inverse threshold function of the expression level of *cI*, so that the output is ON if *cI* expression is below the threshold τ and OFF if above it.

In order to obtain the function realized by each of the networks, this system of differential equations is solved for the steady state for the 4 possible states of the I_1 and I_2 (0,0), (0,1), (1,0), and (1,1). For each network, the expression levels are normalized by the state with the highest expression and then a threshold is applied to the expression level of *cI* to determine if the network is in an ON or OFF state (see above). For mathematical simplicity and to reduce the

number of free parameters we assumed: $k_L = k_C = 1$, $\delta=1$, $K_L = K_T = K$, leaving essentially two free parameters: the half-repression point K and the read-through rate μ.

In order to obtain the parameters region that allows this system of ordinary differentials equations to fit the experimental data, we performed a grid search in these two parameters. The results for τ =2/5 can be seen on Fig. S4C.

Bacterial strains and growth conditions

All strains used in this study are derivatives of *E. coli* MG1655 and are listed in Table S1. Plasmids are listed in Table S3. Strain and plasmid construction is described in detail below. M9CA+glycerol medium (1x M9 salts, 0.5% glycerol, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.5% casamino acids) was routinely used for bacterial growth unless otherwise stated. Selective media contained ampicillin at 100 µg ml^{-1}, kanamycin at 50 µg ml^{-1} (for plasmid located resistance cassette) or 25 µg ml^{-1} (for chromosomally located resistance cassette), and chloramphenicol at 15 µg ml^{-1}. Solid media additionally contained 1.5% (w/v) agar. β-D-thiogalactopyranoside (IPTG) was used at 1 mM unless stated otherwise, anhydrotetracycline (aTc) at 100 ng ml^{-1}.

Strain and plasmid construction

All strains and plasmids used in this study are listed in Tables S1 and S3, respectively. Strain used for measuring GRN behavior was based on TB201, which is an *E. coli* MG1655 derivative carrying a att_{P21}::P_R-*yfp* allele. Δ*lacI*785::*kanR* and Δ*lacZ*4787::*rrnB*-3 alleles were transduced (P1) into TB201 from JW0336, and the Kan^R marker removed (pCP20 (Cherepanov and

Wackernagel, 1995)), resulting in strain **ASE023**. *recA* gene of strain ASE023 was in-frame deleted using λ Red recombination (Datta et al., 2006). Kan^R marker was amplified from pKD13 (Datsenko and Wanner, 2000) and introduced into the *recA* gene of ASE023. Kan^R cassette was subsequently removed (pCP20 (Cherepanov and Wackernagel, 1995)) resulting in strain **KT131**. Strain **KT132** was used for measuring behavior of networks containing *yfp* instead of *cI* and was constructed as described above, with the parent strain being MG1655 instead of TB201. Strain **Frag1B** was used to supply constitutively expressed *tetR* encoded on the chromosome and measure orientation-dependent repression of $P_{Ltet\mathrm{O}1}$.

Plasmids were constructed by using standard cloning techniques (Sambrook and Russell, 2001) with enzymes and buffers from New England Biolabs, according to the respective protocols. All primer sequences used for this study are listed in Table S2. For cloning and plasmid propagation *E. coli* Frag1D was grown routinely in lysogeny broth (LB) at 30°C with agitation (240 rpm). All plasmids and strains were verified by sequencing.

To facilitate directional cloning of the DNA fragments, and at the same time reduce the background of clones containing empty plasmids, we inserted a DNA fragment encoding *mCherry* flanked with BglI sites into the cloning vector. This facilitated isolation of completely cut vector as well as identification of background clones on a plate due to their fluorescence. The fragment encoding *mCherry* was amplified from vector pBS3Clux (Radeck et al., 2013) and cloned into PstI and SmaI sites of vector pLA2 (Haldimann and Wanner, 2001), creating plasmid **pAS017**. The plasmid used for cloning of GRN permutations, **pAS019**, was constructed by amplifying the vector backbone (consisting of kanamycin resistance cassette, and SC101* origin

of replication) of pZS*2R-*gfp* and inserting the BglI sites-flanked mCherry cassette from pAS017 into its ScaI and EcoRI sites.

Fragments encoding $P_{Llac\mathrm{O}1}$-*tetR*, $P_{Llac\mathrm{O}1}$-*lacI*, and $P_{Ltet\mathrm{O}1}$-*cI* were amplified from the original D052 plasmid (Guet et al., 2002) and cloned into XhoI and XbaI sites of vector pZS*12-*gfp* (simultaneously removing the *gfp* cassette), resulting in plasmids **pAS014**, **pAS015** and **pAS016**, respectively. All three repressors are tagged with *ssrA* degradation tag to reduce the half-life of the proteins (Keiler et al., 1996).

For construction of the set of gene order permutations, fragments containing $P_{Llac\mathrm{O}1}$-*tetR*, $P_{Llac\mathrm{O}1}$-*lacI*, and $P_{Ltet\mathrm{O}1}$-*cI* were amplified from pAS014, pAS015 and pAS016, respectively. The fragment containing $P_{Ltet\mathrm{O}1}$-*yfp* was amplified from pZS*11-*yfp*. The primers were designed to create BglI restriction sites flanking the genes so that directional and ordered cloning was possible. Equimolar amounts of the fragments were ligated using T4 DNA ligase. The respective trimer was purified from an agarose gel and cloned into BglI sites of pAS019, resulting in plasmids **pN1-54**. **pKT10**, the empty control plasmid, was constructed by removing *mCherry* from pAS019 using XhoI and SalI and ligating the compatible overhangs. Mutations in the -10 promoter element of plasmids pAS016 and pZS*11-*yfp* were introduced by site-directed mutagenesis. Primer design and mutagenesis were performed according to the manufacturer's instructions for the QuikChange II site-directed mutagenesis kit (Agilent Technologies) resulting in plasmids **pAS023** and **pAS024**. Plasmid pAS023 served as template for construction of network plasmids with $P_{tet\text{-}10}$ mutation (**pAS026**, **pAS045-7**, and **pAS050-1**). Terminators T1T2, T*crp* and T*tonB* were cloned into XbaI site of plasmid pAS015, resulting in plasmids **pAS020**, **pAS021** and **pAS038**, respectively. These plasmids served as template for construction of

network plasmids with exchanged terminators (**pAS039**, **pAS040**, **pAS053** and **pAS055**). Promoterless *yfp* gene, and $P_{Ltet O1}$-*yfp* were amplified from pZS*11-*yfp* and cloned with BglI into pAS019 resulting in plasmids **pAS035**, **pAS036** and **pAS037**. DNA fragments between repressor genes containing terminators T1, T*crp* and T*tonB* were cloned into plasmid pAS035, resulting in plasmids **pAS041**, **pAS042** and **pAS043**, respectively.

Strains with GRNs P_{LtetO1}-*cI*-P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI* (CTL), P_{LlacO1}-*lacI*-P_{LtetO1}-*cI*-P_{LlacO1}-*tetR* (LCT) and P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*-P_{LtetO1}-*cI* (TLC) integrated into the chromosome were constructed using λ Red recombination (Datta et al., 2006). Appropriate fragments including Kan^R cassette were amplified from plasmids pN2, pN3 and pN6 and integrated into phage HK022 attachment site of ASE023 resulting in strains **ASE031**, **ASE032** and **ASE033**, respectively. Strain with three repressors integrated into separate loci on the chromosome originated from strain KT131, which was subsequently transformed with plasmids **pKT12** (carrying P_{LlacO1}-*lacI* and integrating into phage HK022 attachment site), and **pAS022** (carrying P_{LlacO1}-*tetR* and integrating into phage λ attachment site). Both pKT12 and pAS022 are based on modified CRIM plasmids (Haldimann and Wanner, 2001; Pleška et al., 2016). After each round of transformation, the Cam^R marker was removed (pCP20 (Cherepanov and Wackernagel, 1995)). The P_{LtetO1}-*cI* fragment was integrated into *26old* locus using λ Red recombination (Datta et al., 2006), resulting in strain **ASE030**.

Strain with $P_{lac(-131-410)}$-*yfp* originated from HG105, which is an *E. coli* MG1655 derivative carrying a Δ*lacZYA ΔlacI* allele (Garcia et al., 2011). HG105 was transformed with plasmid **pCC01** (carrying $P_{lac(-131-410)}$-*yfp* and integrating into phage λ attachment site) resulting in strain **ASE039**. pCC01 is based on modified CRIM plasmid (Haldimann and Wanner, 2001; Pleška et al., 2016).

534 Cam^R marker was subsequently removed (pCP20 (Cherepanov and Wackernagel, 1995)). To
535 facilitate transduction of *lacI* gene, Cam^R cassette was integrated downstream of *lacI* in strain
536 MG1655 using λ Red recombination (Datta et al., 2006), resulting in strain **ASE041**. *lacI* gene
537 and Cam^R cassette were then integrated into *flhC*, *yeaH*, and *asnT* loci, followed by Cam^R
538 cassette removal, resulting in strains **ASE046**, **ASE047**, and **ASE048**, respectively.
539
540 *Fluorescence assays*
541 YFP fluorescence of *E. coli* strains harboring different permutational GRN variants was assayed
542 using a Synergy H1 microplate reader (BioTek). Strains were grown in a 96-well plate at 30°C
543 with aeration on a microplate shaker in the dark. Routinely, M9CA+glycerol medium was used.
544 Strains ASE031, ASE032 and ASE033 were grown in LB. Overnight cultures started from single
545 colonies were diluted 1:1000 into fresh medium (supplemented with aTc and/or IPTG as
546 indicated) and grown to reach exponential phase. OD_{600} and fluorescence (excitation 515,
547 emission 545; endpoint-reads; gain 90; emission side: bottom) were recorded. For strains
548 grown in LB, cells were centrifuged and resuspended in PBS (supplemented with 1 mM $MgSO_4$
549 and 0.1 $CaCl_2$ (Tomasek, K. et al., 2018)) prior to measurements. Specific fluorescence activity is
550 given by the raw fluorescence output normalized by cell density. For GRN permutations
551 fluorescence is reported normalized to the fully unrepressed P_R promoter.
552
553 *Total RNA purification*
554 Strains were inoculated from single colonies and grown in 10 ml M9CA+glycerol medium for 8
555 hours with aeration in the dark. Total RNA was extracted from approximately $5*10^8$ cells using

RNAprotect Bacteria reagent and RNeasy Mini kit (Qiagen). Briefly, after removal of the growth medium the cells were resuspended in 500 µl M9CA+glycerol medium and 2 volumes of RNAprotect Bacteria reagent. Thereafter cells were enzymatically lysed and digested with lysozyme and Proteinase K according to the manufacturer's recommendation.

The extracted total RNA was purified from residual plasmid DNA using the DNA-*free*™ DNA removal kit (Thermo Fisher Scientific) using 4 U rDNase at 37°C for 1 hour in total. First 2 U rDNAse were added and after 30 minutes 2 more units were added for another 30 minutes incubation time. The RNA concentration was measured using the NanoDrop 200 UV-Vis spectrophotometer (NanoDrop products, Wilmington, DE) and the integrity of the purified RNA was verified on an agarose gel. RNA purity was verified using 1x One*Taq* 2x master mix and 0.2 µM primers KTp38 and KTp39 for plasmid networks and primers AS271-2, AS277 and AS280 for chromosomal *lacI* strains running an end-point PCR.

cDNA preparation and quantitative real-time PCR

cDNA was reverse transcribed using the iScript™ cDNA synthesis kit (BioRad) supplemented with random hexamers. 1 µg total RNA were used as template in a 20 µl reaction yielding approximately 50 ng/µl cDNA. As no reverse transcriptase control the reverse transcriptase reaction was performed with all components except the reverse transcriptase to verify the absence of DNA contaminations. The products of the reverse transcriptase reaction were column purified.

Measurement of transcript abundance was performed by quantitative real-time RT-PCR using the Go*Taq* qPCR Master Mix (Promega, Mannheim, D) supplemented with SYBR Green

578 according to the manufacturer's procedure with minor modifications. 100 pg cDNA for plasmid
579 networks and 500 pg of cDNA for chromosomal strains was used. Primer pairs (Table S2) were
580 designed to quantify the transcription level of *cI* and *lacI*. Expression of the kanamycin
581 resistance marker on the network plasmids and *cysG* on the chromosome were monitored as
582 constitutive references. The qPCR reaction was carried out on the BIO-RAD qPCR C1000 system
583 using 0.3 µM of the respective primers for *cI* and *kanR* amplification and 0.3 µM and 0.9 µM of
584 the respective primers for *lacI* and *cysG* amplification at an annealing temperature of 62°C. The
585 amplification efficiency, the linearity, including the slope and the R^2, and specificity of each
586 primer pair was determined by amplifying experimental triplicates of a serially dilution mixture
587 of pN1 plasmid or genomic DNA of one of the chromosomal strains (1 ng to 1 pg). Using the
588 conditions mentioned above the amplification efficiency was almost equal for all primer pairs.
589 Expression of *cI* and *lacI* was calculated as fold changes using the comparative C_T method ($\Delta\Delta C_T$)
590 (Livak and Schmittgen, 2001).
591
592 *Northern blot assay*
593 After overnight ethanol precipitation (0.1 volume 1M sodium acetate, 2.5 volume ethanol) 10
594 µg of total RNA were denatured with 2x RNA loading buffer (4% 10x TBE-DEPC, 0.02% Xylene
595 cyanol, 0.02% Bromphenol blue, 94% Formamide) for 15 minutes at 65°C. The high range RNA
596 molecular weight marker (RiboRuler High Range RNA Ladder, Thermo Fisher Scientific) was
597 dephosphorylated with FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific)
598 and 5'-end labelled with γ-32P-ATP and T4 Polynucleotide Kinase (Thermo Fisher Scientific)
599 according to the manufacturer's instructions. The denatured RNA samples and high range RNA

molecular weight marker were separated on a 1.2% denaturing agarose gel (1.2x MOPS, 19.4% formaldehyde; running buffer: 1x MOPS, 16.2% formaldehyde) for 2 hours. After rinsing the gel three times with DEPC water the RNA was blotted onto a Nylon membrane (Hybond-N+, GE Healthcare, Amersham, UK) overnight using 20x SSC (3 M NaCl, 0.3 M sodium citrate). The membrane was exposed to UV light, rinsed with 6x SSC and again exposed to UV before pre-hybridizing for 2 hours at 55°C. The pre-hybridization solution contains 0.04% BSA, 0.04% PVP, 0.04% Ficoll, 10 mM EDTA, 5x SSC, 0.2% SDS, 0.1% dextran and 0.1 mg ml-1 salmon sperm DNA. Hybridization was performed overnight at 55°C using 2 pmol of *cI* and *lacI* specific 32P 5'-end labelled oligonucleotides, respectively. The probes were labelled using γ-32P-ATP and 10 U T7 polynucleotide kinase (Thermo Fisher Scientific) for 1 hour at 37°C and column purified afterwards. The hybridization solution was identical to the pre-hybridization solution but without dextran and salmon sperm DNA. After rinsing two times with 0.5x SSC and washing 20 minutes with 1x SSC + 0.1% SDS and 15 minutes with 0.5x SSC + 0.1% SDS, the membrane was dried for 30 minutes at room temperature. The blot was exposed up to several days, and the hybridization signals were detected using Phosphorimager from Molecular Dynamics. As reference signal for normalization the kanamycin resistant gene on the network plasmids was used.

FACS

Flow cytometry was performed on a FACSCanto II analyzer (BD Biosciences, San Jose, CA). Sensitivity of the lasers was determined within the daily setup using BD FACS 7-color setup beads. For scatter detection the 488-nm laser was used: the forward scatter (FSC) detector was

set to 560 V, the side scatter (SSC) detector was set to 374 V. Both signals were collected through a 488/10-nm band-pass filter. Cells were plotted on a log scale with thresholding on FSC and SSC at 1,000. The green emission from the FITC-H channel was collected through a 530/30-nm band-pass filter using 488-nm laser and the detector was set to 473 V. The fluorescence signal observed from a physiologically distinct subpopulation, gated on FSC-H and SSC-H, was biexponentially transformed. Cells were grown as for fluorescence population measurements, and after 6 hours of growth 15 µl aliquots were frozen overnight adding 15 µl 30% glycerol in M9 buffer (1x M9 salts with Ca/Mg). After thawing, the samples were diluted in cold M9 buffer to reach an event rate of approximately 500 events/sec at medium flow rate. 20,000 events were recorded using high throughput sampler (HTS). The mean fluorescence of approximately 10,000 gated cells similar in size and shape (FSC-H) and cellular complexity (SSC-H) was determined. Events were gated and values were extracted using FlowJo software (version 10.0.7, FlowJo LLC, Tree Star).

Statistical analysis

To analyse *yfp* fluorescence data measured to assess whether genomic location affected response to IPTG, we conducted an analysis of variance (ANOVA) at each concentration separately, with the measured fluorescence as the response factor and the genomic concentration as the fixed factor (Table S4). We followed up these with a series of Tukey's multiple comparisons tests performed separately for each IPTG concentration, to directly compare the expression levels between genomic locations (Fig. S9)

644 *Data availability*

645 DNA sequences of all GRNs cloned in this study can be found in IST Research Depository under

646 DOI 10.15479/AT:ISTA:8951.

Acknowledgements

We thank J. Bollback, L. Hurst, M. Lagator, C. Nizak, O. Rivoire, M. Savageau, G. Tkacik and B. Vicozo for helpful discussions; A. Dolinar and A. Greshnova for technical assistance; T. Bollenbach for supplying the strain JW0336; C. Rusnac, and members of the Guet lab for comments. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° 628377 (ANS) and an ANR-FWF grant (CCG).

Author contributions

ANS and CCG designed research, ANS and KT constructed plasmids and strains, conducted experiments, and analyzed results, CCC assisted with plasmid and strain construction as well as data analysis, BK assisted with data analysis, ES performed Northern blot assays, TP developed model, ANS, KT and TP prepared the figures, ANS, CCG, KT, and TP wrote manuscript.

Competing interests

No competing interests declared.

References

Apirion D, Miczak A. 1993. RNA processing in prokaryotic cells. *BioEssays* **15**:113–120. doi:10.1002/bies.950150207

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**:2006.0008. doi:10.1038/msb4100050

Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA. 2004. Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* **14**:283–291. doi:10.1016/j.sbi.2004.05.004

Balaji S, Babu MM, Aravind L. 2007. Interplay between network structures, regulatory modes and sensing mechanisms of transcription factors in the transcriptional regulatory network of E. coli. *J Mol Biol* **372**:1108–1122. doi:10.1016/j.jmb.2007.06.084

Beckwith JR, Signer ER, Epstein W. 1966. Transposition of the Lac region of E. coli. *Cold Spring Harb Symp Quant Biol* **31**:393–401. doi:10.1101/sqb.1966.031.01.051

Berger M, Gerganova V, Berger P, Rapiteanu R, Lisicovas V, Dobrindt U. 2016. Genes on a Wire: The Nucleoid-Associated Protein HU Insulates Transcription Units in Escherichia coli. *Sci Rep* **6**:31512. doi:10.1038/srep31512

Block DHS, Hussein R, Liang LW, Lim HN. 2012. Regulatory consequences of gene translocation in bacteria. *Nucleic Acids Res* **40**:8979–8992. doi:10.1093/nar/gks694

Brinig MM, Cummings CA, Sanden GN, Stefanelli P, Lawrence A, Relman DA. 2006. Significant gene order and expression differences in Bordetella pertussis despite limited gene content variation. *J Bacteriol* **188**:2375–2382. doi:10.1128/JB.188.7.2375-2382.2006

Browning DF, Busby SJ. 2004. The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**:57–65. doi:10.1038/nrmicro787

Browning DF, Busby SJW. 2016. Local and global regulation of transcription initiation in bacteria. *Nat Rev Microbiol* **14**:638–650. doi:10.1038/nrmicro.2016.103

Bryant JA, Sellars LE, Busby SJW, Lee DJ. 2014. Chromosome position effects on gene expression in Escherichia coli K-12. *Nucleic Acids Res* **42**:11383–11392. doi:10.1093/nar/gku828

Cambray G, Guimaraes JC, Mutalik VK, Lam C, Mai Q-A, Thimmaiah T, Carothers JM, Arkin AP, Endy D. 2013. Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res* **41**:5139–5148. doi:10.1093/nar/gkt163

Cardinale S, Arkin AP. 2012. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnol J* **7**:856–866. doi:10.1002/biot.201200085

Chan LY, Kosuri S, Endy D. 2005. Refactoring bacteriophage T7. *Mol Syst Biol* **1**. doi:10.1038/msb4100025

Chen HD, Jewett MW, Groisman EA. 2012. An allele of an ancestral transcription factor dependent on a horizontally acquired gene product. *PLoS Genet* **8**:e1003060. doi:10.1371/journal.pgen.1003060

Cherepanov PP, Wackernagel W. 1995. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**:9–14.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**:6640–6645. doi:10.1073/pnas.120163297

Datta S, Costantino N, Court DL. 2006. A set of recombineering plasmids for gram-negative bacteria. *Gene* **379**:109–115. doi:10.1016/j.gene.2006.04.018

Díaz-Maldonado H, Gómez MJ, Moreno-Paz M, San Martín-Úriz P, Amils R, Parro V, López de Saro FJ. 2015. Transposase interaction with the β sliding clamp: effects on insertion sequence proliferation and transposition rate. *Sci Rep* **5**:13329. doi:10.1038/srep13329

709 Douillard FP, Ribbera A, Xiao K, Ritari J, Rasinkangas P, Paulin L, Palva A, Hao Y, de Vos WM. 2016.
710 Polymorphisms, Chromosomal Rearrangements, and Mutator Phenotype Development during
711 Experimental Evolution of Lactobacillus rhamnosus GG. *Appl Environ Microbiol* **82**:3783–3792.
712 doi:10.1128/AEM.00255-16
713 Dubnau D, Losick R. 2006. Bistability in bacteria. *Mol Microbiol* **61**:564–572. doi:10.1111/j.1365-
714 2958.2006.05249.x
715 Garcia HG, Lee HJ, Boedicker JQ, Phillips R. 2011. Comparison and calibration of different reporters for
716 quantitative analysis of gene expression. *Biophys J* **101**:535–544. doi:10.1016/j.bpj.2011.06.026
717 Georg J, Hess WR. 2011. cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol*
718 *Biol Rev MMBR* **75**:286–300. doi:10.1128/MMBR.00032-10
719 Guet CC, Elowitz MB, Hsing W, Leibler S. 2002. Combinatorial synthesis of genetic networks. *Science*
720 **296**:1466–1470. doi:10.1126/science.1067407
721 Haldimann A, Wanner BL. 2001. Conditional-replication, integration, excision, and retrieval plasmid-host
722 systems for gene structure-function studies of bacteria. *J Bacteriol* **183**:6384–6393.
723 doi:10.1128/JB.183.21.6384-6393.2001
724 Hudson RE, Bergthorsson U, Roth JR, Ochman H. 2002. Effect of Chromosome Location on Bacterial
725 Mutation Rates. *Mol Biol Evol* **19**:85–92. doi:10.1093/oxfordjournals.molbev.a003986
726 Igler C, Lagator M, Tkačik G, Bollback JP, Guet CC. 2018. Evolutionary potential of transcription factors
727 for gene regulatory rewiring. *Nat Ecol Evol* **2**:1633–1643. doi:10.1038/s41559-018-0651-y
728 Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, Raineri E, Garriga-Canut M, Serrano L. 2008.
729 Evolvability and hierarchy in rewired bacterial gene networks. *Nature* **452**:840–845.
730 doi:10.1038/nature06847
731 Jacob F, Monod J. 1961. On the Regulation of Gene Activity. *Cold Spring Harb Symp Quant Biol* **26**:193–
732 211. doi:10.1101/SQB.1961.026.01.024
733 Junier I. 2014. Conserved patterns in bacterial genomes: a conundrum physically tailored by
734 evolutionary tinkering. *Comput Biol Chem* **53 Pt A**:125–133.
735 doi:10.1016/j.compbiolchem.2014.08.017
736 Junier I, Rivoire O. 2016. Conserved Units of Co-Expression in Bacterial Genomes: An Evolutionary Insight
737 into Transcriptional Regulation. *PloS One* **11**:e0155740. doi:10.1371/journal.pone.0155740
738 Junier I, Unal EB, Yus E, Lloréns-Rico V, Serrano L. 2016. Insights into the Mechanisms of Basal
739 Coordination of Transcription Using a Genome-Reduced Bacterium. *Cell Syst* **2**:391–401.
740 doi:10.1016/j.cels.2016.04.015
741 Keiler KC, Waller PR, Sauer RT. 1996. Role of a peptide tagging system in degradation of proteins
742 synthesized from damaged messenger RNA. *Science* **271**:990–993.
743 Kuhlman TE, Cox EC. 2012. Gene location and DNA density determine transcription factor distributions
744 in Escherichia coli. *Mol Syst Biol* **8**:610. doi:10.1038/msb.2012.42
745 Kwok R. 2010. Five hard truths for synthetic biology. *Nature* **463**:288–290. doi:10.1038/463288a
746 Lagomarsino MC, Espéli O, Junier I. 2015. From structure to function of bacterial chromosomes:
747 Evolutionary perspectives and ideas for new experiments. *FEBS Lett* **589**:2996–3004.
748 doi:10.1016/j.febslet.2015.07.002
749 Lawrence JG, Roth JR. 1996. Selfish operons: horizontal transfer may drive the evolution of gene
750 clusters. *Genetics* **143**:1843–1860.
751 Lim HN, Lee Y, Hussein R. 2011. Fundamental relationship between operon organization and gene
752 expression. *Proc Natl Acad Sci U S A* **108**:10626–10631. doi:10.1073/pnas.1105692108
753 Liu LF, Wang JC. 1987. Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A*
754 **84**:7024–7027.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods San Diego Calif* **25**:402–408. doi:10.1006/meth.2001.1262

Lutz R, Bujard H. 1997. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**:1203–1210.

Mangan S, Zaslaver A, Alon U. 2003. The Coherent Feedforward Loop Serves as a Sign-sensitive Delay Element in Transcription Networks. *J Mol Biol* **334**:197–204. doi:10.1016/j.jmb.2003.09.049

Meyer S, Reverchon S, Nasser W, Muskhelishvili G. 2018. Chromosomal organization of transcription: in a nutshell. *Curr Genet* **64**:555–565. doi:10.1007/s00294-017-0785-5

Mirkin EV, Castro Roa D, Nudler E, Mirkin SM. 2006. Transcription regulatory elements are punctuation marks for DNA replication. *Proc Natl Acad Sci U S A* **103**:7276–7281. doi:10.1073/pnas.0601127103

Mukherji S, van Oudenaarden A. 2009. Synthetic biology: Understanding biological design from synthetic circuits. *Nat Rev Genet* **10**:859–871. doi:10.1038/nrg2697

Narula J, Devi SN, Fujita M, Igoshin OA. 2012. Ultrasensitivity of the Bacillus subtilis sporulation decision. *Proc Natl Acad Sci U S A* **109**:E3513-3522. doi:10.1073/pnas.1213974109

Nocedal I, Mancera E, Johnson AD. 2017. Gene regulatory network plasticity predates a switch in function of a conserved transcription regulator. *eLife* **6**. doi:10.7554/eLife.23250

Orosz A, Boros I, Venetianer P. 1991. Analysis of the complex transcription termination region of the Escherichia coli rrnB gene. *Eur J Biochem* **201**:653–659.

Palmer AC, Egan JB, Shearwin KE. 2011. Transcriptional interference by RNA polymerase pausing and dislodgement of transcription factors. *Transcription* **2**:9–14. doi:10.4161/trns.2.1.13511

Periwal V, Scaria V. 2015. Insights into structural variations and genome rearrangements in prokaryotic genomes. *Bioinforma Oxf Engl* **31**:1–9. doi:10.1093/bioinformatics/btu600

Pleška M, Qian L, Okura R, Bergmiller T, Wakamoto Y, Kussell E, Guet CC. 2016. Bacterial Autoimmunity Due to a Restriction-Modification System. *Curr Biol CB* **26**:404–409. doi:10.1016/j.cub.2015.12.041

Prud'homme B, Gompel N, Rokas A, Kassner VA, Williams TM, Yeh S-D, True JR, Carroll SB. 2006. Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* **440**:1050–1053. doi:10.1038/nature04597

Ptashne M. 2004. A genetic switch: phage lambda revisited, 3rd ed. ed. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press.

Radeck J, Kraft K, Bartels J, Cikovic T, Dürr F, Emenegger J, Kelterborn S, Sauer C, Fritz G, Gebhard S, Mascher T. 2013. The Bacillus BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with Bacillus subtilis. *J Biol Eng* **7**:29. doi:10.1186/1754-1611-7-29

Raeside C, Gaffé J, Deatherage DE, Tenaillon O, Briska AM, Ptashkin RN, Cruveiller S, Médigue C, Lenski RE, Barrick JE, Schneider D. 2014. Large chromosomal rearrangements during a long-term evolution experiment with Escherichia coli. *mBio* **5**:e01377-01314. doi:10.1128/mBio.01377-14

Reynolds R, Bermúdez-Cruz RM, Chamberlin MJ. 1992. Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators. *J Mol Biol* **224**:31–51.

Riordan, J. 2003. Introduction to Combinatorial Analysis. Dover Books on Mathematics.

Saito T, Chibazakura T, Takahashi K, Yoshikawa H, Sekine Y. 2010. Measurements of transposition frequency of insertion sequence IS1 by GFP hop-on assay. *J Gen Appl Microbiol* **56**:187–192.

Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual, 3rd ed. ed. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press.

San Millan A, Depardieu F, Godreuil S, Courvalin P. 2009. VanB-Type Enterococcus faecium Clinical Isolate Successively Inducibly Resistant to, Dependent on, and Constitutively Resistant to Vancomycin. *Antimicrob Agents Chemother* **53**:1974–1982. doi:10.1128/AAC.00034-09

Scholz SA, Diao R, Wolfe MB, Fivenson EM, Lin XN, Freddolino PL. 2019. High-Resolution Mapping of the Escherichia coli Chromosome Reveals Positions of High and Low Transcription. *Cell Syst* **8**:212-225.e9. doi:10.1016/j.cels.2019.02.004

Schuster P, Fontana W, Stadler PF, Hofacker IL. 1994. From sequences to shapes and back: a case study in RNA secondary structures. *Proc Biol Sci* **255**:279–284. doi:10.1098/rspb.1994.0040

Shearwin KE, Callen BP, Egan JB. 2005. Transcriptional interference – a crash course. *Trends Genet* **21**:339–345. doi:10.1016/j.tig.2005.04.009

Shubin N, Tabin C, Carroll S. 2009. Deep homology and the origins of evolutionary novelty. *Nature*. doi:10.1038/nature07891

Smits WK, Eschevins CC, Susanna KA, Bron S, Kuipers OP, Hamoen LW. 2005. Stripping Bacillus: ComK auto-stimulation is responsible for the bistable response in competence development. *Mol Microbiol* **56**:604–614. doi:10.1111/j.1365-2958.2005.04488.x

Sobetzko P. 2016. Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes. *Nucleic Acids Res* **44**:1514–1524. doi:10.1093/nar/gkw007

Sobetzko P, Travers A, Muskhelishvili G. 2012. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc Natl Acad Sci U S A* **109**:E42-50. doi:10.1073/pnas.1108229109

Sousa A, Bourgard C, Wahl LM, Gordo I. 2013. Rates of transposition in Escherichia coli. *Biol Lett* **9**:20130838. doi:10.1098/rsbl.2013.0838

Sousa C, de Lorenzo V, Cebolla A. 1997. Modulation of gene expression through chromosomal positioning in Escherichia coli. *Microbiol Read Engl* **143 (Pt 6)**:2071–2078. doi:10.1099/00221287-143-6-2071

Steinrueck M, Guet CC. 2017. Complex chromosomal neighborhood effects determine the adaptive potential of a gene under selection. *eLife* **6**. doi:10.7554/eLife.25100

Stern DL, Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution? *Evolution* **62**:2155–2177. doi:10.1111/j.1558-5646.2008.00450.x

Szeberényi J, Roy MK, Vaidya HC, Apirion D. 1984. 7S RNA, containing 5S ribosomal RNA and the termination stem, is a specific substrate for the two RNA processing enzymes RNase III and RNase E. *Biochemistry (Mosc)* **23**:2952–2957.

Tomanek I, Grah R, Lagator M, Andersson AMC, Bollback JP, Tkačik G, Guet CC. 2020. Gene amplification as a form of population-level gene expression regulation. *Nat Ecol Evol* **in revision**.

Tomasek, K., Bergmiller, T., Guet, C. C. 2018. Lack of cations in flow cytometry buffers affect fluorescence signals by reducing membrane stability and viability of Escherichia coli strains. *J Biotechnol* **268**:40–52.

Vora T, Hottes AK, Tavazoie S. 2009. Protein occupancy landscape of a bacterial genome. *Mol Cell* **35**:247–253. doi:10.1016/j.molcel.2009.06.035

Wagner GP, Lynch VJ. 2010. Evolutionary novelties. *Curr Biol* **20**:R48–R52. doi:10.1016/j.cub.2009.11.010

Weigand MR, Peng Y, Loparev V, Batra D, Bowden KE, Burroughs M, Cassiday PK, Davis JK, Johnson T, Juieng P, Knipe K, Mathis MH, Pruitt AM, Rowe L, Sheth M, Tondella ML, Williams MM. 2017. The History of Bordetella pertussis Genome Evolution Includes Structural Rearrangement. *J Bacteriol* **199**. doi:10.1128/JB.00806-16

Wolf DM, Arkin AP. 2003. Motifs, modules and games in bacteria. *Curr Opin Microbiol* **6**:125–134.

Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* **8**:206–216. doi:10.1038/nrg2063

Wu H-Y, Fang M. 2003. DNA supercoiling and transcription control: a model from the study of suppression of the leu-500 mutation in Salmonella typhimurium topA- strains. *Prog Nucleic Acid Res Mol Biol* **73**:43–68.

Wu K, Rao CV. 2010. The role of configuration and coupling in autoregulatory gene circuits. *Mol Microbiol* **75**:513–527. doi:10.1111/j.1365-2958.2009.07011.x

Yeung E, Dy AJ, Martin KB, Ng AH, Del Vecchio D, Beck JL, Collins JJ, Murray RM. 2017. Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks. *Cell Syst* **5**:11-24.e12. doi:10.1016/j.cels.2017.06.001

Zipser D. 1969. Polar Mutations and Operon Function. *Nature* **221**:21–25. doi:10.1038/221021a0

Table S1. Strains used in this study.

Name	Genotype[a]	Source
DH5α	F$^-$,*Φ80lacZΔM15, Δ(lacZYA-argF), U169, recA1, endA1, hsdR17 (rK–, mK+), phoA supE44, thi-1, gyrA96, relA1, λ*$^-$	Laboratory stock
DH5α λ*pir*+	F$^-$,*Φ80lacZΔM15, Δ(lacZYA-argF), U169, recA1, endA1, hsdR17 (rK–, mK+), phoA supE44, thi-1, gyrA96, relA1, λpir*$^+$	Laboratory stock
HG105	MG1655 Δ*lacZYA* Δ*lacI*	(Garcia et al., 2011)
JW0336	F$^-$, Δ(*araD-araB*)567, Δ*lacZ*4787(::*rrnB*-3), Δ*lacI*785::kan, λ$^-$, *rph*-1, Δ(*rhaD-rhaB*)568, *hsdR*514	(Baba et al., 2006)
MG1655	F-, λ$^-$, *ilvG*-, *rfb*-50, *rph*-1	Laboratory stock
TB201	MG1655 att$_{P21}$::P$_R$-*yfp*	(Pleška et al., 2016)
Frag1B	F$^-$, *lacZ82*(Am), λ$^-$, *rha-4, thiE*, *gal-33, P$_{N25}$/tetR placi^q/lacI Sp^R*	Laboratory stock
Frag1D	F$^-$, *lacZ82*(Am), λ$^-$, *rha-4, thiE*, *gal-33, P$_{N25}$/tetR placi^q/lacI Sp^R*, Δ*recA*	Laboratory stock
ASE023	MG1655 att$_{P21}$::P$_R$-*yfp* Δ*lacI*785 Δ*lacZ*4787::*rrnB*-3	This study
ASE030	MG1655 att$_{P21}$::P$_R$-*yfp* Δ*lacI*785 Δ*lacZ*4787::*rrnB*-3 Δ*recA* att$_{HK022}$::P$_{lacO1}$-*lacI* att$_\lambda$::P$_{lacO1}$-*tetR* 3939*old*::P$_{tetO1}$-*cI* Cam^R	This study
ASE031	MG1655 att$_{P21}$::P$_R$-*yfp* Δ*lacI*785 Δ*lacZ*4787::*rrnB*-3 att$_{HK022}$::P$_{LtetO1}$-*cI*-P$_{LlacO1}$-*tetR*-P$_{LlacO1}$-*lacI* (CTL)	This study
ASE032	MG1655 att$_{P21}$::P$_R$-*yfp* Δ*lacI*785 Δ*lacZ*4787::*rrnB*-3 att$_{HK022}$::P$_{LlacO1}$-*lacI*-P$_{LtetO1}$-*cI*-P$_{LlacO1}$-*tetR* (LCT)	This study
ASE033	MG1655 att$_{P21}$::P$_R$-*yfp* Δ*lacI*785 Δ*lacZ*4787::*rrnB*-3 att$_{HK022}$::P$_{LlacO1}$-*tetR*-P$_{LlacO1}$-*lacI*-P$_{LtetO1}$-*cI* (TLC)	This study
ASE039	MG1655 Δ*lacZYA* Δ*lacI* att$_\lambda$::P$_{lac\text{-}131\text{-}410}$-*yfp*	This study
ASE041	MG1655 P*lac*::Cam^R-FRT	This study
ASE046	MG1655 Δ*lacZYA* Δ*lacI* att$_\lambda$::P$_{lac\text{-}131\text{-}410}$-*yfp fhlC*::*lacI*	This study
ASE047	MG1655 Δ*lacZYA* Δ*lacI* att$_\lambda$::P$_{lac\text{-}131\text{-}410}$-*yfp yeaH*::*lacI*	This study
ASE048	MG1655 Δ*lacZYA* Δ*lacI* att$_\lambda$::P$_{lac\text{-}131\text{-}410}$-*yfp asnT*::*lacI*	This study
KT131	MG1655 att$_{P21}$::P$_R$-*yfp* Δ*lacI*785 Δ*lacZ*4787::*rrnB*-3 Δ*recA*	This study
KT132	MG1655 Δ*lacI*785 Δ*lacZ*4787::*rrnB*-3 Δ*recA*	This study

[a] Amp – ampicillin resistance, Kan – kanamycin resistance, Sp – spectinomycin resistance, Cam – chloramphenicol resistance

Table S2. Oligonucleotides used in this study.

Name	Sequence (5'-3')[a]	Use
AS096	GTACCTGCAGGACGTCAGTACTGCCA*CGG*AGGC**ATACGCAAACCGCCTCTCCC**	*mCherry* up
AS097	GTACCCCGGGAGATCTGTCGACGCCT*AGG*TGGC**CTACTAGTATATAAACGCAG**	*mCherry* down
AS098	ATTCGCCA*CGG*AGGCA**CCTTTCGTCTTCACC**	Fragment 1F, up[b]
AS099	GACAGCCT*GAA*TGGC**TCTAGGGCGGCGGATTTG**	Fragment 1F, down[b]
AS100	ATTCGCCA*TTC*AGGCA**CCTTTCGTCTTCACC**	Fragment 2F, up[b]
AS101	GACAGCCT*CAC*TGGC**TCTAGGGCGGCGGATTTG**	Fragment 2F, down[b]
AS102	ATTCGCCA*GTG*AGGCA**CCTTTCGTCTTCACC**	Fragment 3F, up[b]
AS103	GACAGCCT*AGG*TGGC**TCTAGGGCGGCGGATTTG**	Fragment 3F, down[b]
AS104	**TTCACCTCGAGAATTGTGAGC**	P_{LlacO1}-*lacI/tetR*, up
AS105	GACA**TCTAGATTAAGCTGCTAAAGCGTAG**	P_{LlacO1}-*lacI/tetR*, down
AS106	**TTCACCTCGAGTCCCTATCAG**	P_{LtetO1}-*cI*, up
AS107	**GGATCCTCTAGATCAAGCTGC**	P_{LtetO1}-*cI*, down
AS129	ATTCGCCA*CGG*AGGC**TCTAGGGCGGCGGATTTG**	Fragment 1R, up[b]
AS130	GACAGCCT*GAA*TGGCT**CCTTTCGTCTTCACC**	Fragment 1R, down[b]
AS131	ATTCGCCA*TTC*AGGC**TCTAGGGCGGCGGATTTG**	Fragment 2R, up[b]
AS132	GACAGCC*TCA*CTGGCTC**CTTTCGTCTTCACC**	Fragment 2R, down[b]
AS133	ATTCGCCA*GTG*AGGC**TCTAGGGCGGCGGATTTG**	Fragment 3R, up[b]
AS134	GACAGCCT*AGG*TGGCTC**CTTTCGTCTTCACC**	Fragment 3R, down[b]
AS135	GCATGAATTC**GGCTGTTCTGGTGTTGCTAG**	pZS*2 backbone, up
AS136	GTCAAGTACT**CTCGAGGTGAAGACGAAAGG**	pZS*2 backbone, down
AS142	CTAGAAAAAGCCTCCGACCGGAGGCTTTTGT	T*tonB*, fwd
AS143	CTAGACAAAAGCCTCCGGTCGGAGGCTTTTT	T*tonB*, rev
AS144	CTAGATGGCGCGTTACCTGGTAGCGCGCCATTTTGTTTT	T*crp*, fwd
AS145	CTAGAAAACAAAATGGCGCGCTACCAGGTAACGCGCCAT	T*crp*, rev
AS147	GACAGCCT*CAC*TGGCA**CCTCTAGAAAACAAAATGGC**	Fragment 2F-T*crp*, down[b]
AS153	GACAGCCT*CAC*TGGCA**CCTCTAGACAAAAGCCTCCG**	Fragment 2F-T*tonB*, down[b]
AS159	**TCCTCCTTAGTTCCTATTCC**	Integration of P_{tetO1}-*cI* into *folD*, *cam* up
AS160	ACTTCGGAATAGGAACTAAGGAGGA**AAATAGGCGTATCACGAGGC**	Integration of P_{tetO1}-*cI* into *folD*, *cI* up
AS161	CATCAATAATAAGGCTTTATGCTAGATGCATTCCGCTTTGCGACTCAACC**ACTAGCAACACCAGAACAGC**	Integration of networks into att_{HK022}, up
AS167	CATCCAGAGTCTTCGGGTCAGGGTTAAATTCACGGTCGGTGCACTTTAG**GGCTTACCCGTCTTACTGTCC**	Integration of networks into att_{HK022}, down
AS168	CACCGTCGCTGAGACTGAAAGCTTCATTTTTCGTCCATGATGGCGTTGTA**GAAAAGTGCCACCTGCATCG**	Integration of P_{LtetO1}-*cI* into *folD*, *cam* down

AS170	TTCTTAAATTATCTTAATCCTTAGACAAGGAAATAAATCAGTTCC AGATT**TAGATCAAGCTGCTAAAGCG**	Integration of P_{LtetO1}-*cI* into *folD*, *cI* down
AS171	**CATCCCTATCAGTGATAGAG**CT**AC**G**GAGCACATCAGCAGGAC GC**	A-11C and T-7G substitution, fwd
AS172	**GCGTCCTGCTGATGTGCTC**C**GTA**G**CTCTATCACTGATAGGGAT G**	A-11C and T-7G substitution, rev
AS195	GACAGCCT*AGG*TGGC**ATTAAAGAGGAGAAAGGTACC**	Promoterless *yfp*, fwd
AS198	ATTAGCATGC**TAATAGGTATCCTATGATTA**	wt P_{lac}, fwd
AS199	TAATGAATTC**TCCTTCTCGATCCGAGACGA**	wt P_{lac}, rev
AS206	CATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGG GCAGTGA**CATATGAATATCCTCCTTAG**	Integration of Cam^R into P_{lac}, fwd
AS207	AAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC GTTGCGC**ACAGCTGCAGGCATGCAAGC**	Integration of Cam^R into P_{lac}, rev
AS221	TAAAACTATCAGCCAGGTCATTATCGCCTGGCTGATTTTTAGCT TACTGT**CACCATCGAATGGCGCAAAA**	Integration of *lacI* into *yeaH* locus, fwd
AS222	TCATATTTAAAGCGATTGTAAGCTAATGTATGTAATAAATGAGA TAATTT**ACAGCTGCAGGCATGCAAGC**	Integration of *lacI* into *yeaH* locus, rev
AS225	CATCTTAAGCGCCCTCGACCTTTATGGTTGAGGGCGTTTTGCTA TGAACG**CACCATCGAATGGCGCAAAA**	Integration of *lacI* into *asnT* locus, fwd
AS226	GAGACTACTGAATAACTCAAGTTTTATAATCGAGGGGAAAATG GTGATGG**ACAGCTGCAGGCATGCAAGC**	Integration of *lacI* into *asnT* locus, rev
AS229	GCAACATTCCAGCAGCGGTAACGACGTACCGCTGCTTTTTTTTG CCCCAA**CACCATCGAATGGCGCAAAA**	Integration of *lacI* into *flhC* locus, fwd
AS230	TCCACTGTTGACCATGACAGGATGTTCAGTCGTCAGGCGTTAAC GCGCGA**ACAGCTGCAGGCATGCAAGC**	Integration of *lacI* into *flhC* locus, rev
AS267	TTGTCGGCGGTGGTGATGTC	qPCR *cysG*, up
AS268	ATGCGGTGAACTGTGGAATAAACG	qPCR *cysG*, down
AS271	CGTGCACATCAGACATTGTG	*yeaH-lacI* junction, fwd
AS272	CGTTTTCGCAGAAACGTGGC	*lacI* junctions, rev
AS277	GCCTGCAGCTTATGTCAACC	*flhC-lacI* junction, fwd
AS280	GTCACTGACCTTAGTTGAAC	*asnT-lacI* junction, fwd
AS282	GACAGCCT*CAC*TGGCT**CTAGGGAAGAGTTTGT**	Fragment 2F-T1T2, down[b]
AS283	TCGACTCTAGATGGCGCGTTACCTGGTAGCGCGCCATTTTGTTT TCTAGAGGTGCCAGTGAGGCACCTTTCGTCTTCACCTCGAGG	T*crp*, fwd
AS284	TCGACCTCGAGGTGAAGACGAAAGGTGCCTCACTGGCACCTCT AGAAAACAAAATGGCGCGCTACCAGGTAACGCGCCATCTAGA G	T*crp*, rev
AS285	TCGACTCTAGAAAAAGCCTCCGACCGGAGGCTTTTGTCTAGAG GTGCCAGTGAGGCACCTTTCGTCTTCACCTCGAGG	T*tonB*, fwd
AS286	TCGACCTCGAGGTGAAGACGAAAGGTGCCTCACTGGCACCTCT AGACAAAAGCCTCCGGTCGGAGGCTTTTTCTAGAG	T*tonB*, rev
KTp38	**TCAGTGATAGAGATTGACATCCCT**	RNA purity, *cI* up

KTp39	**CCCCACAACGGAACAACTCT**	RNA purity, *cI* down
KTp45	GACAGCCT*CAC*TGGCTC**TTAAGCTGCTAAAGCGTAG**	Fragment 2 without T1, down[b]
KTp46	ATTCGCCA*TTC*AGGCTC**TTAAGCTGCTAAAGCGTAG**	Fragment 2R without T1, up[b]
KTp65	**GCTGTTGAGCCAGGTGATTT**	qPCR *cI*, up
KTp66	**GGGATCATTGGGTACTGTGG**	qPCR *cI*, down
KTp67	**AATACGCAAACCGCCTCTC**	qPCR *lacI*, up
KTp68	**CAGTCGGGAAACCTGTCGT**	qPCR *lacI*, down
KTp71	**GTTGTCACTGAAGCGGGAAG**	qPCR *kanR*, up
KTp72	**GCAAGGTGAGATGACAGGAGA**	qPCR *kanR*, down
KTp73	**ACTCATCACCCCCAAGTCTG**	Northern blot, *cI* probe 1
KTp74	**GGATCATTGGGTACTGTGGG**	Northern blot, *cI* probe 2
KTp75	**CCTGACTGCCCCATCCCC**	Northern blot, *cI* probe 3
KTp76	**CTCGTCCTGCAGTTCATTCA**	Northern blot, *kanR* probe 1
KTp77	**GCCAACGCTATGTCCTG**	Northern blot, *kanR* probe 2
KTp90	ATGGATCCT**ATTAAGCTGCTAAAGC**	P_{LlacO1}-*tetR/lacI*, down
KTp91	ATGCATGC**TCGAGAATTGTGAGC**	P_{LlacO1}-*tetR/lacI*, up
KTp93	**CGGTTTGCGTATTGGGCG**	Northern blot, *lacI* probe 1
KTp94	**AGAAGATTGTGCACCGCC**	Northern blot, *lacI* probe 2
5'_recA	TGACTATCCGGTATTACCCGGCATGACAGGAGTAAAAATG**GGGGATCCGTCGACCTGCAGTT**	*recA* deletion, up
3'_recA	AAGGGCCGCAGATGCGACCCTTGTGTATCAAACAAGACGA**TGTAGGCTGGAGCTGCTTC**	*recA* deletion, down

[a] Restriction sites are underlined; sequences of overhangs produced after restriction with BglI are italicized, annealing sequences are shown in bold.

[b] Fragment number refers to gene position in the three gene array; F stands for "forward" orientation (opposite to the orientation of the kanamycin cassette), R for "reverse" orientation.

Table S3. Plasmids used in this study.

Name	Description[a]	Primers and enzymes used for cloning	Source
pAS014	*pSC101* ori*, Amp^R, P_{LlacO1}-*tetR* (T)	AS104/AS105; XhoI+XbaI	This study
pAS015	*pSC101* ori*, Amp^R, P_{LlacO1}-*lacI* (L)	AS104/AS105; XhoI+XbaI	This study
pAS016	*pSC101* ori*, Amp^R, P_{LtetO1}-*cI* (C)	AS106/AS107; XhoI+XbaI	This study
pAS017	*R6K* ori, Kan^R, *attP*$_\lambda$, P_{lac}-*mCherry*	AS096/AS097; PstI+SmaI	This study
pAS019	*pSC101* ori*, Kan^R, *mCherry*	AS135/AS136; ScaI+EcoRI	This study
pAS020	*pSC101* ori*, Amp^R, P_{LlacO1}-*lacI*-T*crp*-T1	AS142/AS143; XbaI	This study
pAS021	*pSC101* ori*, Amp^R, P_{LlacO1}-*lacI*-T*tonB*-T1	AS144/AS145; XbaI	This study
pAS022	*R6K* ori, Cam^R-*frt*, *attP*$_\lambda$, P_{LlacO1}-*tetR* (from pKT11)	BamHI+SphI	This study
pAS023	*pSC101* ori*, Amp^R, $P_{LtetO1(A\text{-}11C,\ T\text{-}7G)}$-*cI*	AS171/AS172	This study
pAS024	*pSC101* ori*, Amp^R, $P_{LtetO1(A\text{-}11C,\ T\text{-}7G)}$-*yfp*	AS171/AS172	This study
pAS026	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*-$P_{LtetO1(A\text{-}11C,\ T\text{-}7G)}$-*cI*	AS098/AS099, AS100/AS101, AS102/AS103; BglI	This study
pAS035	*pSC101* ori*, Kan^R, *yfp*$_{rev}$	AS129/AS195; BglI	This study
pAS036	*pSC101* ori*, Kan^R, P_{LtetO1}-*yfp*	AS098/AS103; BglI	This study
pAS037	*pSC101* ori*, Kan^R, P_{LtetO1}-*yfp*$_{rev}$	AS129/AS134; BglI	This study
pAS038	*pSC101* ori*, Amp^R, P_{LlacO1}-*lacI*-T1T2	XbaI	This study
pAS039	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI-T1T2*-P_{LtetO1}-*cI*	AS098/AS099, AS100/AS282, AS102/AS103; BglI	This study
pAS040	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-T*crp*-T1-P_{LlacO1}-*lacI*-P_{LtetO1}-*cI*	AS098/AS099, AS100/AS101, AS102/AS103; BglI	This study
pAS041	*pSC101* ori*, Kan^R, T1-*yfp*	SalI	This study
pAS042	*pSC101* ori*, Kan^R, T*crp-yfp*	AS283/AS284; SalI	This study

pAS043	*pSC101* ori*, Kan^R, T*tonB-yfp*	AS285/AS286; SalI	This study
pAS045	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-$P_{LtetO1(A-11C,T-7G)}$-*cI*-P_{LlacO1}-*tetR*	AS098/AS099, AS100/AS101, AS102/AS103; BglI	This study
pAS046	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*$_r$-$P_{LtetO1(A-11C,T-7G)}$-*cI*$_r$-P_{LlacO1}-*lacI*$_r$	AS129/AS130, AS131/AS132, AS133/AS134; BglI	This study
pAS047	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-$P_{LtetO1(A-11C,T-7G)}$-*cI*-P_{LlacO1}-*lacI*	AS098/AS099, AS100/AS101, AS102/AS103; BglI	This study
pAS050	*pSC101* ori*, Kan^R, $P_{LtetO1(A-11C,T-7G)}$-*cI*$_r$-P_{LlacO1}-*lacI*$_r$-PLlacO1-*tetR*$_r$	AS129/AS130, AS131/AS132, AS133/AS134; BglI	This study
pAS051	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*$_r$-$P_{LtetO1(A-11C,T-7G)}$-*cI*$_r$-PLlacO1-*tetR*$_r$	AS129/AS130, AS131/AS132, AS133/AS134; BglI	This study
pAS053	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*-T*crp*-P_{LtetO1}-*cI*	AS098/AS099, AS100/AS147, AS102/AS103; BglI	This study
pAS055	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*-T*tonB*-P_{LtetO1}-*cI*	AS098/AS099, AS100/AS153, AS102/AS103; BglI	This study
pCC01	*R6K* ori, Cam^R-*frt*, *attP*$_\lambda$, P_{lac}-*yfp*	AS198/AS199; EcoRI+SphI	This study
pKT10	*pSC101* ori*, Kan^R	XhoI+SalI	This study
pKT11	*R6K* ori, Cam^R-*frt*, *attP*$_{P21}$, P_{LlacO1}-*tetR*	KTp90/KTp91; BamHI+SphI	This study
pKT12	*R6K* ori, Cam^R-*frt*, *attP*$_{HK022}$, P_{LlacO1}-*lacI*	KTp90/KTp91; BamHI+SphI	This study
D052	*pSC101* ori*, Amp^R, P_{LlacO1}-*lacI*$_f$-P_{LtetO1}-*cI*-P_{LlacO1}-*lacI*-P_R-*gfp*		(Guet et al., 2002)
pN1	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*-P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR* (CLT)	AS098/AS099, AS100/AS101, AS102/AS103; BglI	This study
pN2	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*-P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI* (CTL)		This study

pN3	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-P_{LtetO1}-*cI*-P_{LlacO1}-*tetR* (LCT)		This study
pN4	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR*-P_{LtetO1}-*cI* (LTC)		This study
pN5	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LtetO1}-*cI*-P_{LlacO1}-*lacI* (TCL)		This study
pN6	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*-P_{LtetO1}-*cI* (TLC)		This study
pN9	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*tetR* (LC_rT)	AS098/AS099, KTp46/AS132, AS102/AS103; BglII	This study
pN10	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI* (LT_rC)		This study
pN11	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*lacI* (TC_rL)		This study
pN12	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI* (TL_rC)		This study
pN13	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*-P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR*$_r$ (CLT_r)	AS098/AS099, AS100/KTp45, AS133/AS134; BglII	This study
pN14	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*-P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*$_r$ (CTL_r)		This study
pN16	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR*-P_{LtetO1}-*cI*$_r$ (LTC_r)		This study
pN18	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*-P_{LtetO1}-*cI*$_r$ (TLC_r)		This study
pN19	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*-P_{LlacO1}-*lacI*$_r$-P_{LlacO1}-*tetR*$_r$ (CL_rT_r)	AS098/AS099, KTp46/AS132, AS133/AS134; BglII	This study
pN20	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*-P_{LlacO1}-*tetR*$_r$-P_{LlacO1}-*lacI*$_r$ (CT_rL_r)		This study
pN21	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*tetR*$_r$ (LC_rT_r)		This study
pN22	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI*$_r$ (LT_rC_r)		This study
pN23	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*lacI*$_r$ (TC_rL_r)		This study
pN24	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI*$_r$ (TL_rC_r)		This study
pN25	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*lacI*$_r$-P_{LlacO1}-*tetR*$_r$ ($C_rL_rT_r$)	AS129/AS130, AS131/AS132, AS133/AS134; BglII	This study
pN26	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*tetR*$_r$-P_{LlacO1}-*lacI*$_r$ ($C_rT_rL_r$)		This study
pN27	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*tetR*$_r$ ($L_rC_rT_r$)		This study
pN28	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*$_r$-P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI*$_r$		This study

	($L_rT_rC_r$)		
pN29	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*lacI*$_r$ ($T_rC_rL_r$)		This study
pN30	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*$_r$-P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI*$_r$ ($T_rL_rC_r$)		This study
pN33	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*tetR* (L_rC_rT)	AS129/AS130, AS131/AS132, AS102/AS103; BglI	This study
pN34	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*$_r$-P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI* (L_rT_rC)		This study
pN35	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*lacI* (T_rC_rL)		This study
pN36	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*$_r$-P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI* (T_rL_rC)		This study
pN37	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR* (C_rLT)	AS129/AS130, AS100/AS101, AS102/AS103; BglI	This study
pN39	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI*-P_{LlacO1}-*tetR* (L_rCT)		This study
pN41	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI*-P_{LlacO1}-*lacI* (T_rCL)		This study
pN43	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR*$_r$ (C_rLT_r)	AS129/AS130, AS100/KTp45, AS133/AS134; BglI	This study
pN44	*pSC101* ori*, Kan^R, P_{LtetO1}-*cI*$_r$-P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*$_r$ (C_rTL_r)		This study
pN45	*pSC101* ori*, Kan^R, P_{LlacO1}-*lacI*$_r$-P_{LtetO1}-*cI*-P_{LlacO1}-*tetR*$_r$ (L_rCT_r)		This study
pN47	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*$_r$-P_{LtetO1}-*cI*-P_{LlacO1}-*lacI*$_r$ (T_rCL_r)		This study
pN49	*pSC101* ori*, Kan^R, P_{LtetO1}-*yfp*-P_{LlacO1}-*lacI*-P_{LlacO1}-*tetR* (VLT)	AS098/AS099, AS100/AS101, AS102/AS103; BglI	This study
pN54	*pSC101* ori*, Kan^R, P_{LlacO1}-*tetR*-P_{LlacO1}-*lacI*-P_{LtetO1}-*yfp* (TLV)		This study
pZS*12-*gfp*	*pSC101* ori*, Amp^R, P_{LlacO1}-*gfp*		Laboratory stock
pZS*2R-*gfp*	*pSC101* ori*, Kan^R, P_R-*gfp*		Laboratory stock
pZS*11-*yfp*	*pSC101* ori*, Amp^R, P_{LtetO1}-*yfp*		Laboratory stock
pZA31-CFP	*P15A ori*, Cam^R, P_{LtetO1}-*cfp*		Laboratory stock
pLA2	*R6K ori*, Kan^R, *attP*$_\lambda$		(Haldimann and Wanner, 2001)

pBS3Clux	*pMB1 ori*, Amp^R, P$_{lac}$-*mCherry*, *luxABCDE*		(Radeck et al., 2013)
pKD3	*R6K ori*, Cam^R-*frt*, Amp^R		(Datsenko and Wanner, 2000)
pKD13	*R6K ori*, Kan^R-*frt*, Amp^R		(Datsenko and Wanner, 2000)
pAH68-frt-cat	*R6K* ori, Cam^R-*frt*, *attP*$_{HK022}$		(Pleška et al., 2016)
pAH120-frt-cat	*R6K* ori, Cam^R-*frt*, *attP*$_{\lambda}$		(Pleška et al., 2016)

[a] Amp – ampicillin resistance, Kan – kanamycin resistance, Cam – chloramphenicol resistance

Table S4. ANOVA test statistics.

IPTG [mM]	F(DFn, DFd)	P value
0	$F_{2,24}$=1.957	P=0.1632
0.1	$F_{2,24}$=14.87	P<0.0001
0.2	$F_{2,24}$=13.08	P=0.0001
0.3	$F_{2,18}$=22.53	P<0.0001
0.4	$F_{2,21}$=7.345	P=0.0038
1	$F_{2,24}$=4.500	P=0.0219

861 ***Figure captions***

862 **Figure 1. Architecture and phenotypes of the gene regulatory network. A.** Diagram of
863 interactions between the three independent TUs encoding for the repressors, *lacI*, *tetR*, and *cI*,
864 their respective inducers IPTG and aTc, and the promoters they control, with *yfp* as the GRN
865 output (left). Phenotype of this GRN as predicted by our mathematical model (right). **B.** Genetic
866 architecture of TU permutations of GRN plasmid (left). Cartoon of TU permutations (right).
867 Abbreviations used throughout the text: C stands for *cI*, L for *lacI*, and T for *tetR.* Letter *r*
868 denotes reverse orientation. **C** and **D.** Fluorescence of cells carrying a representative subset of
869 different TU permutations of the GRN plasmid. A binary output value (On or Off) was assigned
870 to each environment which thus defines a logical operation: NOT (aTc) in panel **C** and NOR in
871 panel **D**. Graphs show means and error bars standard deviations for three independent
872 biological replicates.

873

874 **Figure 2. Changes in relative TU order lead to qualitative changes of phenotype.** Fluorescence
875 of cells carrying six different TU permutations of the GRN on a plasmid (**A**), three GRN variants
876 integrated on the chromosome at the phage HK022 attachment site (**B**), and with each of the
877 repressor genes integrated at separate chromosomal loci (**C**). Graphs show population level
878 fluorescence measurements of strains exposed to: no inducer, aTc, IPTG, or a combination of
879 aTc and IPTG (as indicated). Graph shows mean and standard deviations for three independent
880 biological replicates. Flow cytometry histograms of cell fluorescence show 10,000 gated events,
881 corresponding to YFP expressed in a given strain grown without (solid line) and with IPTG

(dotted line). For each strain and condition, three biological replicates are shown. Relative TU order of the three repressors is shown under the respective graphs.

Figure 3. Differences in *cI* expression lead to TU order dependent phenotypes. A. RT-qPCR analysis of *cI* expression. RT-qPCR was performed, using *cI*-specific primers. The induction ratios were calculated relative to the uninduced strain CLT. **B.** YFP levels measured in strains VLT and TLV which carry a *yfp* reporter (V) under control of P_{tet} and differ only in relative TU order. **C.** Heatmaps show P_R promoter activity in six strains carrying plasmids differing in relative TU order (left) and activity of P_{tet} present *in trans* in the same strains on a second plasmid (right). **D.** Relative TU order effects depend on expression from the *lac* promoter in IPTG-dependent manner. Strains CLT and TLC were grown in presence of different concentrations of IPTG. **B - D.** Strains were exposed to aTc, IPTG, or a combination of aTc and IPTG (as indicated). For reasons of clarity, in **C.** the highest expression level for each strain was individually normalized. Graphs show means and error bars standard deviations for three independent biological replicates.

Figure 4. Phenotype of strain TLC can be explained by transcriptional read-through. A. Genetic architecture of plasmid fragments encoding three repressors in strains TLC and TLC_{-10} (carrying mutations in the -10 promoter element of P_{tet}). Promoters are marked as bent arrows, terminators are represented by vertical bars and a circle, monocistronic transcripts are represented by solid arrows, predicted read-through transcripts by dashed arrows (left). Interaction diagrams within the two GRNs. Solid lines represent interactions between transcription factors and the promoters they control, dashed line represents effects resulting

from local genetic context (right). **B.** Fluorescence of cells carrying TLC plasmid with either P_{tet} or P_{tet} carrying a mutation in the -10 promoter element (TLC_{-10}) grown in presence or absence of aTc and IPTG. We expected that if GRN behavior is due to transcriptional read-through, mutating the P_{tet} promoter should not affect the responsiveness to IPTG. If, on the other hand, expression of *cI* was driven only by P_{tet}, mutating the -10 promoter element should lead to an ALL ON phenotype. Lack of repression in strains with P_{tet-10} variant after aTc induction further confirms that this promoter variant is inactive. Graph shows means and error bars standard deviations for three independent biological replicates.

Figure 5. Change of terminator leads to qualitative change in phenotype. A. Genetic architecture of plasmid fragments encoding three repressors in strains TLC with different terminators: T1, T1T2, T*crp* or T*tonB*, preceding *cI*, and fluorescence of cells carrying these plasmids grown in presence or absence of aTc and IPTG. **B.** Genetic architecture of plasmid fragments encoding three repressors in strains LCT with either T1 or a double T*crp*-T1 terminator, and fluorescence of cells carrying these plasmids grown in presence or absence of aTc and IPTG. **A.** and **B.** Promoters are marked as bent arrows, terminators are represented by vertical bars and a circle, operators as rectangles, monocistronic transcripts are represented by solid arrows, predicted read-through transcripts by dashed arrows. Increasing line thickness corresponds to increasing amount of transcript. Graph shows means and error bars standard deviations for three independent biological replicates.

Figure 6. Effect of genetic context of *lacI* repressor chromosomal position on P_{lac} activity. A. Diagram of interactions between LacI repressor and the promoter it represses, as well as genetic architecture of the DNA fragments integrated into MG1655 Δ*lacI* Δ*lacZYA* strain. Promoters are marked as bent arrows, terminators are represented by vertical bars and a circle, operators as rectangles. **B.** P_{lac} activity in cells carrying *lacI* inserted at different loci after a weak (*yeaH*), medium (*flhC*) and strong (*asnT*) terminator grown in different concentrations of IPTG. *yfp* levels were measured in exponentially growing cells. Graph shows means for at least three independent biological replicates. **C.** RT-qPCR quantification of *lacI* transcript in strains described above. RT-qPCR was performed, using *lacI*-specific primers. The induction ratios were calculated relative to the strain with insertion in *yeaH* locus. **D.** Ethidium bromide-stained 1% agarose gel of PCR products obtained with primers spanning the intergenic region between the upstream gene (*yeaH*, *flhC* or *asnT*) and *lacI*. Templates for PCR were: chromosomal DNA (gDNA), cDNA as used for RT-qPCR above, and RNA.

Figure 7. Architecture of a regulon depends on local genetic context. A. Diagram of interactions within our GRN in two different TU arrangements: CLT and TLC. **B.** Regulatory patterns in two regulons with overlapping components. Genes $Gene_1$ and $Gene_2$ are regulated via transcription factor-operator interactions with regulator TF_1. Gene $Gene_3$ is regulated by TF_2 binding and by TF_1 via local genetic context effects. **A. and B.** Solid lines represent interactions between transcription factors and the promoters they control, dashed line represents effects resulting from local genetic context. Promoters are marked as bent arrows, terminators are represented by vertical bars and a circle.

947 **Figure S1. A model of the mechanistic basis of gene expression, including transcriptional read-**
948 **through. A.** The topology of network interactions used in the model, along with the equations
949 that model them. Term in red represent transcriptional read-through. **B.** Two examples of
950 transcriptional read-through. In the first example (top), transcription of *lacI* leads to
951 transcriptional read-through into *tetR* and *cI*. This is modeled by additive terms for the rate of
952 transcription of both *tetR* and *cI*. In the second example (bottom), expression of *cI* leads to
953 transcriptional read-through into *tetR*. Because *lacI* is on the opposite strand from *tetR*,
954 expression of *tetR* is assumed not to influence expression of *lacI*. **C.** Range of values that fit the
955 experimental results. Contour plot for a parameter scan for K in [0,0.5] and μ in [0,1], both with
956 a step size of 0.05. For K in [0,0.21] and μ in [0.4,1], both parameters were scanned with a step
957 size of 0.01. Lighter colors means higher number of networks correctly fitted.
958
959 **Figure S2. GRNs in which phenotype is dependent only on changes in relative TU order and**
960 **orientation.** Fluorescence of cells carrying 26 different TU permutations of the GRN plasmid,
961 grouped in pairs differing only in orientation with respect to plasmid backbone. Graphs show
962 means and error bars standard deviations for three independent biological replicates. Shading
963 marks strains which behave as predicted by the mathematical model incorporating
964 transcriptional read-through (grey, **A.**), white boxes denote strains which do not (**B.**).
965
966 **Figure S3. GRNs in which the influence of plasmid-encoded genetic elements cannot be ruled**
967 **out.** Fluorescence of cells carrying 11 different TU permutations of the GRN plasmid, grouped in
968 pairs differing only in orientation with respect to plasmid backbone. Shading marks strains

969 which behave as predicted by the mathematical model incorporating transcriptional read-
970 through (grey), white boxes denote strains which do not. **A.** GRNs in which pairs do not show
971 consistent phenotype, indicating influence of plasmid-encoded elements. **B.** Pairs in which only
972 one GRN was cloned, making comparison impossible. Graphs show means and error bars
973 standard deviations for three independent biological replicates.
974
975 **Figure S4. Phenotype of strains LCT, $T_rC_rL_r$, TCL, $L_rC_rT_r$, TLC and $C_rL_rT_r$ can be explained by**
976 **transcriptional read-through.** Genetic architecture of plasmid fragments encoding three
977 repressors in strains LCT and $T_rC_rL_r$ (**A.**), TCL and $L_rC_rT_r$ (**B.**), TLC and $C_rL_rT_r$ (**C.**) and their
978 derivatives carrying mutations in the -10 promoter element of P*tet*, together with fluorescence
979 of cells carrying these plasmids grown in presence or absence of aTc and IPTG. We expected
980 that if GRN behavior is due to transcriptional read-through, mutating the P_{tet} promoter should
981 not affect the responsiveness to IPTG. If, on the other hand, expression of *cI* was driven only by
982 P_{tet}, mutating the -10 promoter element should lead to an ALL ON phenotype. Lack of
983 repression in strains with $P_{tet\text{-}10}$ variant after aTc induction further confirms that this promoter
984 variant is inactive. Monocistronic transcripts are represented by solid arrows, predicted read-
985 through transcripts by dashed arrows. Graph shows means and error bars standard deviations
986 for three independent biological replicates.
987
988 **Figure S5. Point mutations in -10 promoter element render P_{tet} inactive.** Fluorescence of cells
989 carrying either wild type P_{tet} or P_{tet} with two point mutations in the -10 element driving *yfp*

990 expression. Strains were grown without inducers. Graphs show means and error bars standard
991 deviations for three independent biological replicates.
992
993 **Figure S6. Northern blot assay shows read-through transcript in strains carrying T*crp* and**
994 **T*tonB***. **A.** Schematic representation of the plasmid fragments encoding three repressors in
995 strains TLC, TL$^{T\mathit{crp}}$C and TL$^{T\mathit{tonB}}$C. Shading marks regions which differ between the three GRNs. **B.**
996 Northern blot analysis of *cI, lacI,* and *kanR* expression. Seven micrograms of total RNA (−, no
997 induction; +, with IPTG induction) was loaded on a formaldehyde gel, and Northern blot
998 analysis was performed as described in Materials and Methods. NC – negative control, strain
999 carrying network plasmid not encoding any transcription factor genes. Ethidium bromide
1000 staining of 23S and 16S rRNA is shown as a loading control.
1001
1002 **Figure S7. Terminator containing DNA fragments between repressor genes show no cryptic**
1003 **promoter activity.** Fluorescence of cells carrying DNA fragments located on GRN plasmids
1004 between *lacI* and *cI*, and encoding T1, T*crp* and T*tonB*, cloned in front of a promoterless *yfp*
1005 gene, and compared to P*tet* activity. Graphs show means and error bars standard deviations
1006 for three independent biological replicates..
1007
1008 **Figure S8. Influence of plasmid genetic elements on gene expression levels. A.** Fluorescence of
1009 cells carrying a promoterless *yfp* gene cloned adjacent to the origin of replication, *ori*. **B.**
1010 Fluorescence of cells carrying P_{tet}-*yfp* cloned in two orientations relative to *ori* and *kanR*
1011 measured in a strain carrying chromosomally encoded *tetR*. **A - B.** Strains were exposed to IPTG

or grown without inducers. Genetic architecture of relevant plasmid fragment is given above the graph. Promoters are marked as bent arrows, terminators are represented by vertical bars and a circle. Graphs show means and error bars standard deviations for three independent biological replicates.

Figure S9. P_{lac} activity differs significantly in cells carrying *lacI* inserted after a weak (*yeaH*), medium (*flhC*) and strong (*asnT*) terminator grown in different concentrations of IPTG. Bars are mean values, circles individual measurements. Error bars are standard deviations. Stars indicate strains that significantly differ from each other (* P-value <0.05, ** <0.005, *** <0.0005, **** <0.0001).

Figure S10. Distribution of phenotypes depending on the threshold applied to define ON and OFF states. A. Definition of the logic operations. Presence or absence of inducer is indicated with + and -, and output by On and Off. **B.** Histogram shows the fraction of networks that can be assigned to a given phenotype depending on threshold. The threshold we used to assign phenotypes is marked with a dashed line.

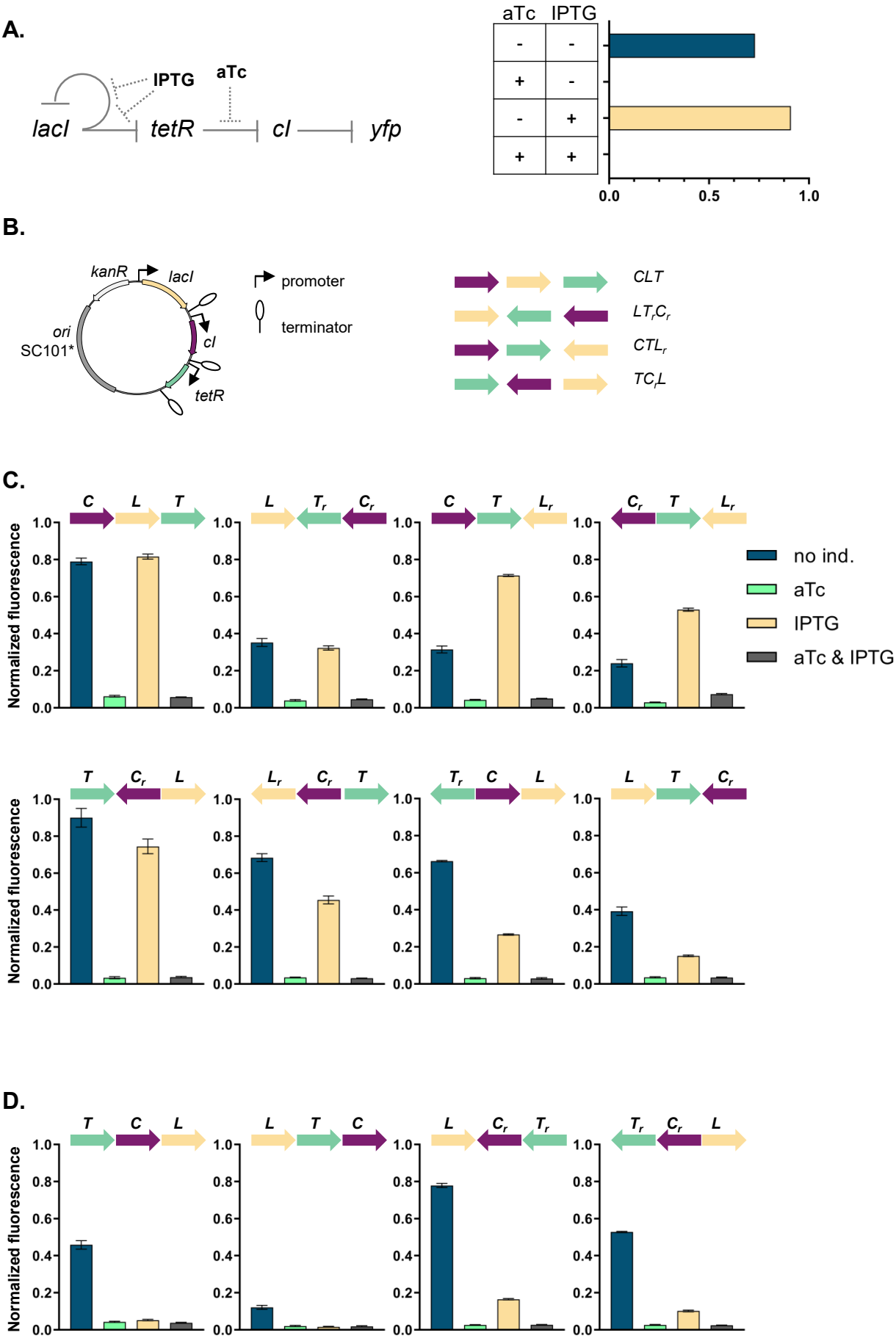
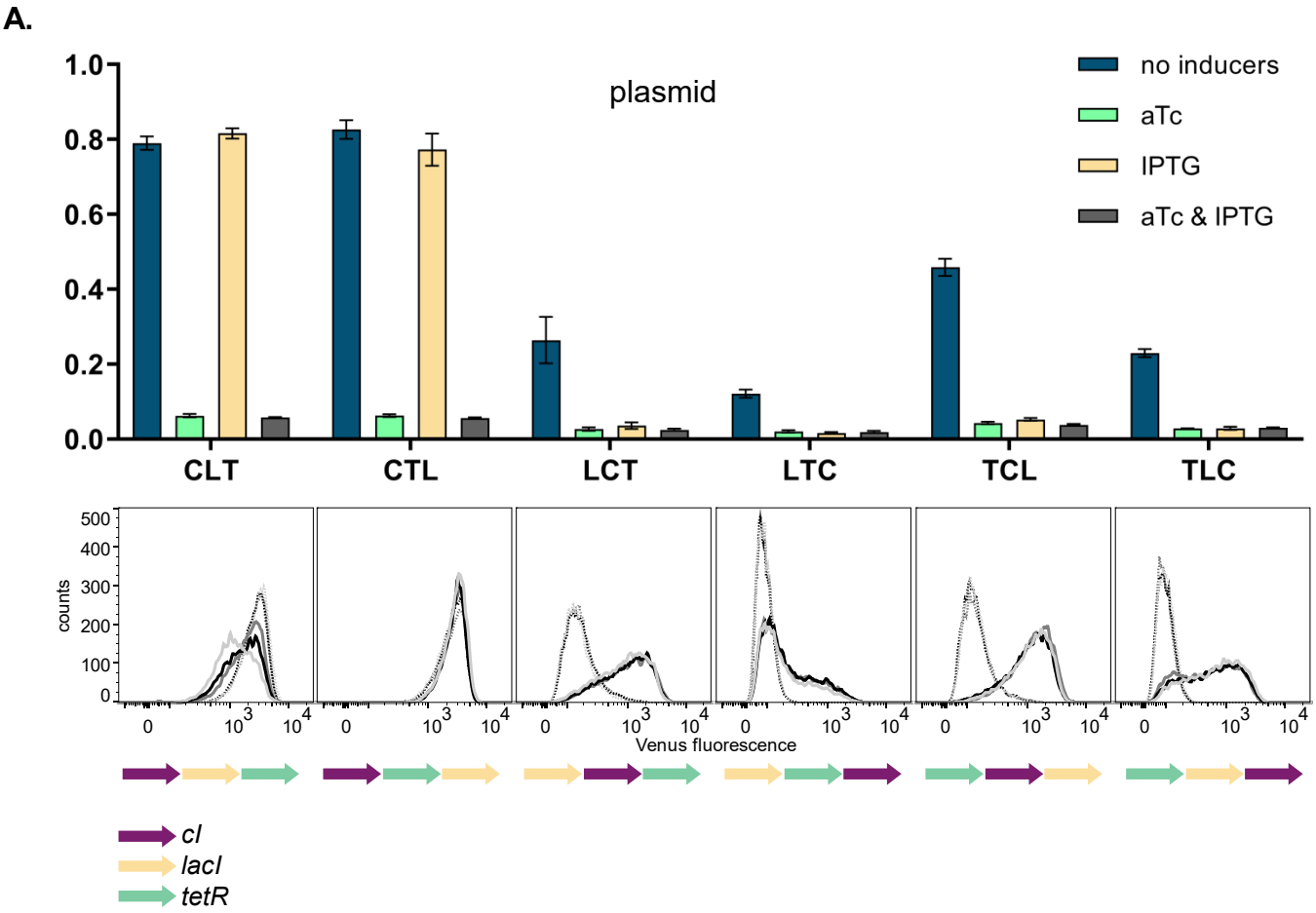
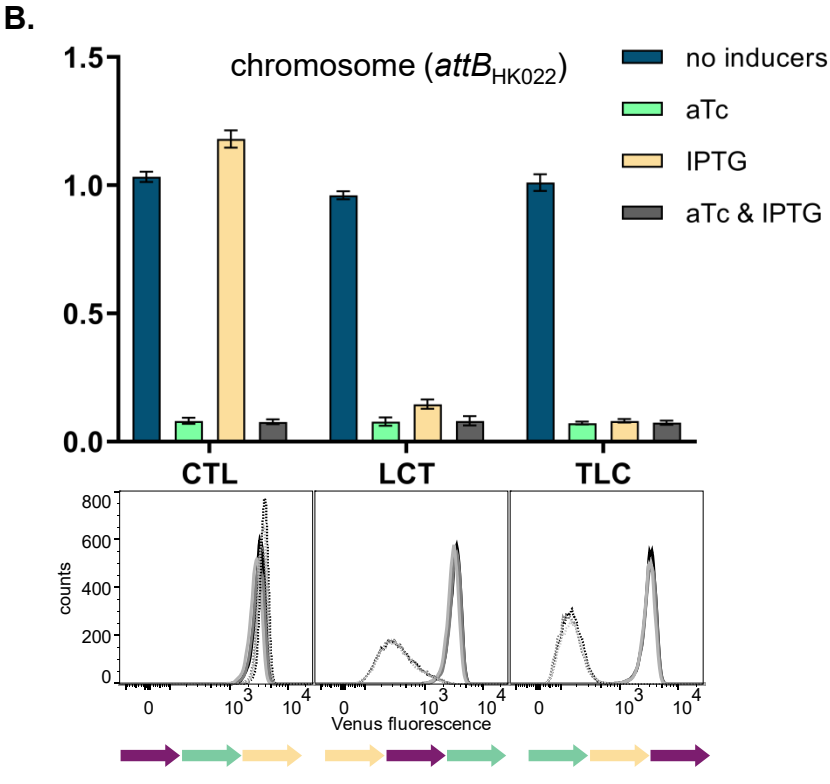


Figure 1





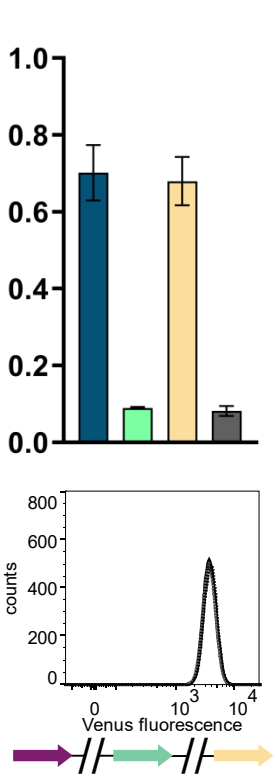


Figure 2

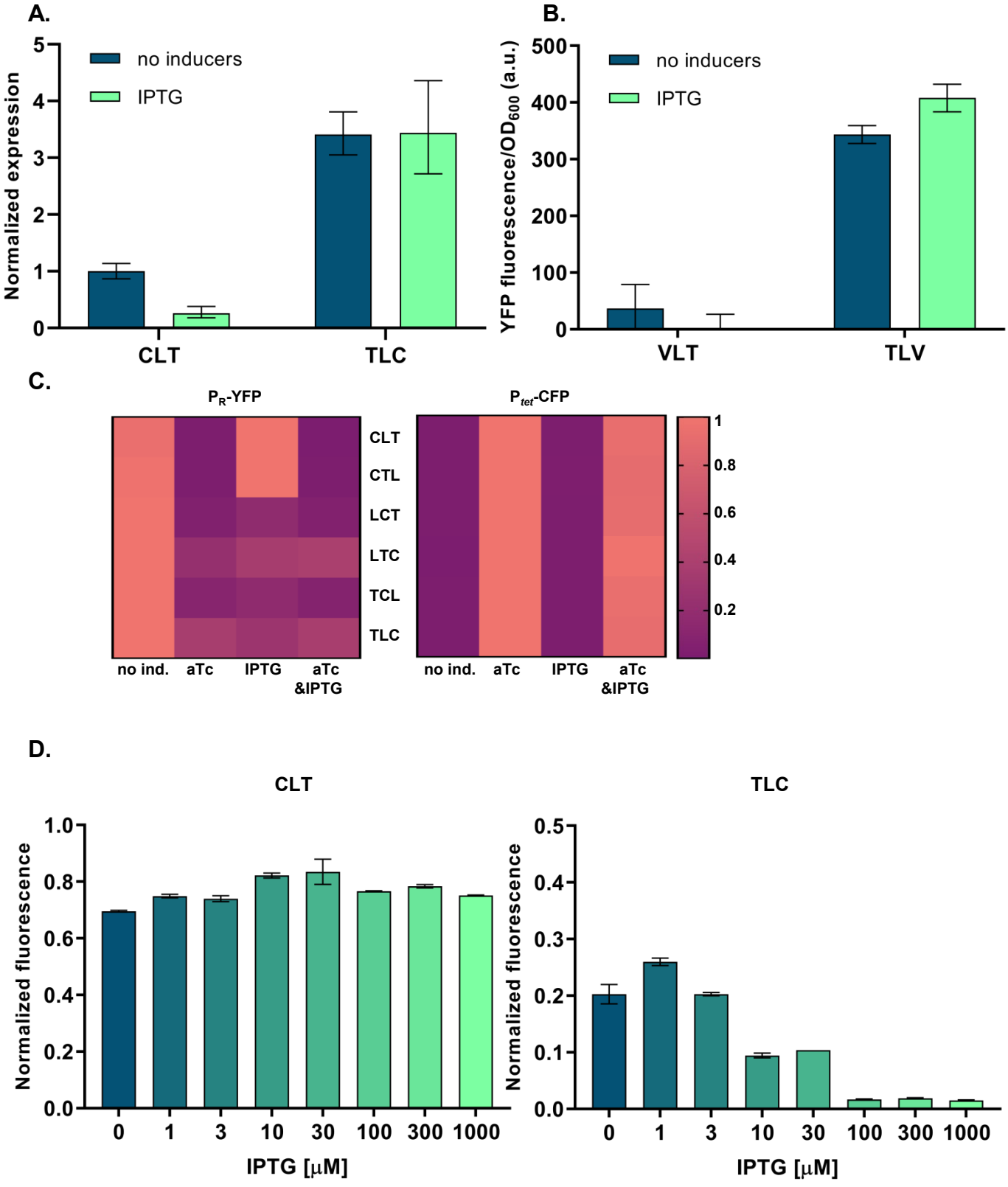


Figure 3

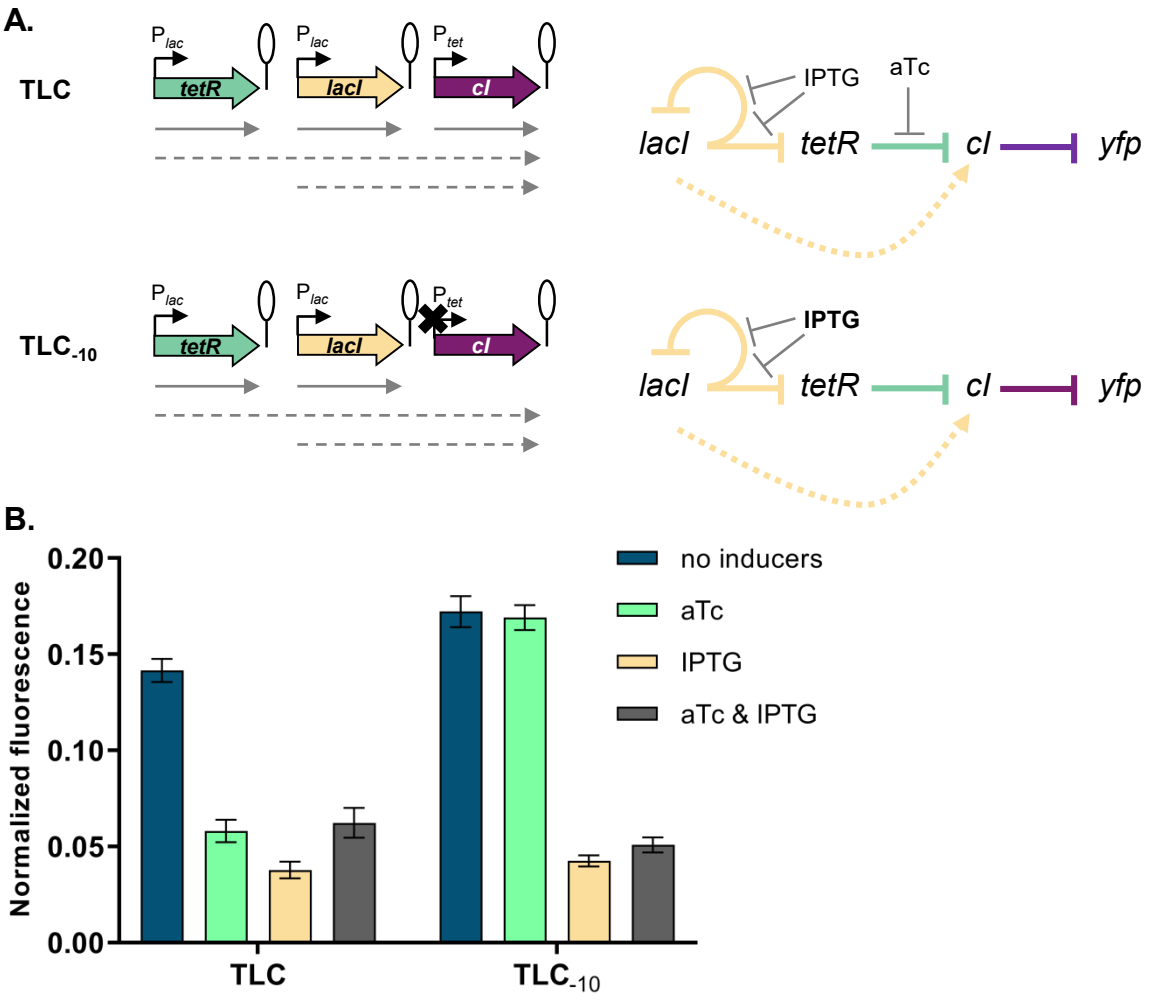


Figure 4

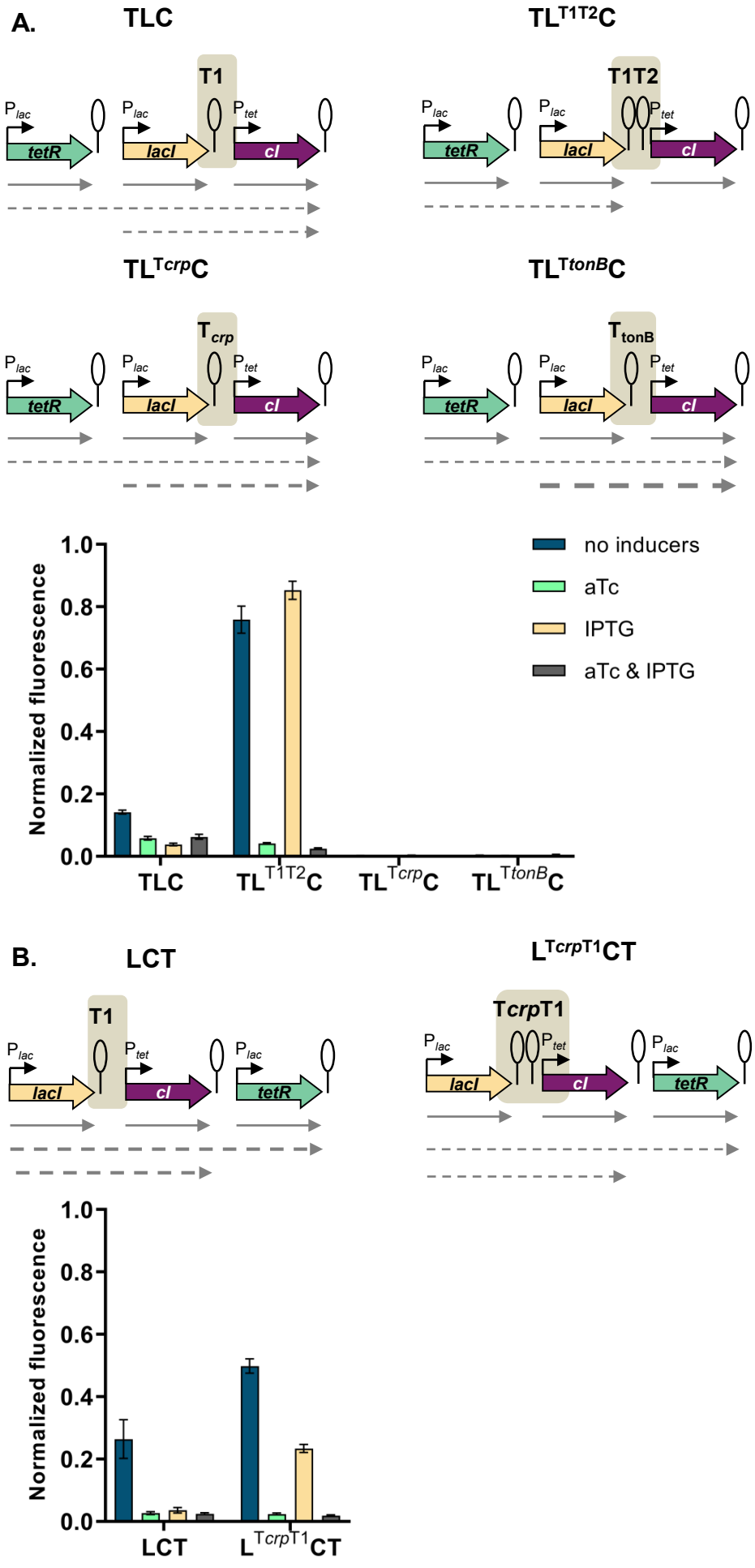
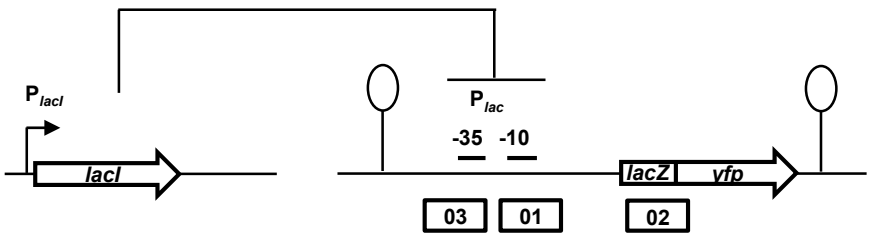


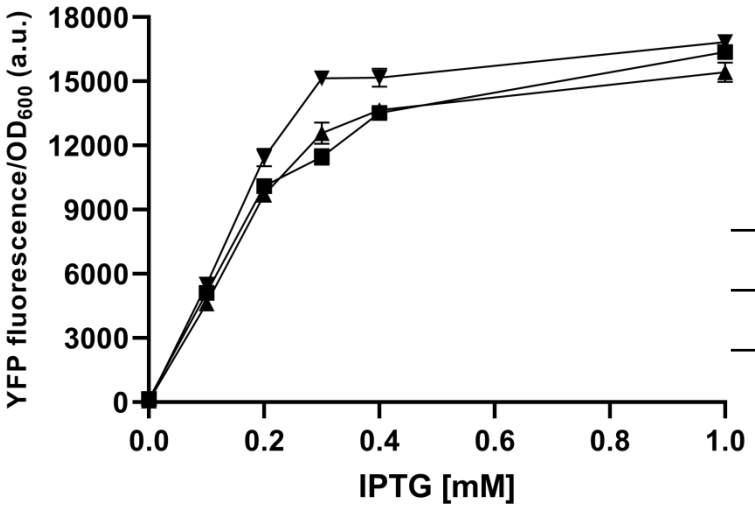
Figure 5

A.



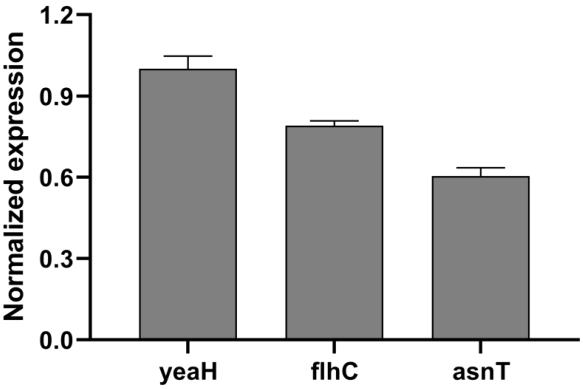
Chromosomal position varied **att$_\lambda$**

B.

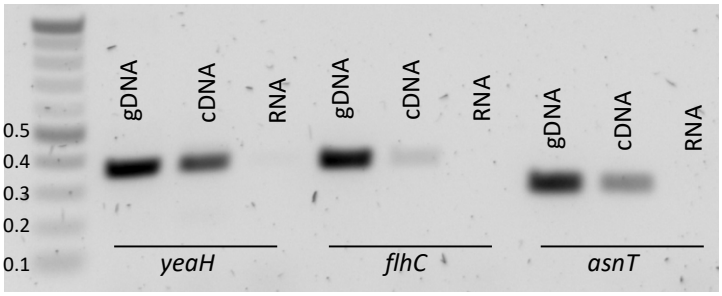


locus	[EC50]	*h*
yeaH	0.1805	1.458
flhC	0.1578	1.999
asnT	0.1412	2.240

C.



D.



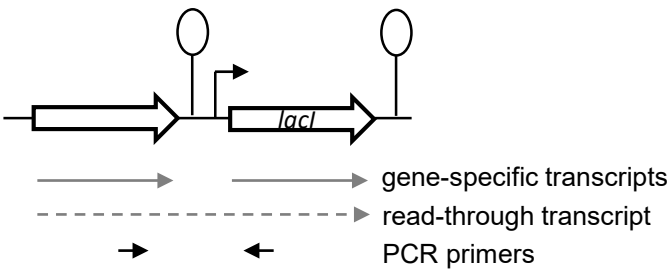


Figure 6

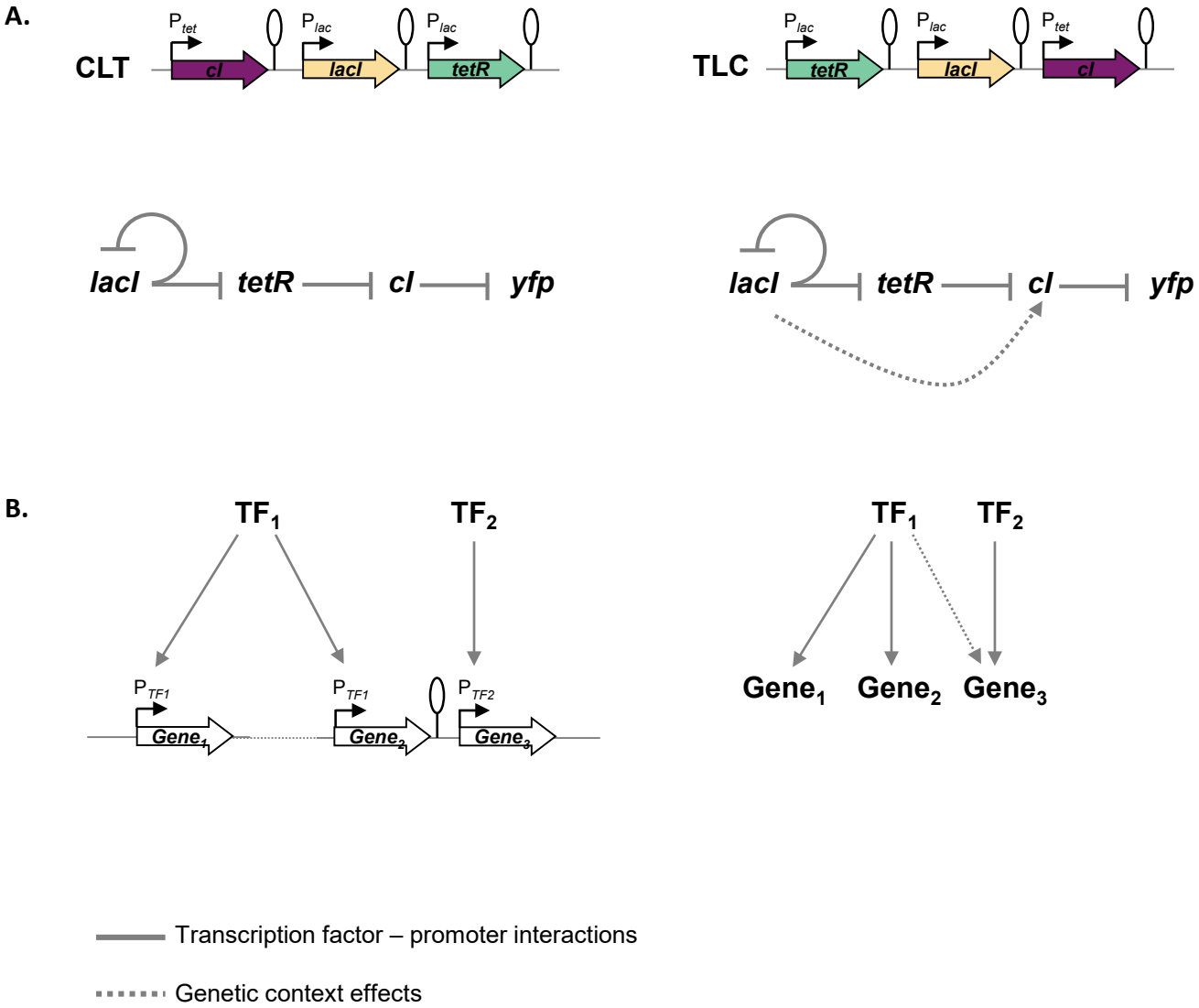
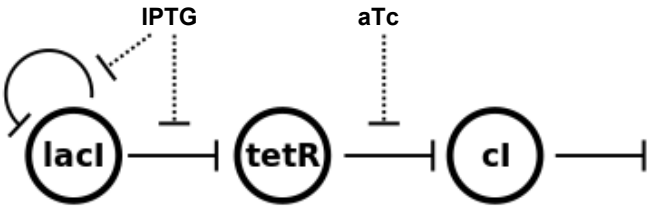


Figure 7

A. **Network interactions**

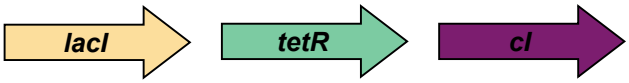


$$\frac{dL}{dt} = \underbrace{k_L - (1-I_1)\frac{L}{K_L+L}}_{L'} - \delta L + r_\chi^L$$

$$\frac{dT}{dt} = \underbrace{k_L - (1-I_1)\frac{L}{K_L+L}}_{T'} - \delta T + r_\chi^T$$

$$\frac{dC}{dt} = \underbrace{k_C - (1-I_2)\frac{T}{K_T+T}}_{C'} - \delta C + r_\chi^C$$

B. **Transcriptional readthrough**



$$r^L_{\vec{L}\vec{T}\vec{C}} = 0$$
$$r^T_{\vec{L}\vec{T}\vec{C}} = \mu L'$$
$$r^C_{\vec{L}\vec{T}\vec{C}} = \mu T'$$

$$r^L_{\vec{L}\vec{T}\vec{C}} = 0$$
$$r^T_{\vec{L}\vec{T}\vec{C}} = \mu C'$$
$$r^C_{\vec{L}\vec{T}\vec{C}} = 0$$

C.

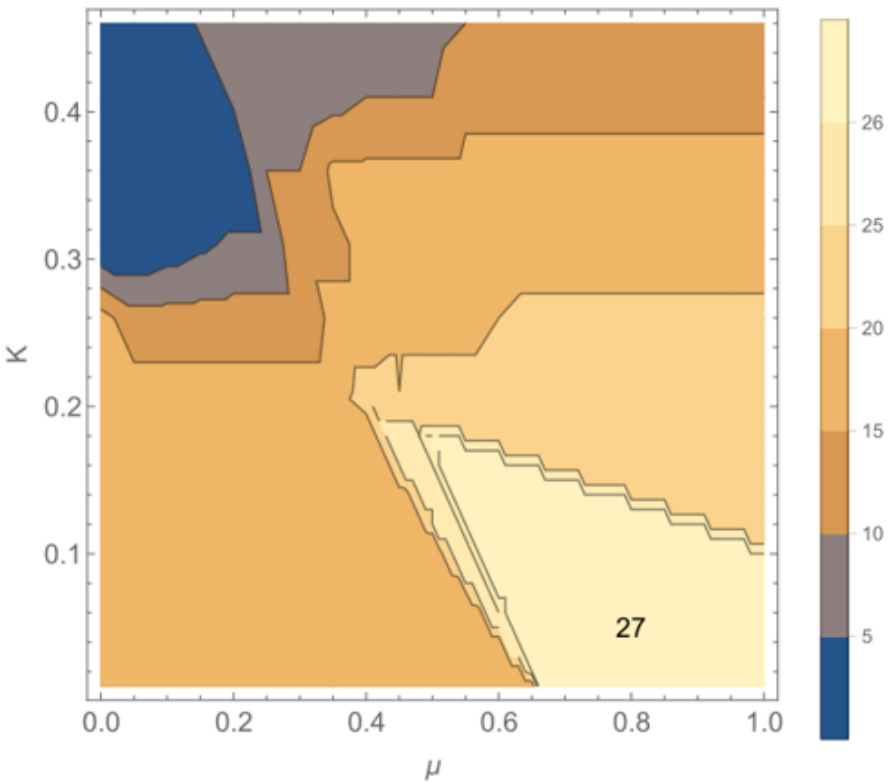
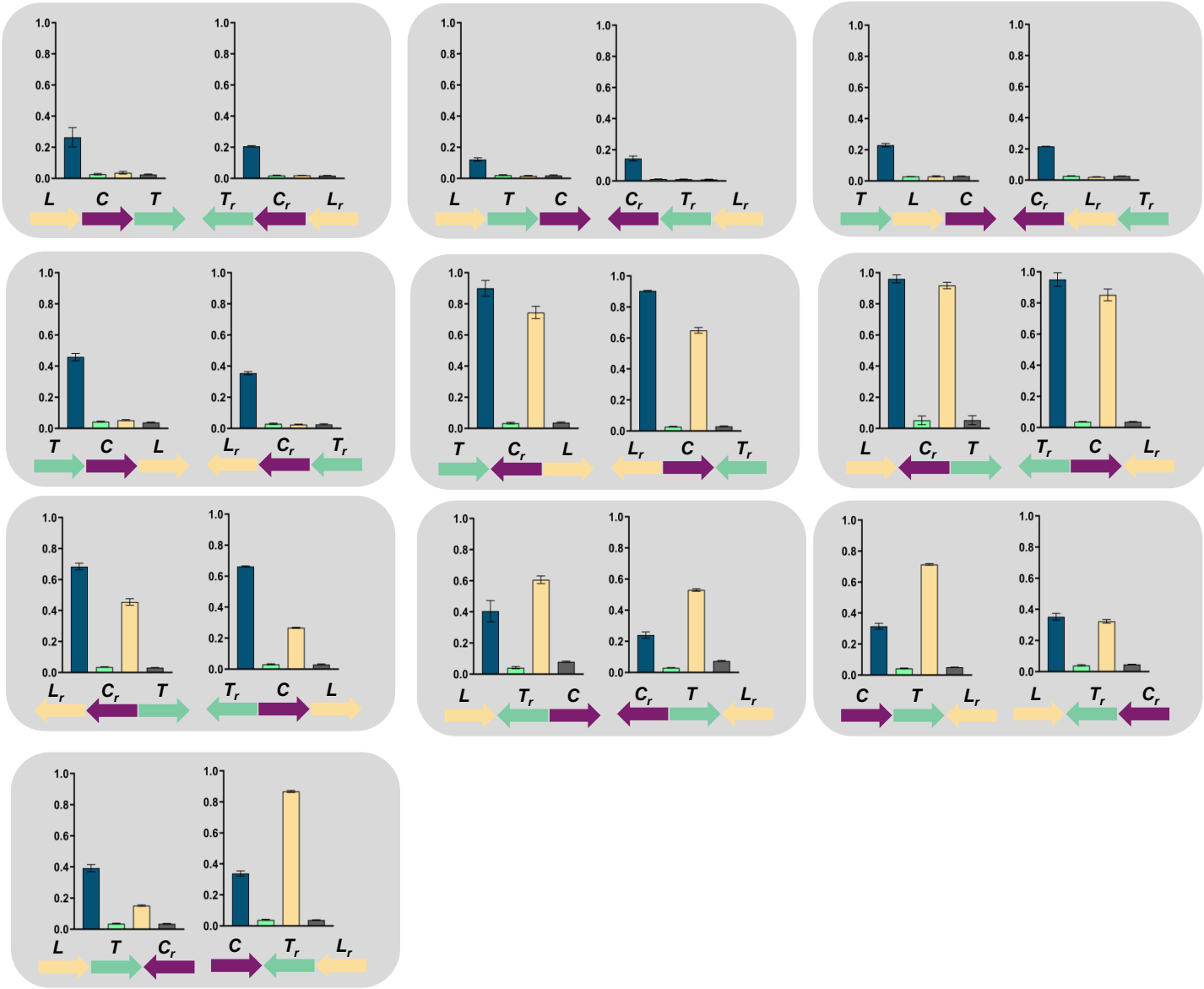


Figure S1

A. Phenotype independent of plasmid elements and predicted by the model



B. Phenotype independent of plasmid elements and not consistent with model predictions

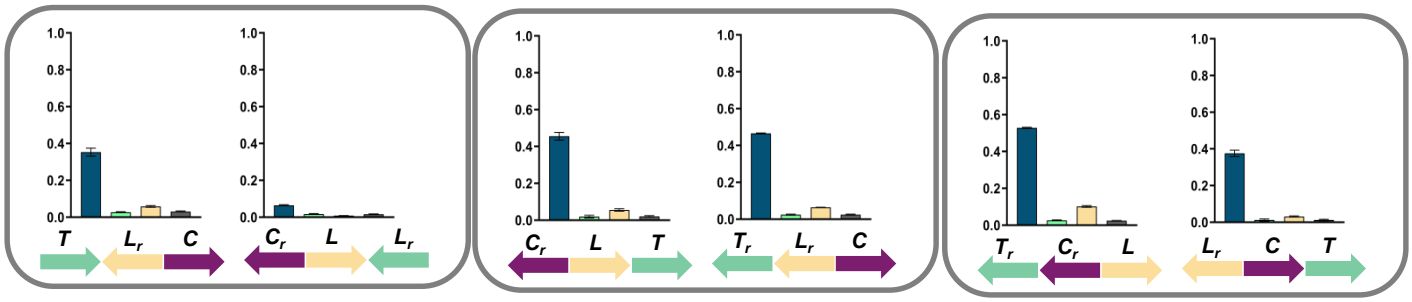
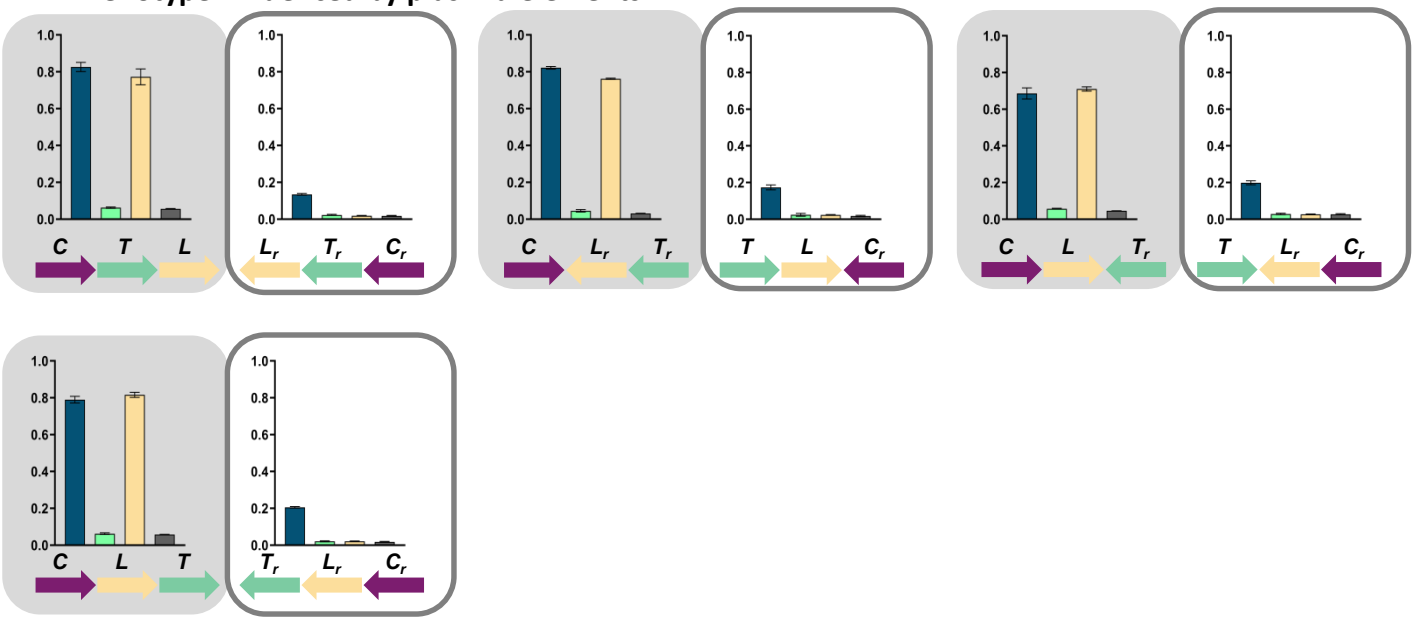


Figure S2

A. Phenotype influenced by plasmid elements



B. Respective pair not cloned

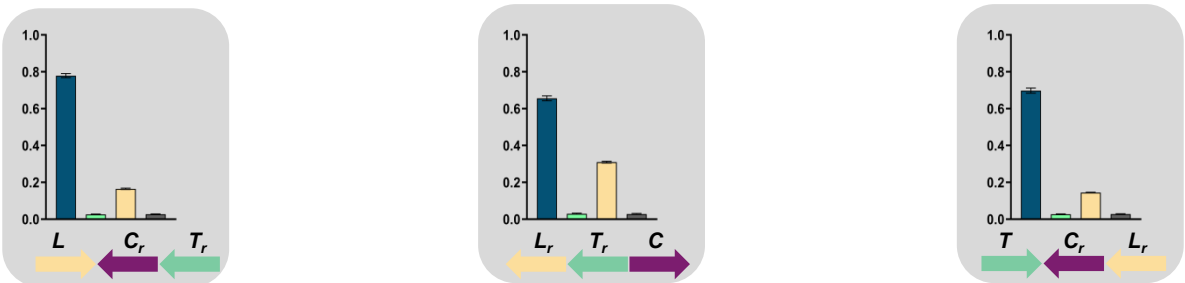


Figure S3

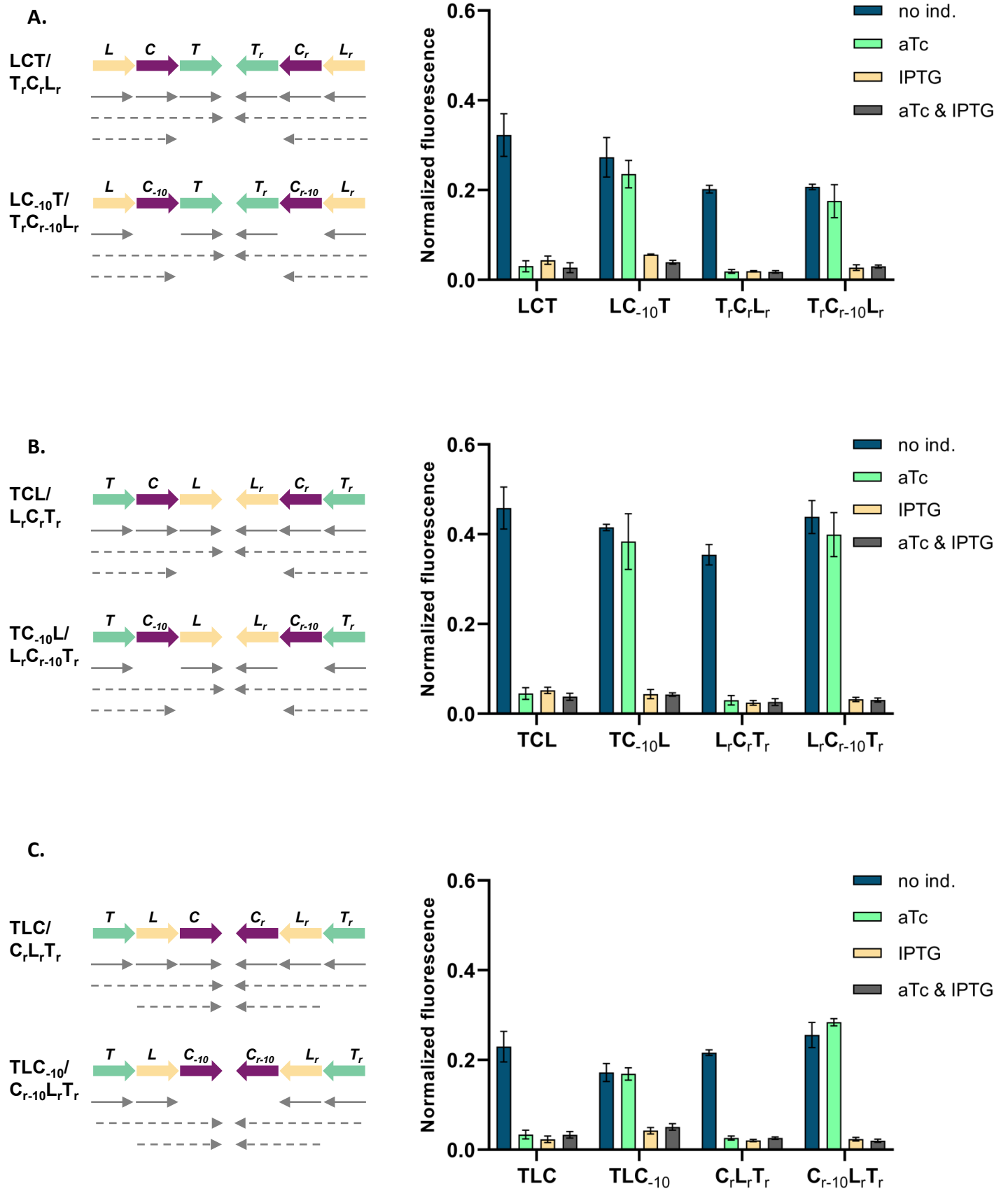


Figure S4

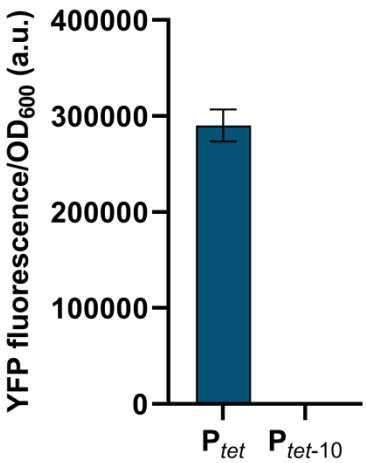
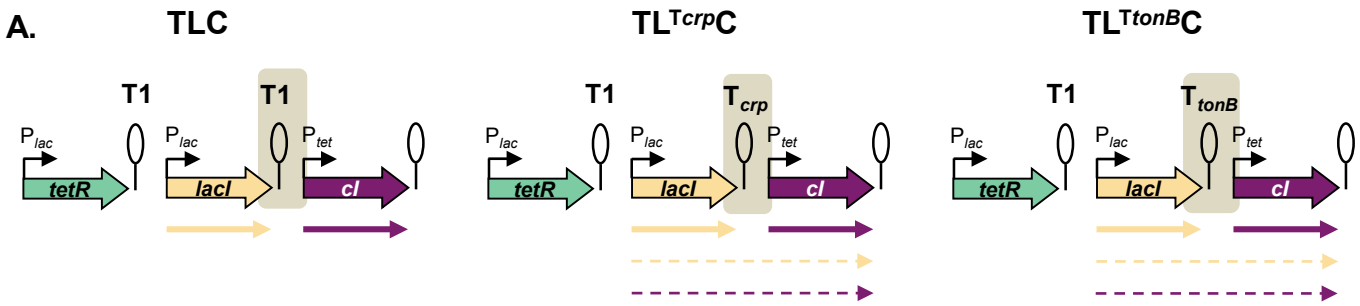
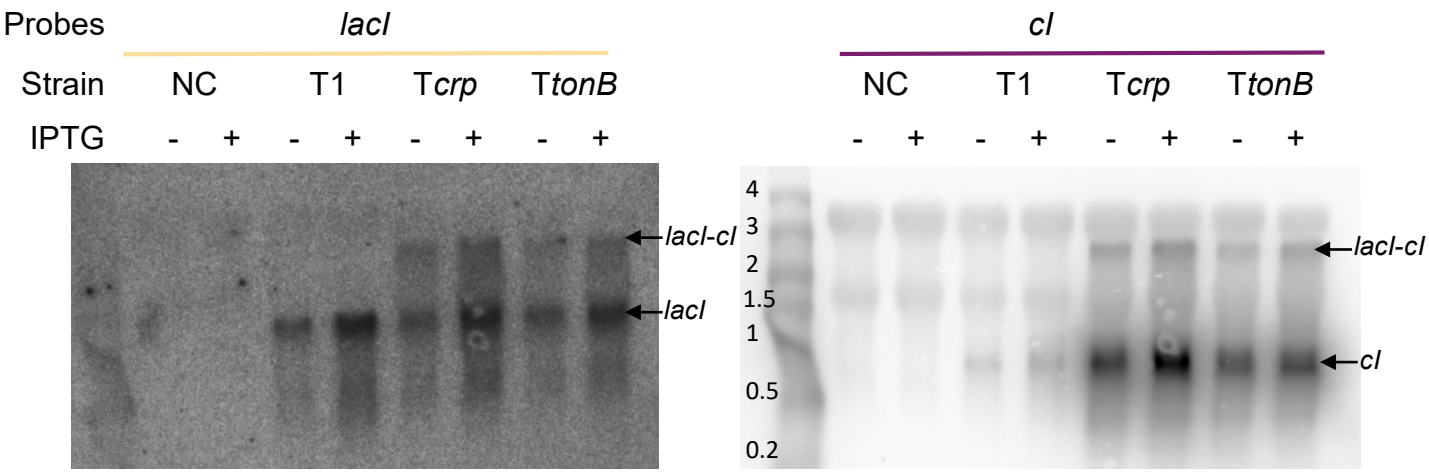


Figure S5



B.



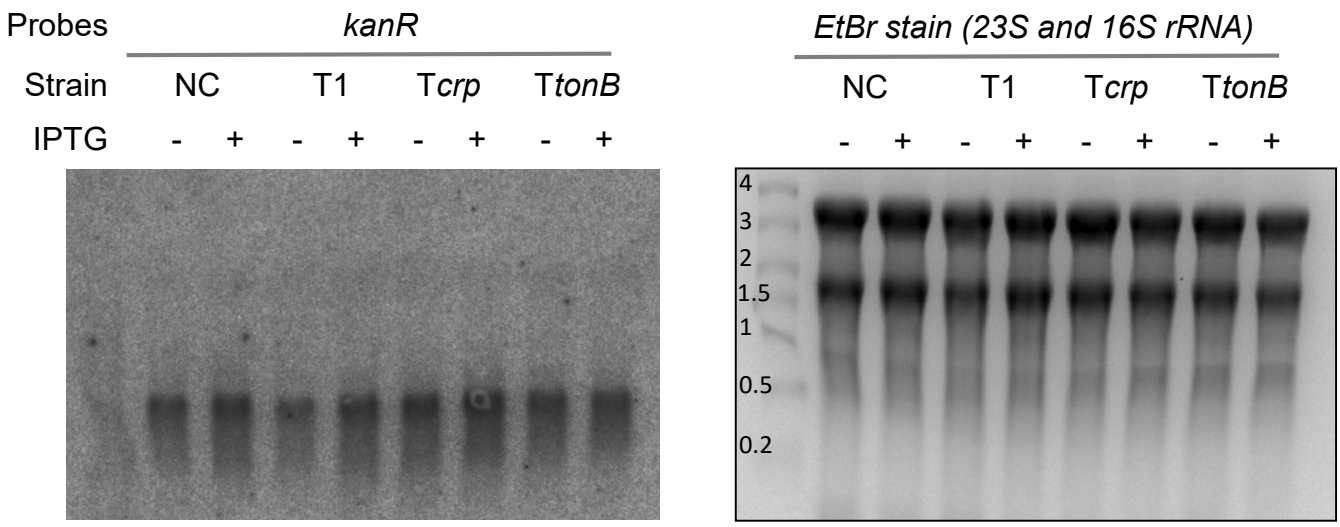


Figure S6

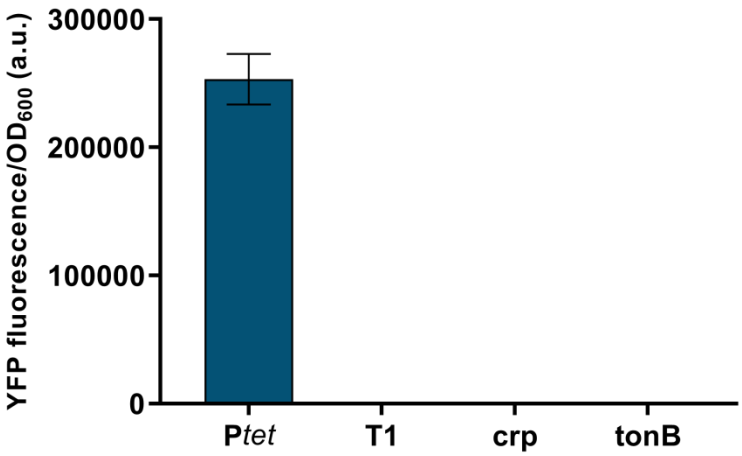


Figure S7

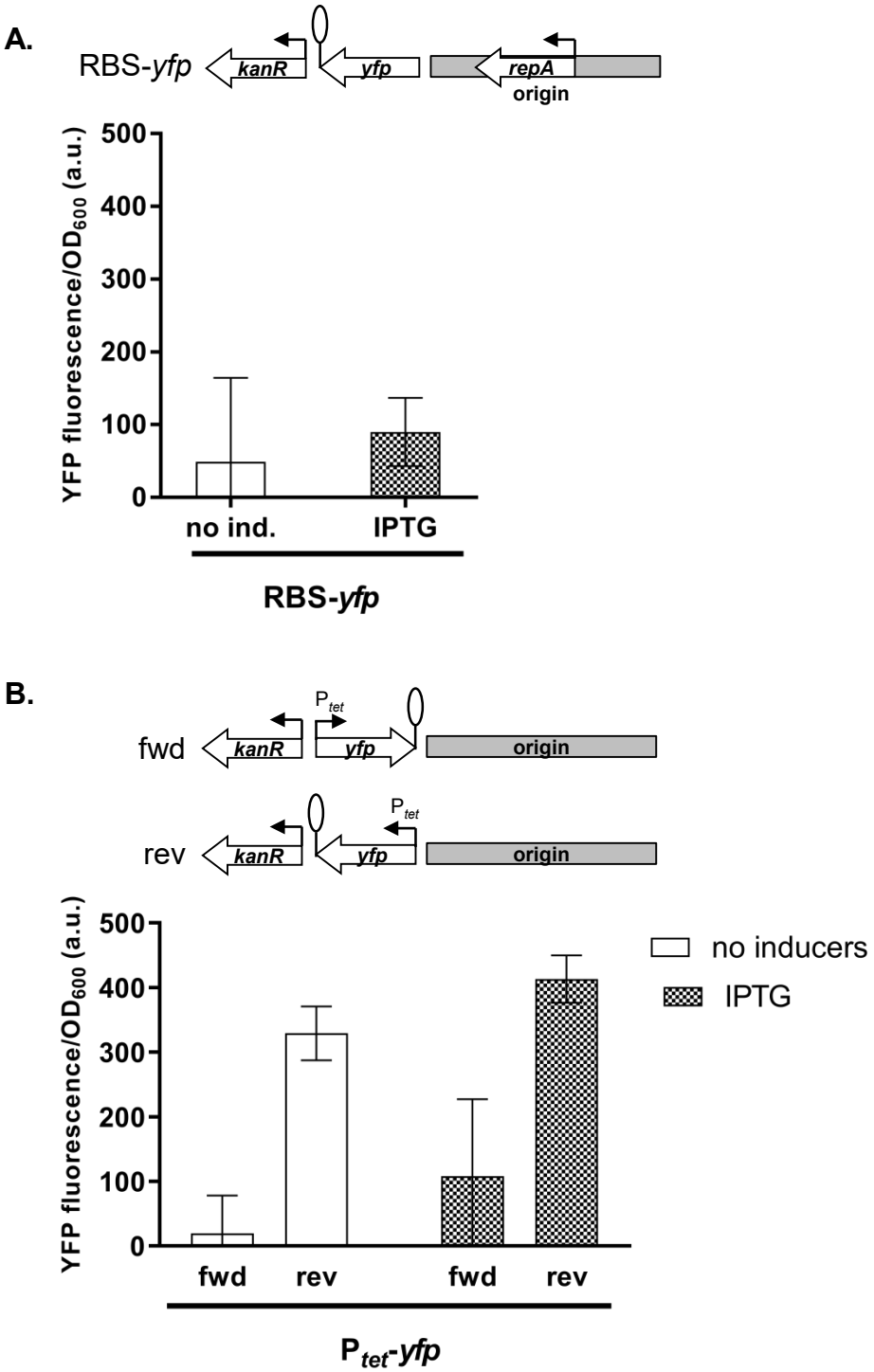


Figure S8

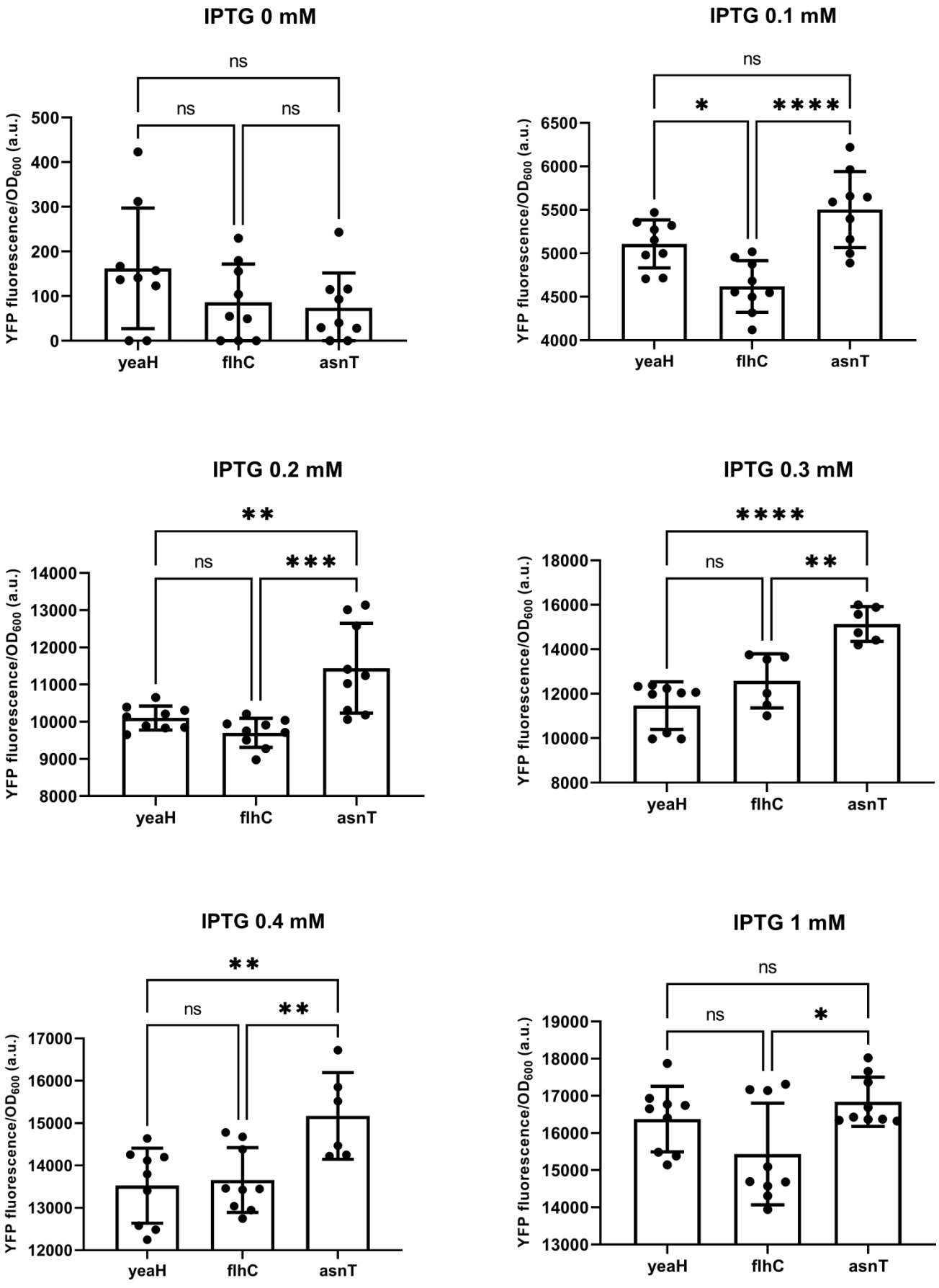


Figure S9

A.

IPTG **aTc**	**-** **-**	**-** **+**	**+** **-**	**+** **+**
NOT (aTc)	On	Off	On	Off
NOR	On	Off	Off	Off
ANDN	Off Off	Off On	On Off	Off Off
ALL ON	On	On	On	On
NAND	On	On	On	Off
ORN	On On	On Off	Off On	On On

B.

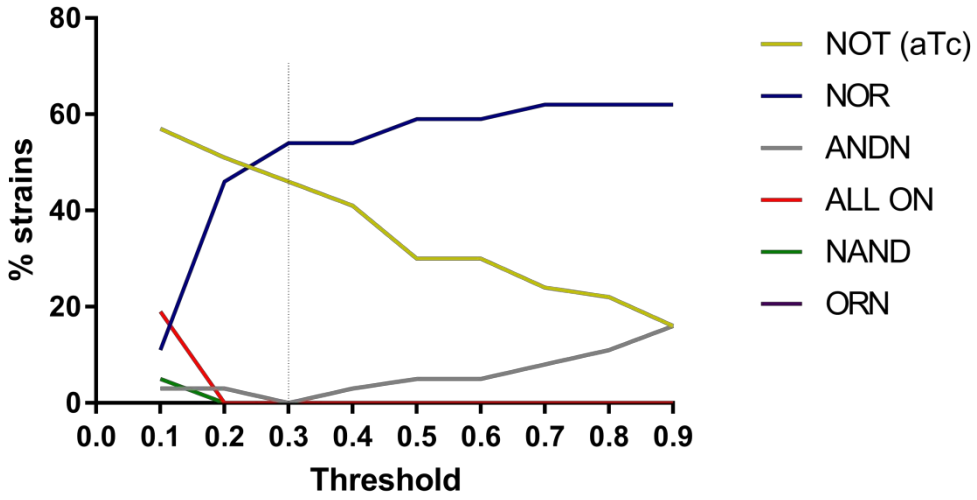


Figure S10