**Extended data, figures**

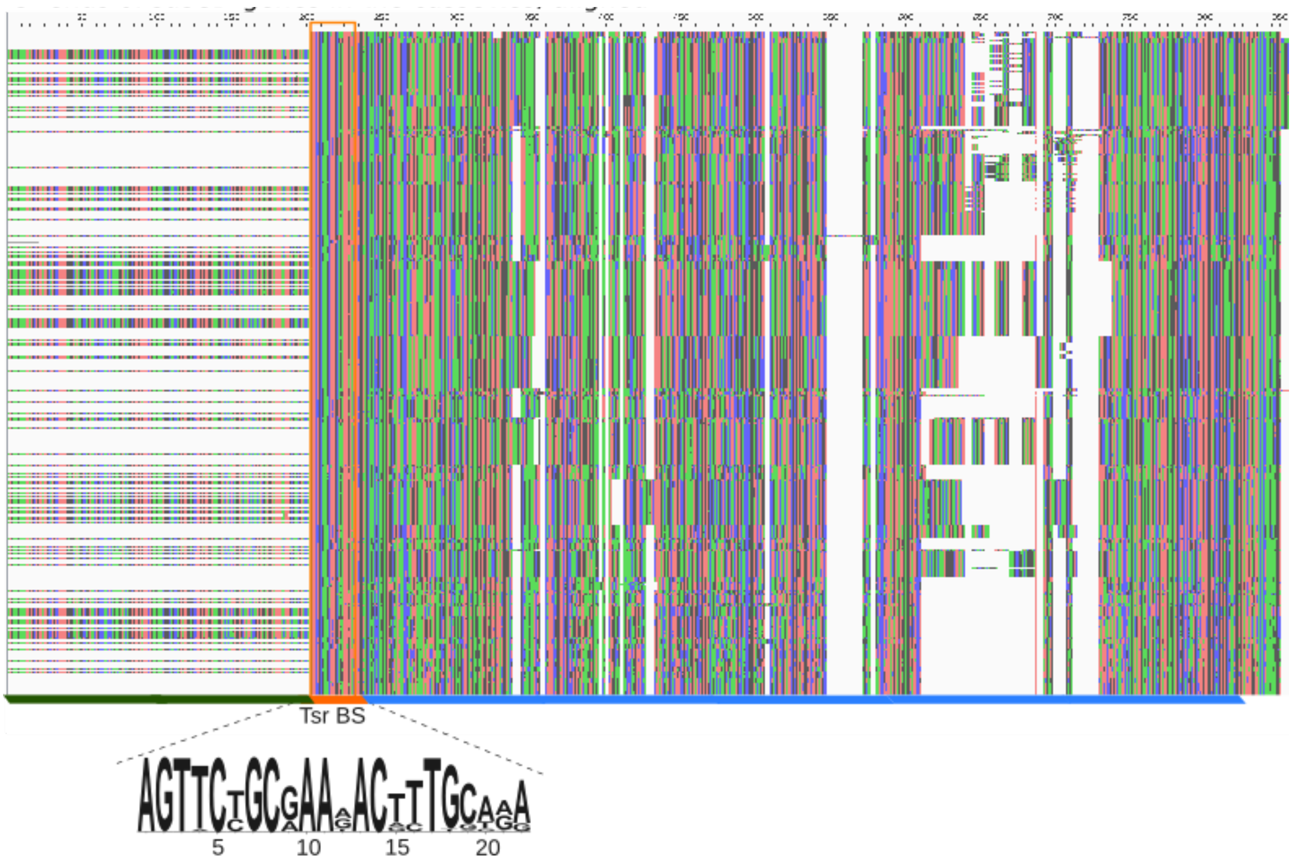


**Figure S1. An illustrative example of the alignment of the *susC* proteins encoded in the *tsr-susCD* shufflons.** A Tsr BS is located at 5' end of each *susC* gene. In each analyzed genome, only one *susCD* transporter gene in the shufflon contains a full-length gene sequence; the other *susCD* transporter genes are shorter, lacking the 5' end sequence. The left side of the alignment shows the presence or absence of the universal N-terminal extension which is characteristic of all functional *susC* proteins. Colors represent different amino acids. The full list of genes and their sequences is given in the Supplementary Table 1.

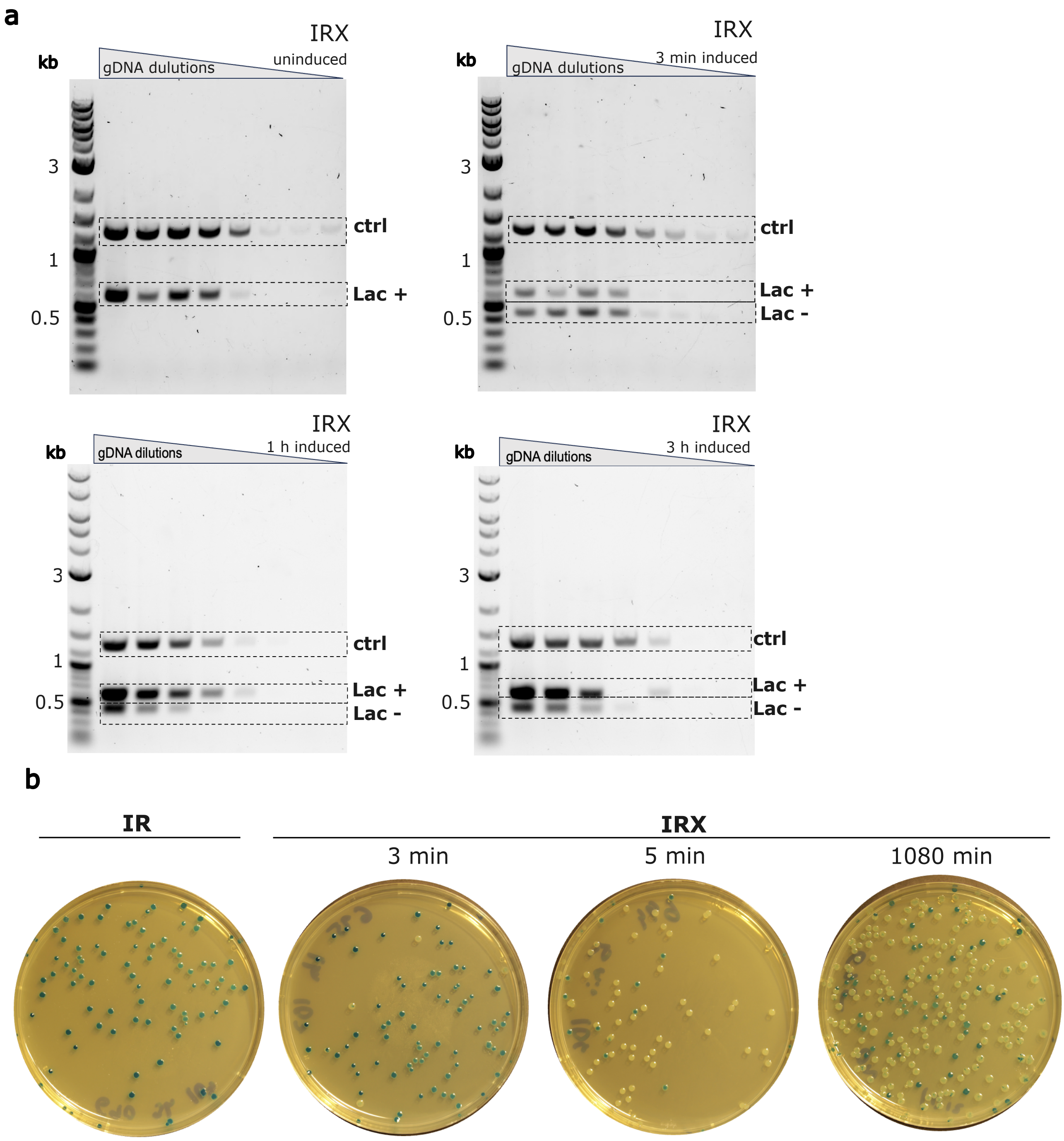


**Figure S2**. **Invertase activity in *E. coli* (a)** Time series of invertase induction in *E. coli* IRX strain. PCR products of serial dilutions of genomic DNA with aTc induction times: 0, 3 min, 1 h, 3 h. **(b)** Example images of plates for the time course of Lac expression during invertase induction quantified via blue-white screening in Figure 4d.

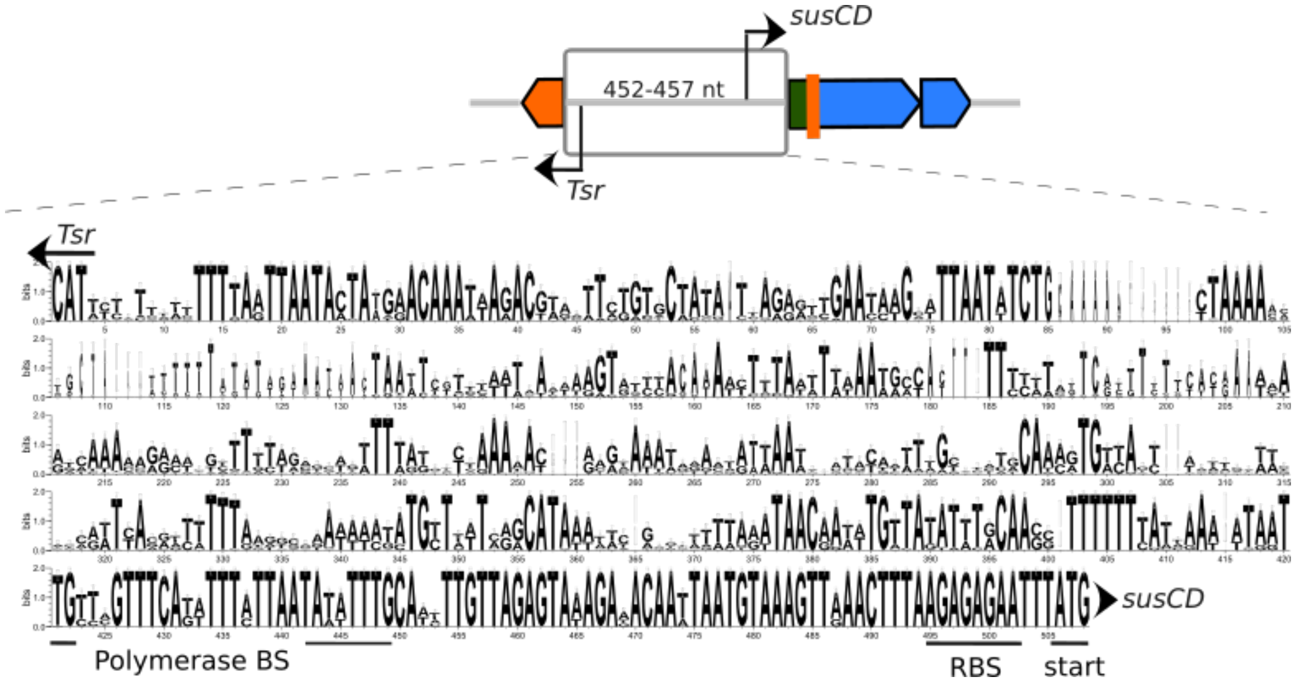


**Figure S3. Logo of the intergenic regions between the *tsr* gene and the expressed *susCD* transporter gene with functional annotation of regulatory elements.**

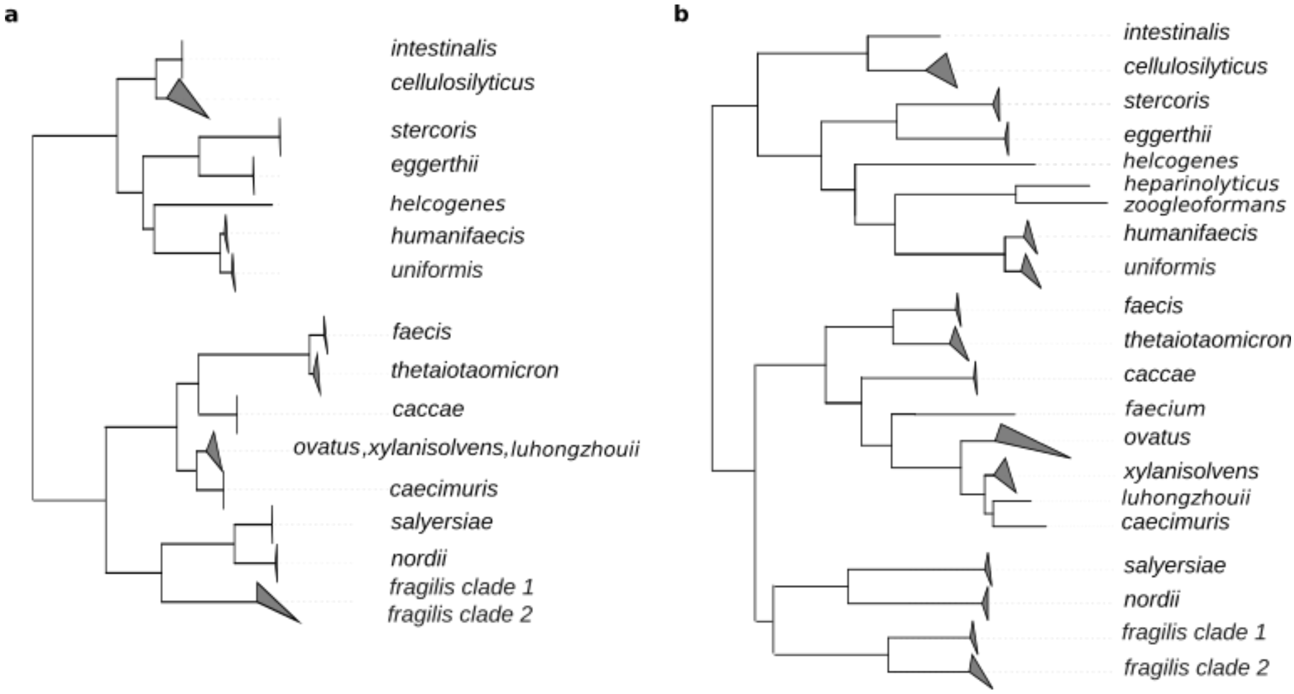


**Figure S4**. **Comparison of the *Bacteroides* phylogenetic trees constructed (a)** based on the concatenated alignment of single-copy common genes and **(b)** alignment of the DNA fragments comprising a *tsr* gene, the shuffled part of the activated *susC* transporter gene, and inter genic region between them.