Genome-wide Transcriptomic Analysis in *Clostridium beijerinckii* 8052 using Single-Nucleotide Resolution RNA-Seq Technology

Yi Wang, **Holger Janssen**, Xiangzhen Li, Yuejian Mao, Hans Blaschek*

Department of Food Science and Human Nutrition;

*Center for Advanced Bioenergy Research (CABER);

University of Illinois at Urbana-Champaign

Urbana, IL, USA

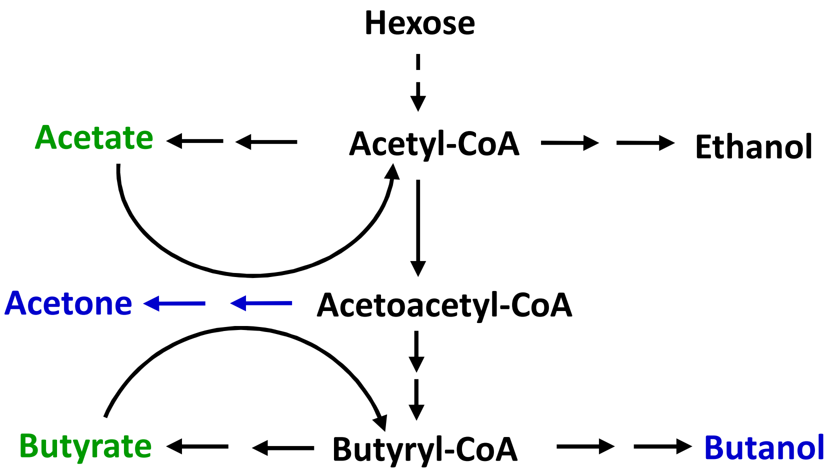
ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

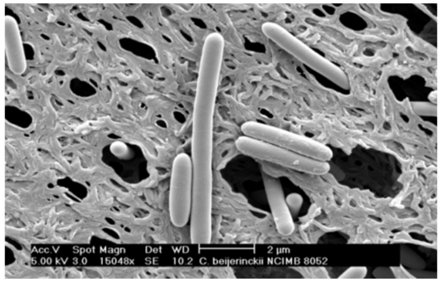
Outline

1. Introduction
    - Research background of *C. beijerinckii* 8052
    - RNA-Seq technology
2. Results
    - Transcriptome structure analysis of *C. beijerinckii* 8052
    - Genome-wide dynamic transcriptional analysis
3. Summary

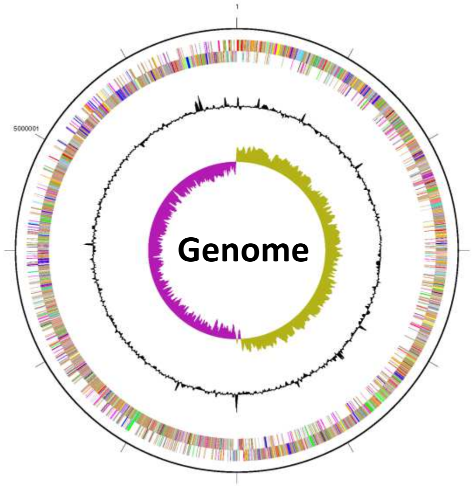
Clostridium beijerinckii 8052

Biphasic fermentation metabolism



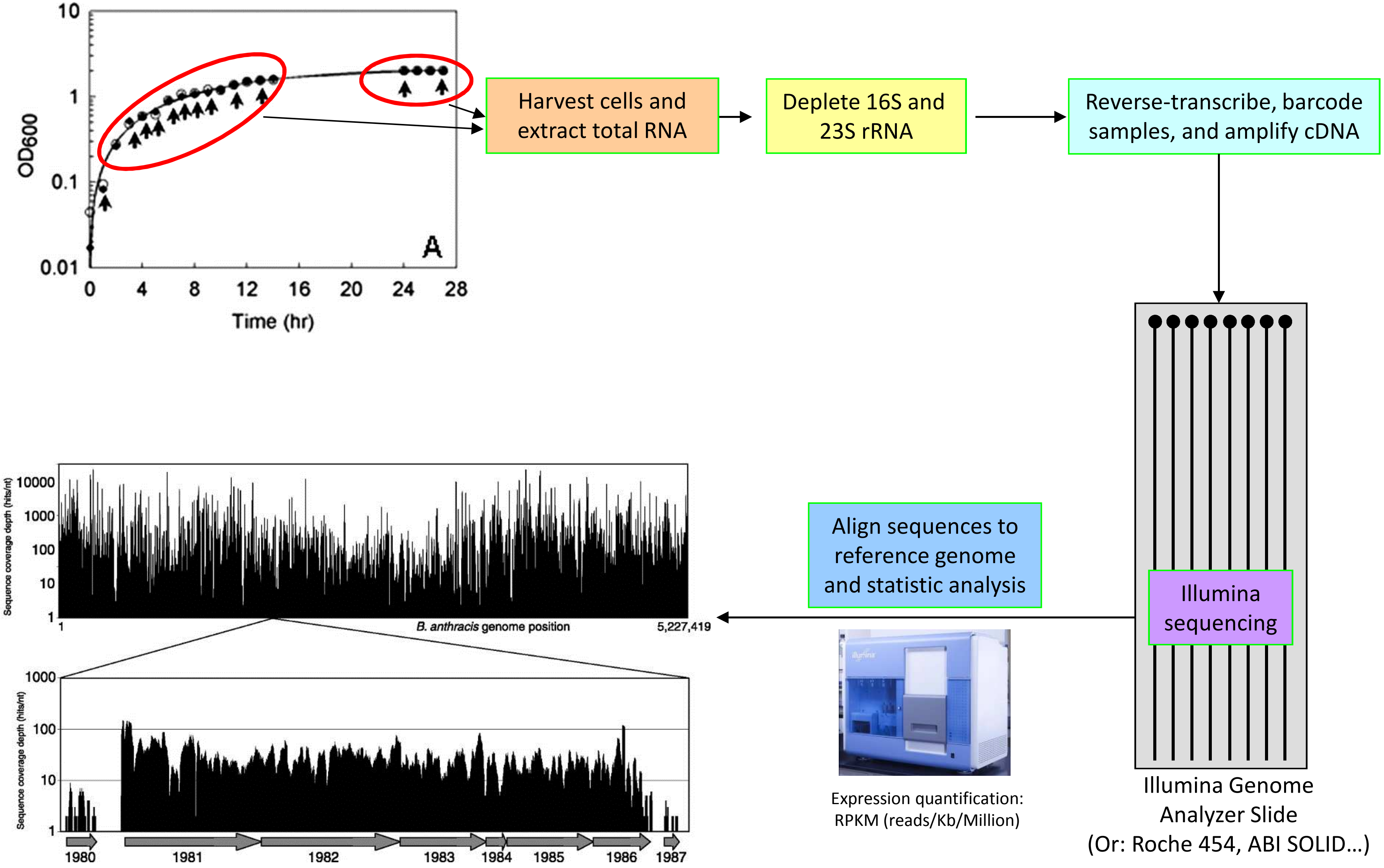


SEM of *C.beijerinckii* 8052



Number of predicted coding sequences	Total bases	Number of genes with predicted function	Number of genes with unknown function
5020	6.000.632	3785	1235

RNA-Seq technology



Harvest cells and extract total RNA

Deplete 16S and 23S rRNA

Reverse-transcribe, barcode samples, and amplify cDNA

Illumina sequencing

Illumina Genome Analyzer Slide
(Or: Roche 454, ABI SOLID…)

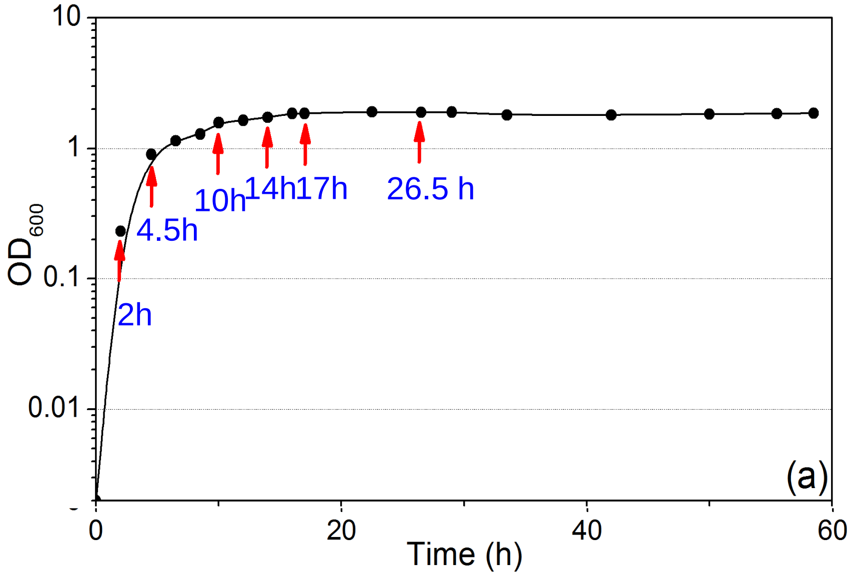
Align sequences to reference genome and statistic analysis

Expression quantification:
RPKM (reads/Kb/Million)

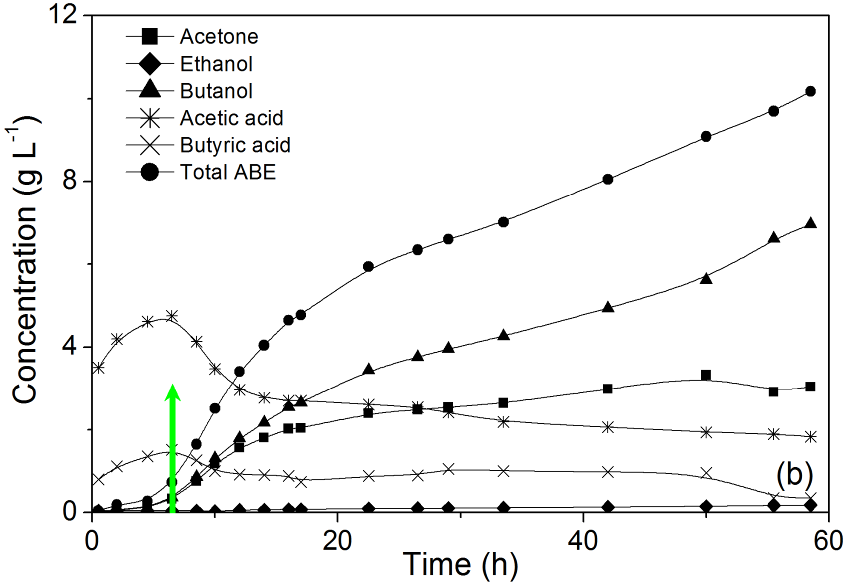
Passalacqua, *et al.* 2009. *J Bacteriol.* 191(10): 3203-3211

Transcriptome structure analysis

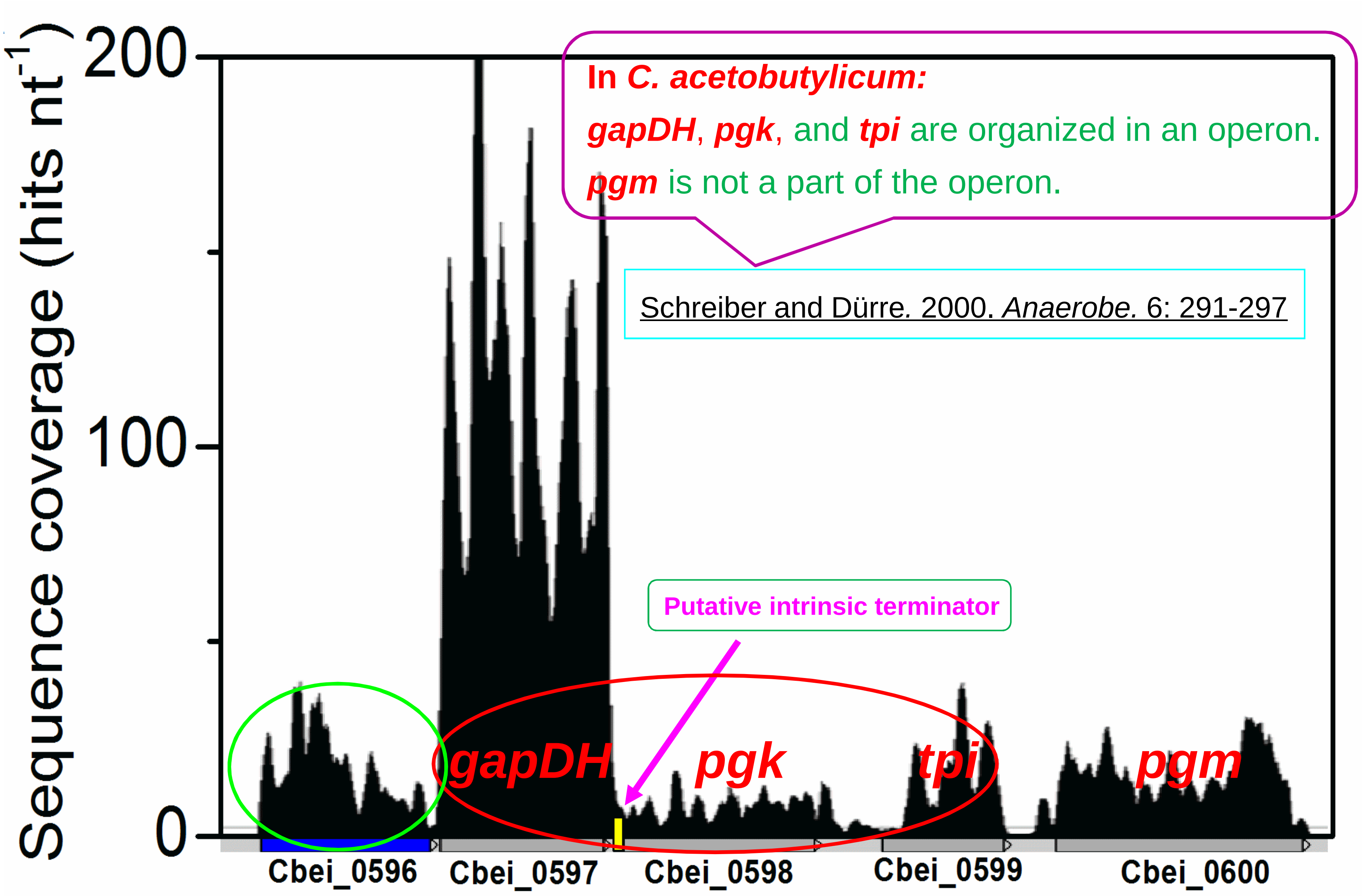
Cell growth



Fermentation products



Glycolysis Gene Operon



In *C. acetobutylicum:*

gapDH, ***pgk***, and ***tpi*** are organized in an operon.

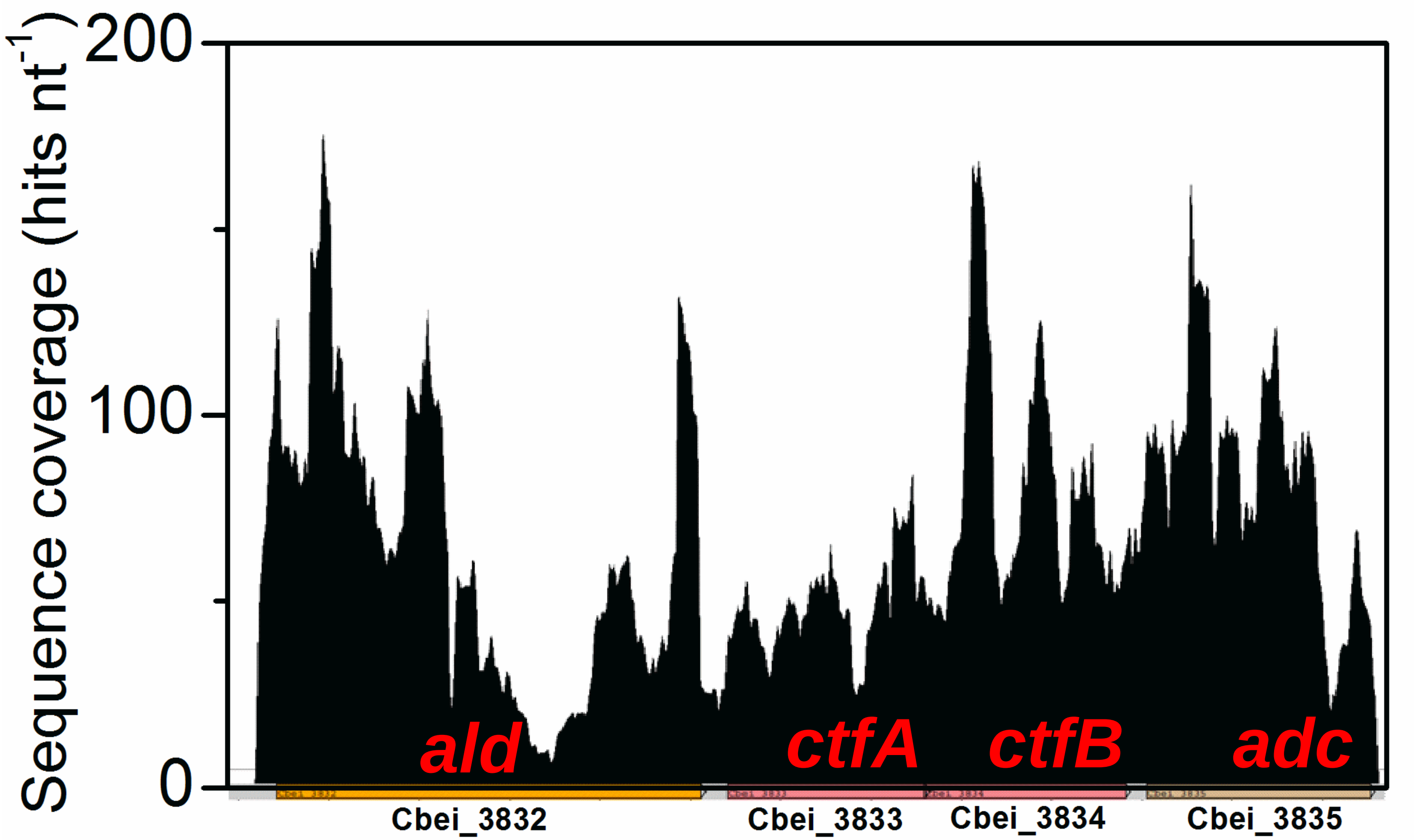
pgm is not a part of the operon.

Schreiber and Dürre. 2000. *Anaerobe.* 6: 291-297

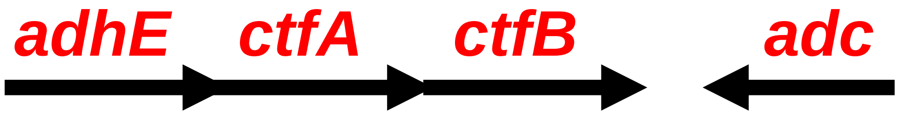
Putative intrinsic terminator

gap: glyceraldehyde-3-phosphate dehydrogenase; ***pgk***: phosphoglycerate kinase;
tpi: triosephosphate isomerase; ***pgm:*** phosphoglycerate mutase

sol operon structure

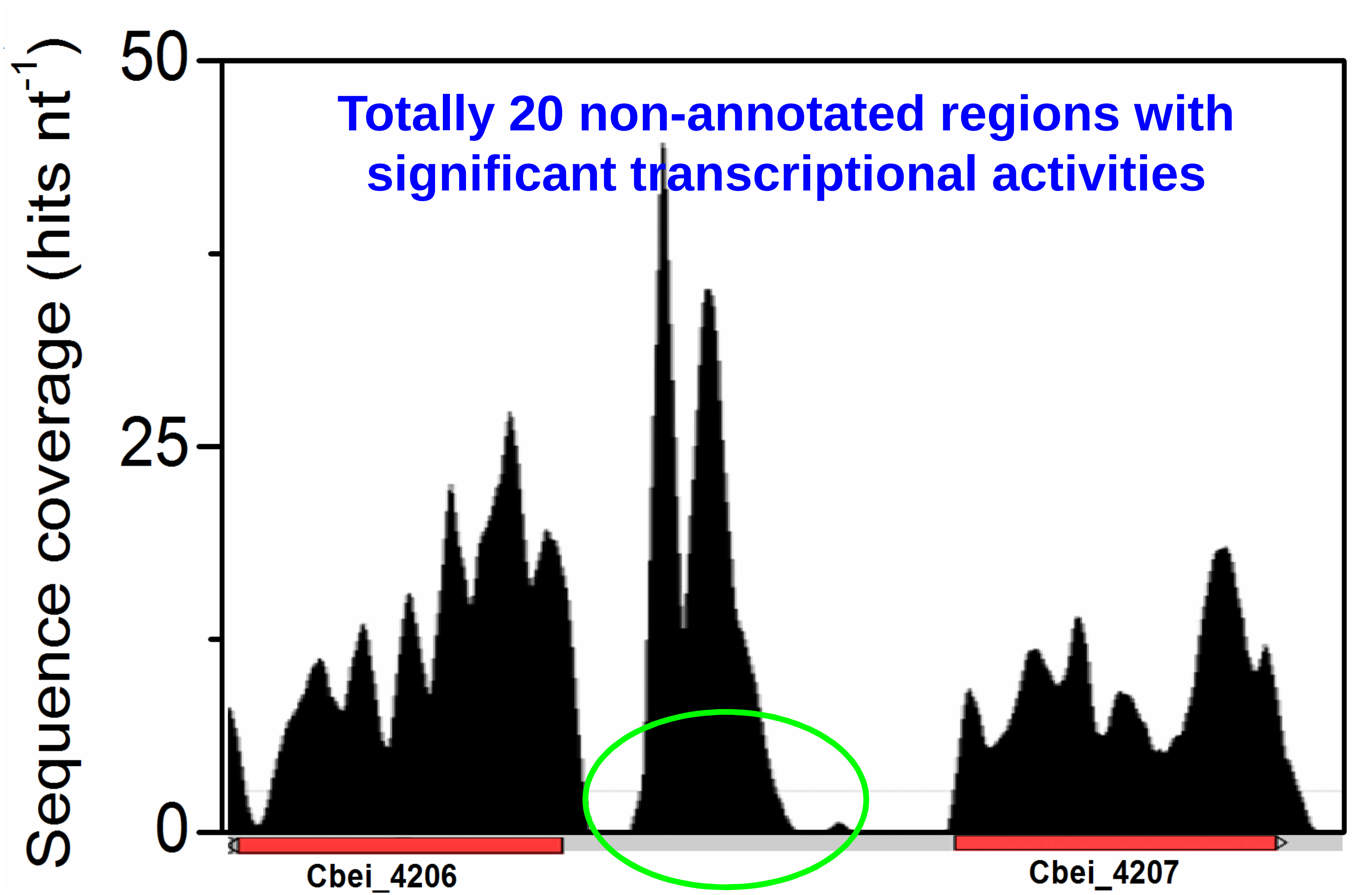


C. acetobutylicum:



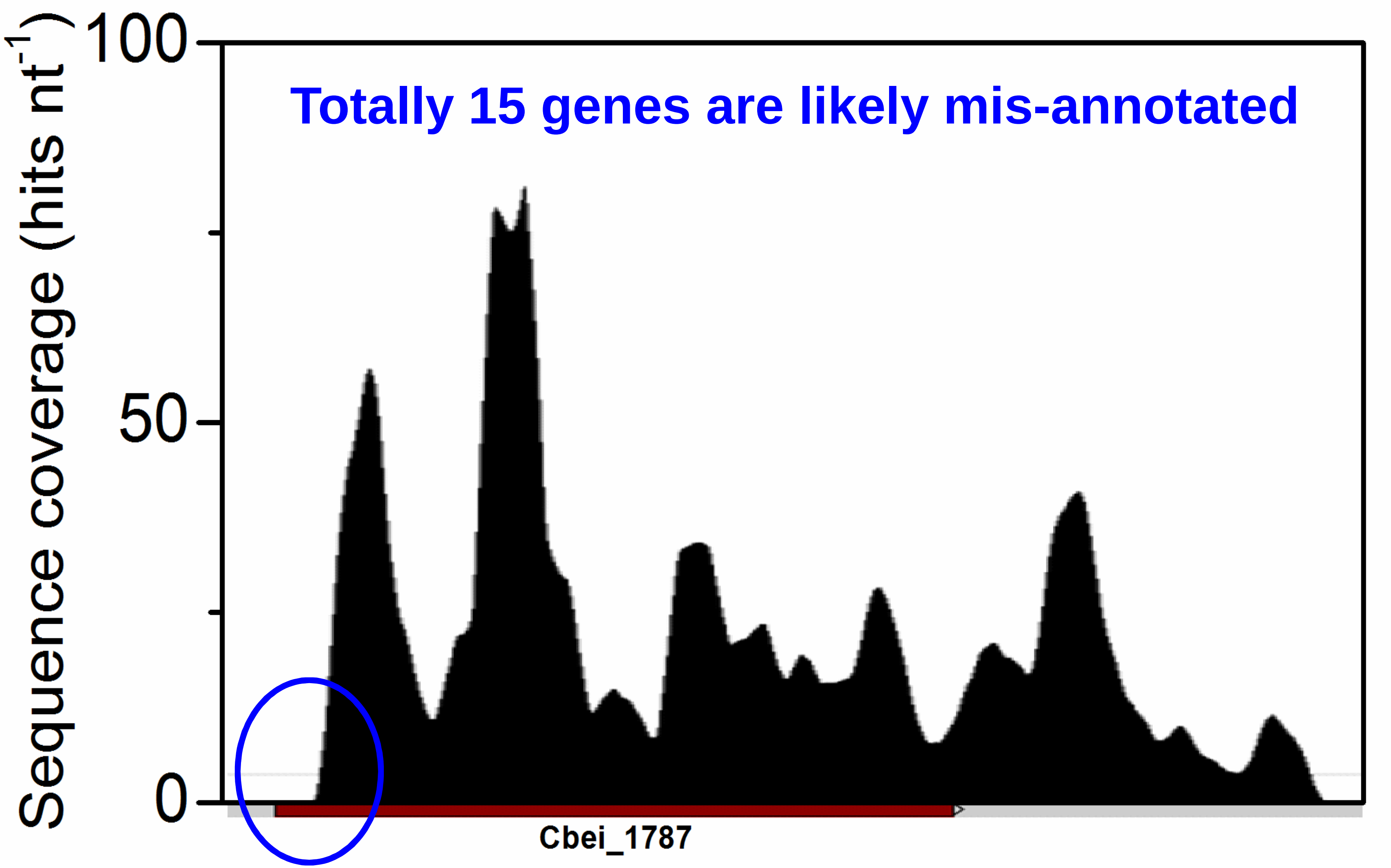
ald: aldehyde dehydrogenase; ***ctfA/B:*** acetoacetyl-CoA: acetate/butyrate-CoA transferase subunit A/B;
adc: acetoacetate decarboxylase ***adhE:*** acetaldehyde-CoA/alcohol dehydrogenase;

Transcription activity **not currently annotated**

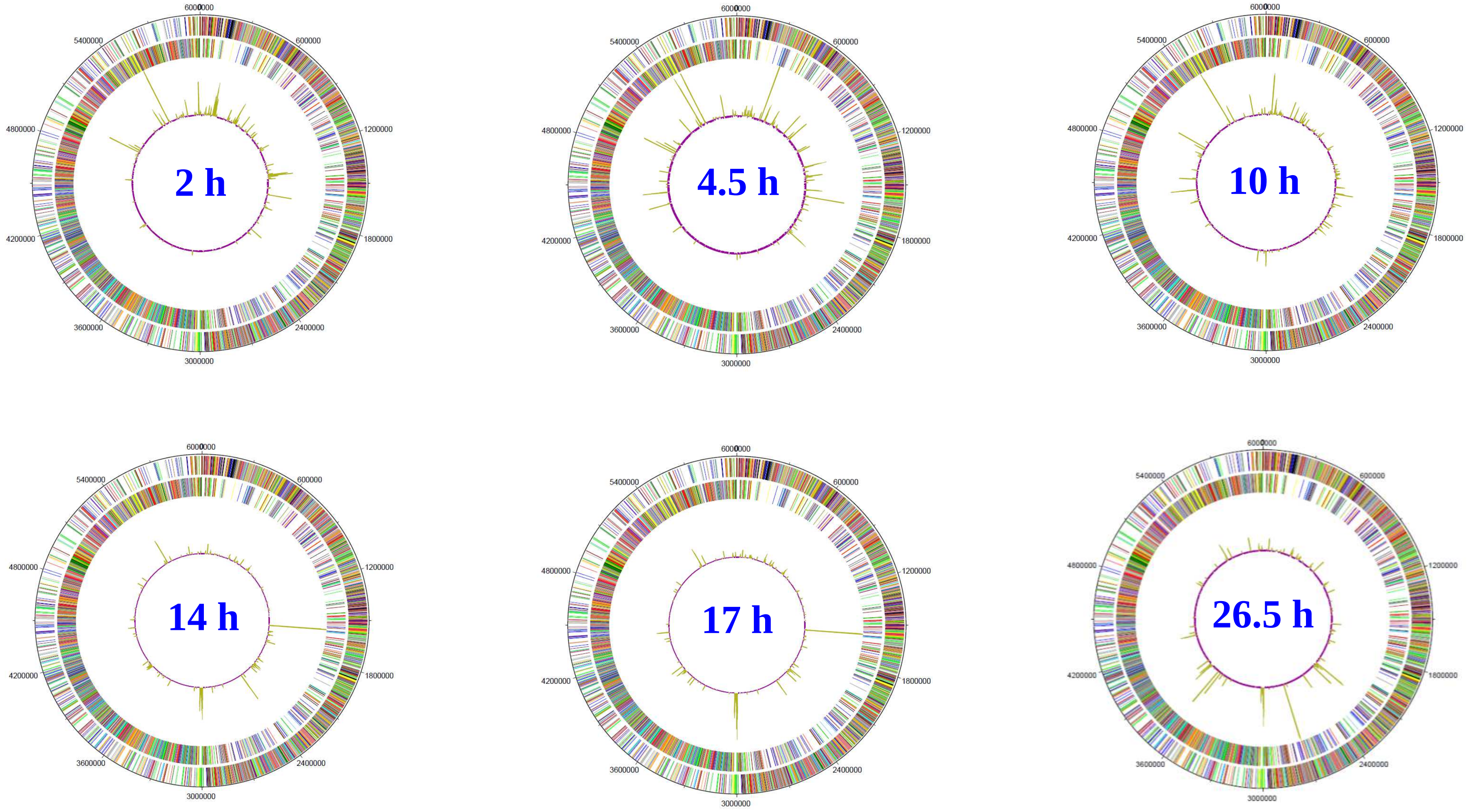


Totally 20 non-annotated regions with significant transcriptional activities

Putative **mis-annotated** regions

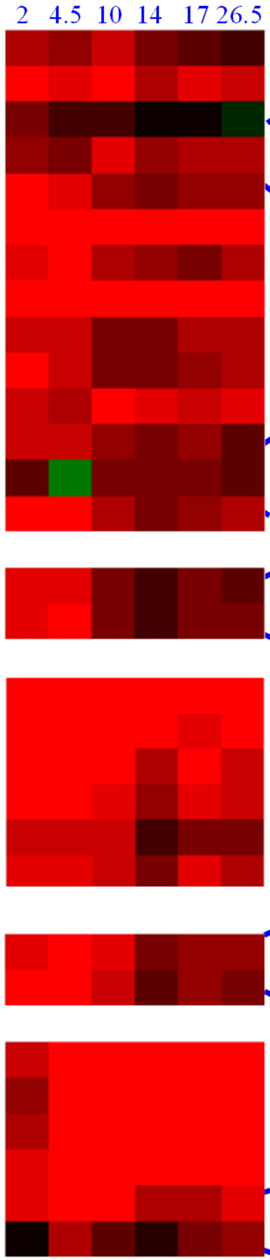


Genome-wide dynamic transcriptional analysis



Circular plots of the sequence data mapping to the *C. beijerinckii* 8052 genome

Expression of **primary metabolic genes**



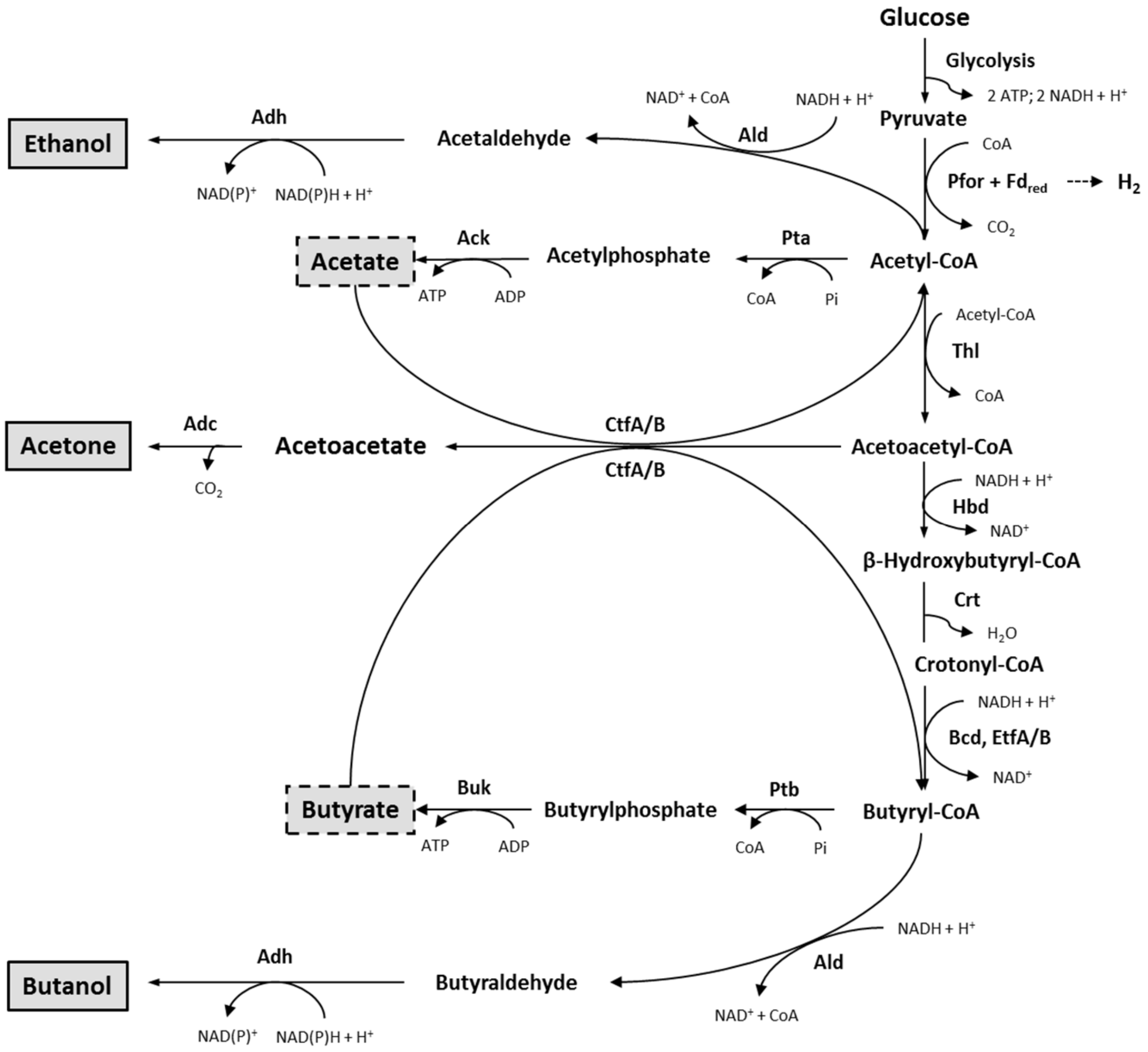
Glycolysis genes
Cbei_4895, *glcK*
Cbei_0341, *pgi*
Cbei_0584, *pfk*
Cbei_0998, *pfk*
Cbei_4852, *pfk*
Cbei_1903, *fba*
Cbei_0599, *tpi*
Cbei_0597, *gap*
Cbei_0598, *pgk*
Cbei_0600, *pgm*
Cbei_0602, *eno*
Cbei_0485, *pyk*
Cbei_1412, *pyk*
Cbei_4851, *pyk*

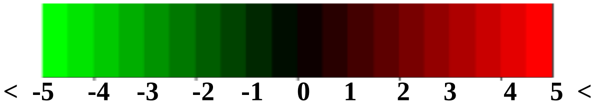
Acetate formation genes
Cbei_1164, *pta*
Cbei_1165, *ack*

Butyryl-CoA formation genes
Cbei_0411, *thlA*
Cbei_0325, *hbd*
Cbei_0321, *crt*
Cbei_0322, *bcd*
Cbei_0324, *etfA*
Cbei_0323, *etfB*

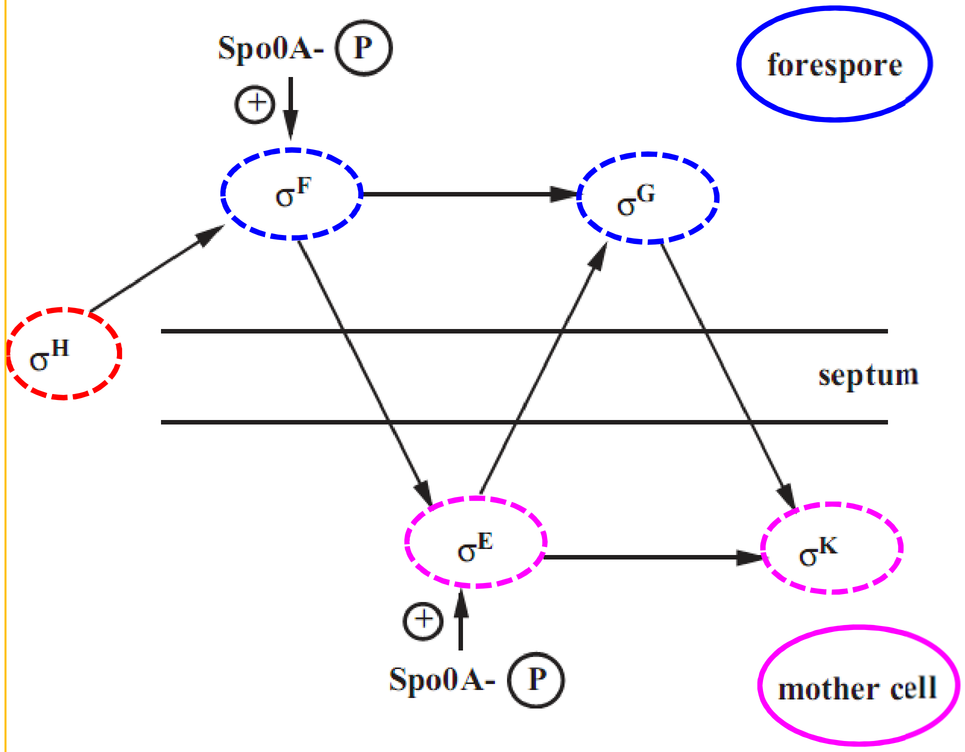
Butyrate formation genes
Cbei_0203, *ptb*
Cbei_0204, *buk*

Solventogenesis genes
Cbei_3832, *ald*
Cbei_3833, *ctfA*
Cbei_3834, *ctfB*
Cbei_3835, *adc*
Cbei_1722, *adh*
Cbei_2181, *adh*

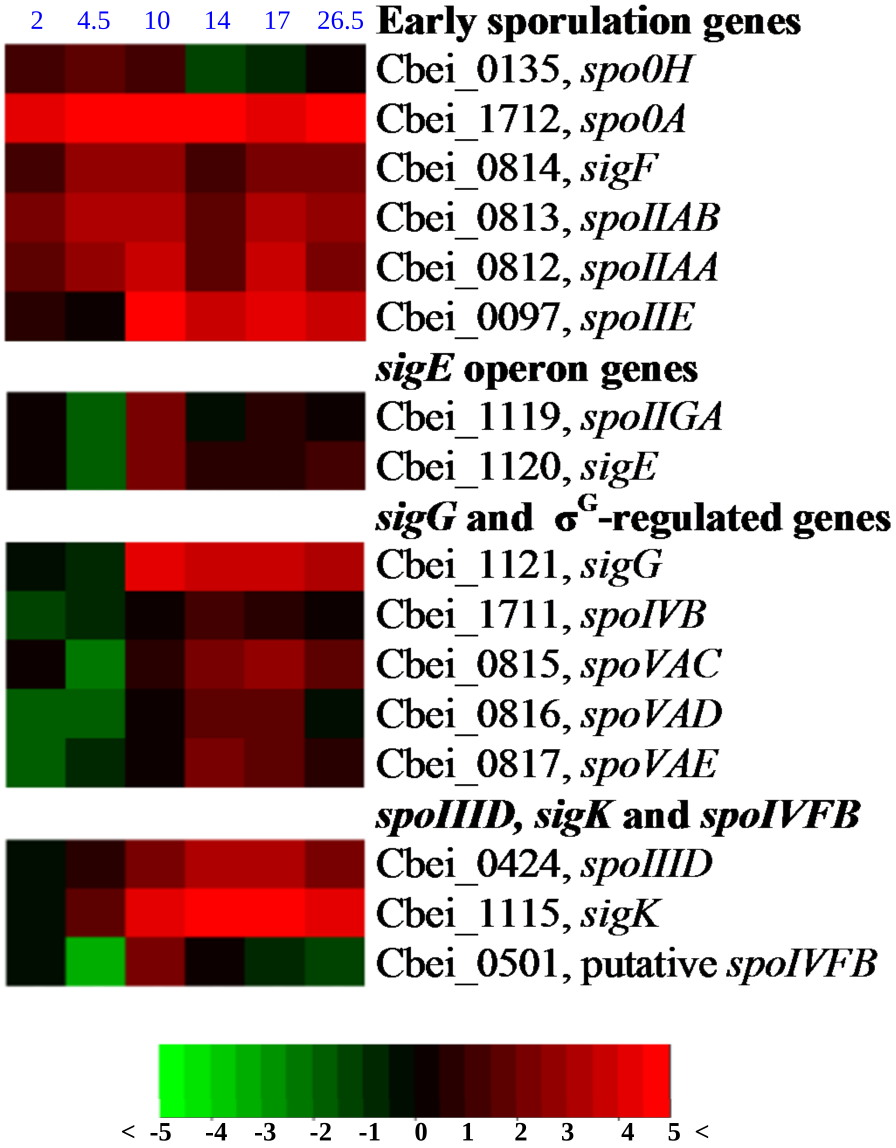




Expression of **sporulation genes**

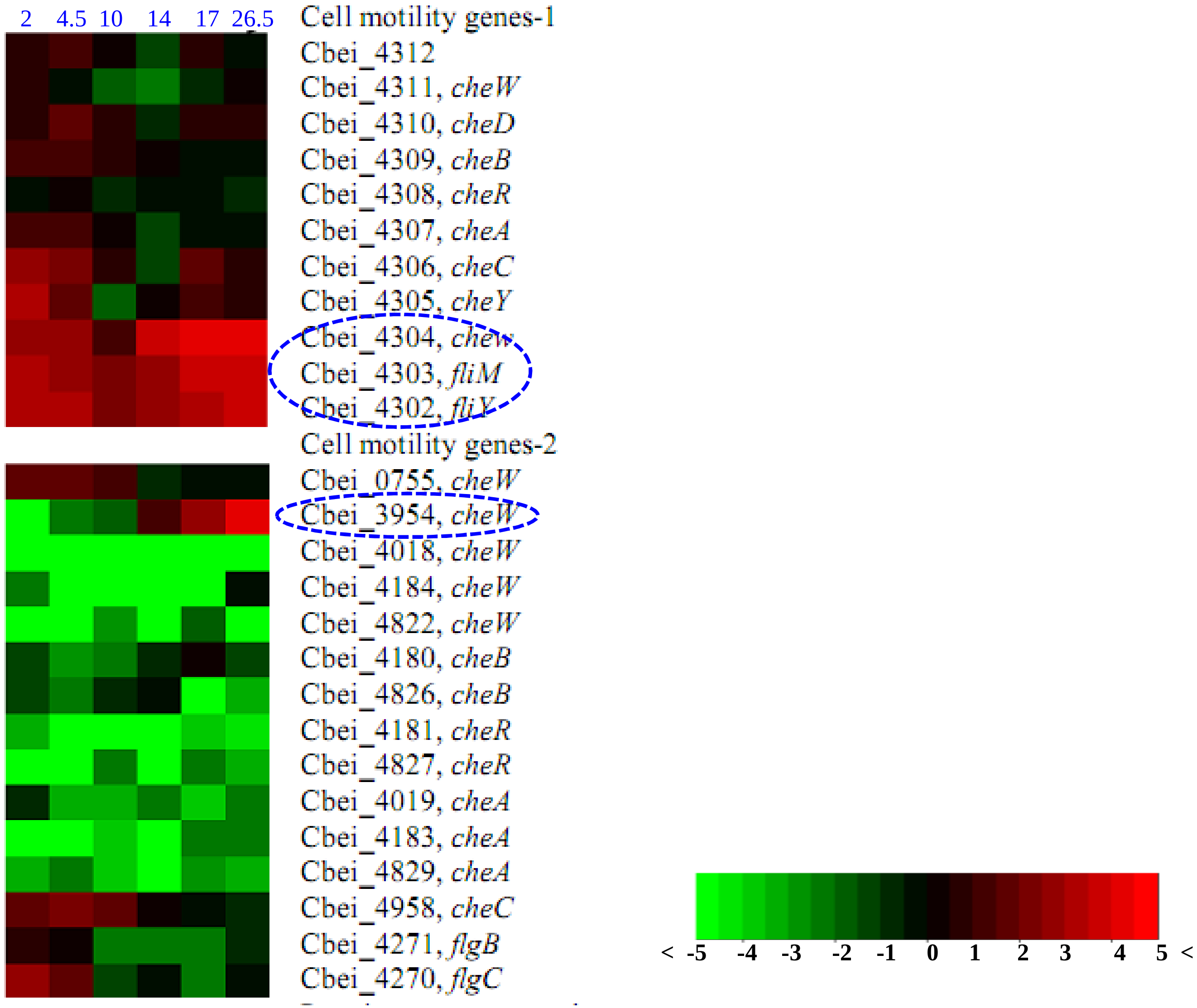


“Crisscross regulation” during endospore formation in *B. subtilis* and *C. acetobutylicum**



*Dürre and Hollergschwandner. 2004. *Anaerobe*. 10:69-74

Expression of **motility genes**



Summary

- RNA-Seq provided both transcriptome structure and quantitative transcriptional information
- Transcriptome analyses using RNA-seq complement and improve genome annotation
- Acid and solvent formation genes exhibited temporal patterns in accordance with switch from acidogenesis to solventogenesis
- Sporulation genes were sequentially induced and demonstrated expression patterns corresponding to the sporulation program

Acknowledgements

- Department of Energy (DOE)
- Dr. Roderick I. Mackie
- Dr. Dylan Dodd
- Dr. Heng Li (Broad Institute, Cambridge, MA)
- Dr. Timothy Perkins (Wellcome Trust Sanger Institute, UK)
- Drs. Nathan Price and Yong-Su Jin & their group members