THE ESSENTIAL GENOME OF A BACTERIUM

2011/10/21
鲁 月

WHAT IS ESSENTIAL GENOME?

Essential genome:

- protein-coding sequences
- essential structural elements
- noncoding RNAs
- regulatory sequences

HOW TO IDENTIFY ESSENTIAL GENOME?

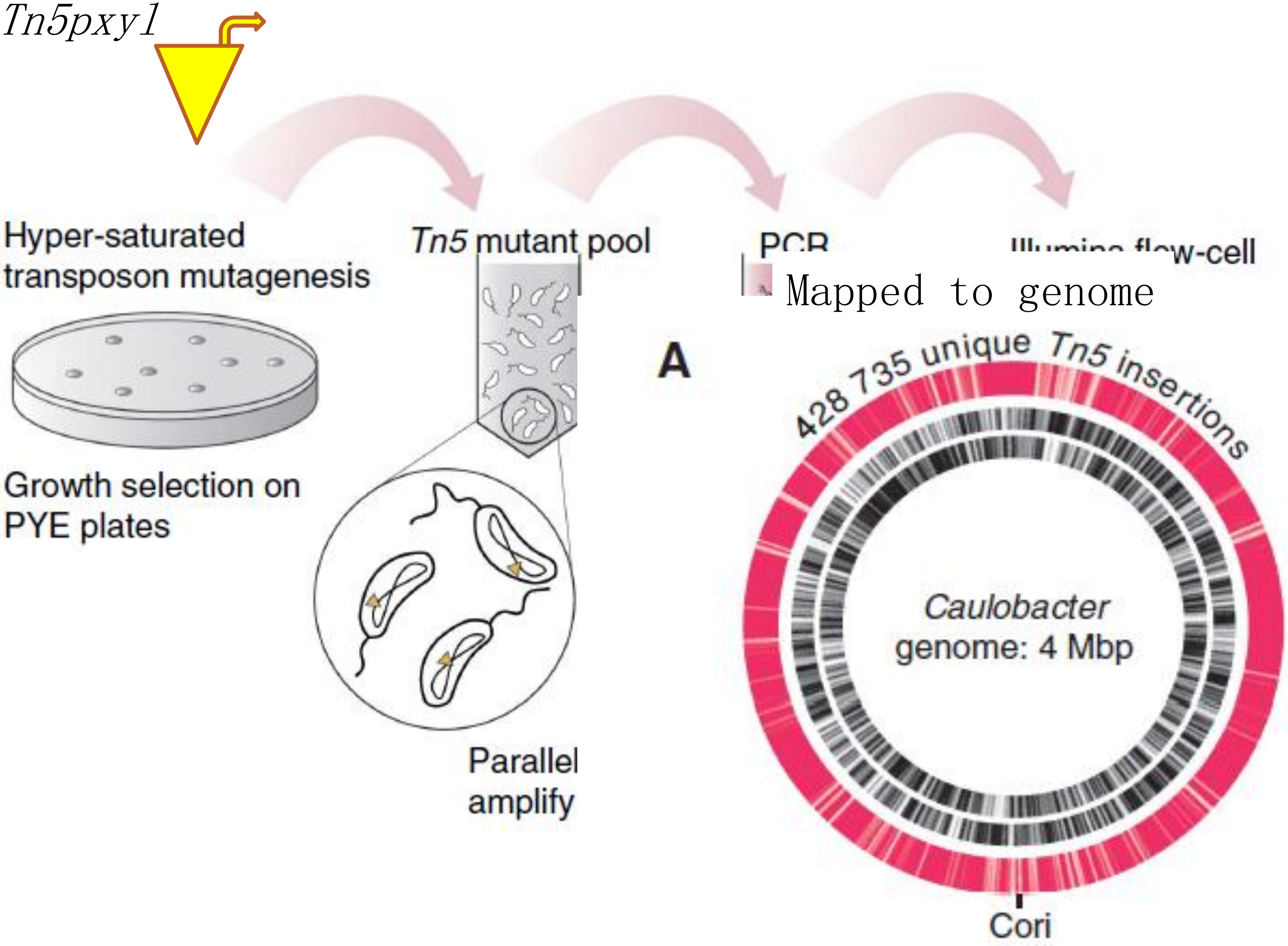
- Previously,

relatively low-throughput transposon mutagenesis (Hutchison et al, 1999; Jacobs et al, 2003; Glass et al, 2006)

in-frame deletion libraries (Kobayashi et al, 2003; Baba et al, 2006)

- In this study
- ultrahigh-resolution transposon mutagenesis+
- high-throughput DNA sequencing

Tn5pxyl



Mapped to genome

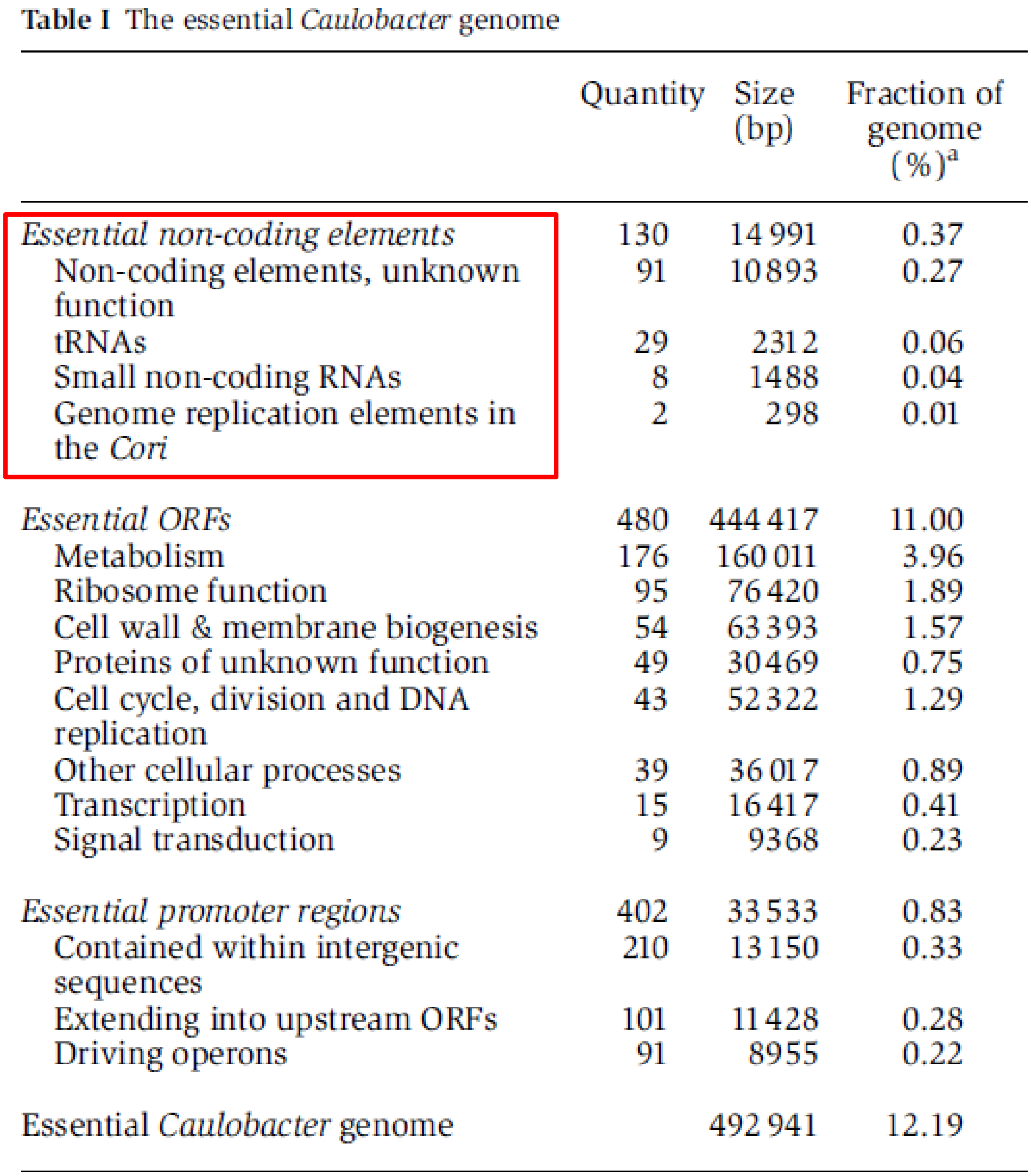
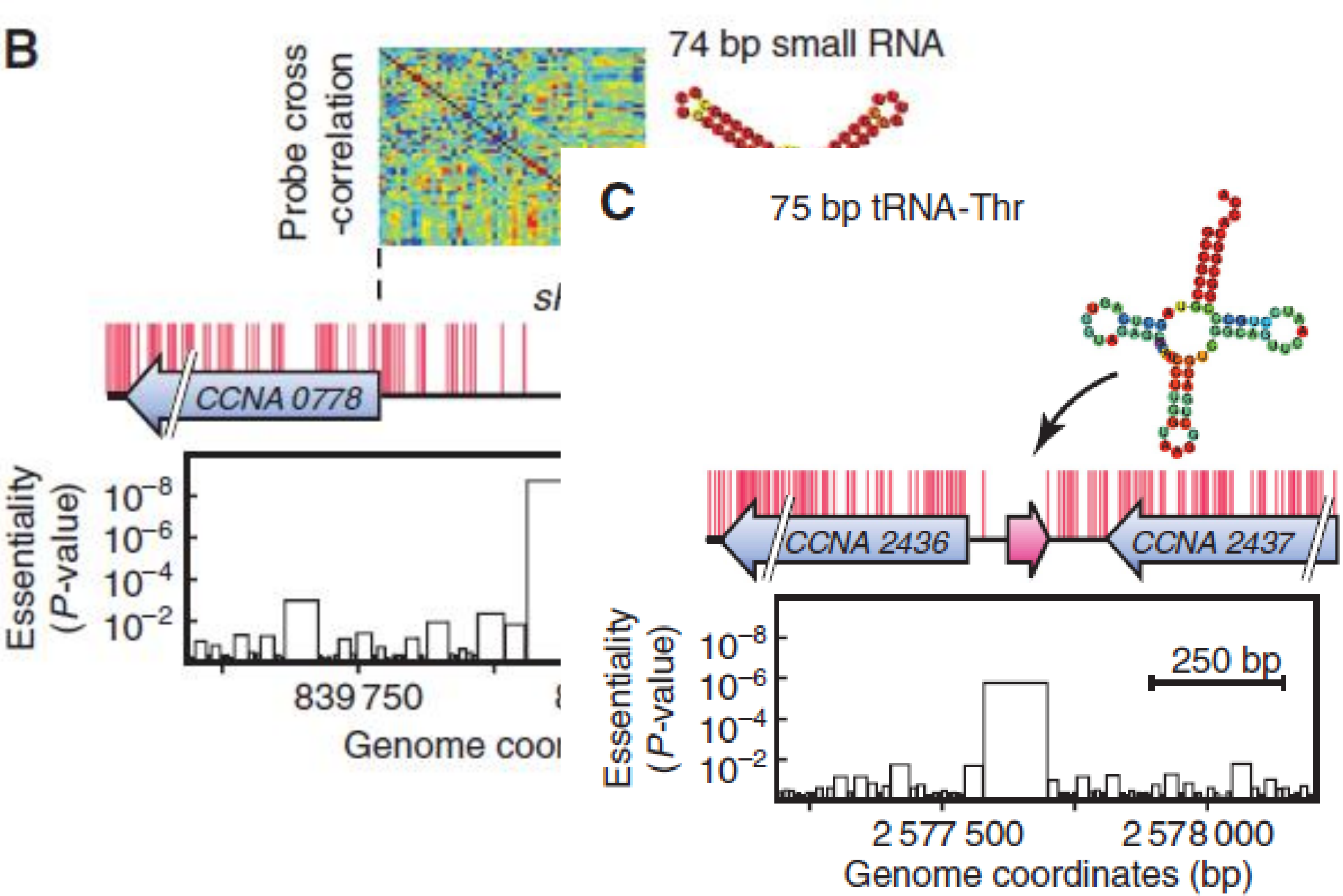
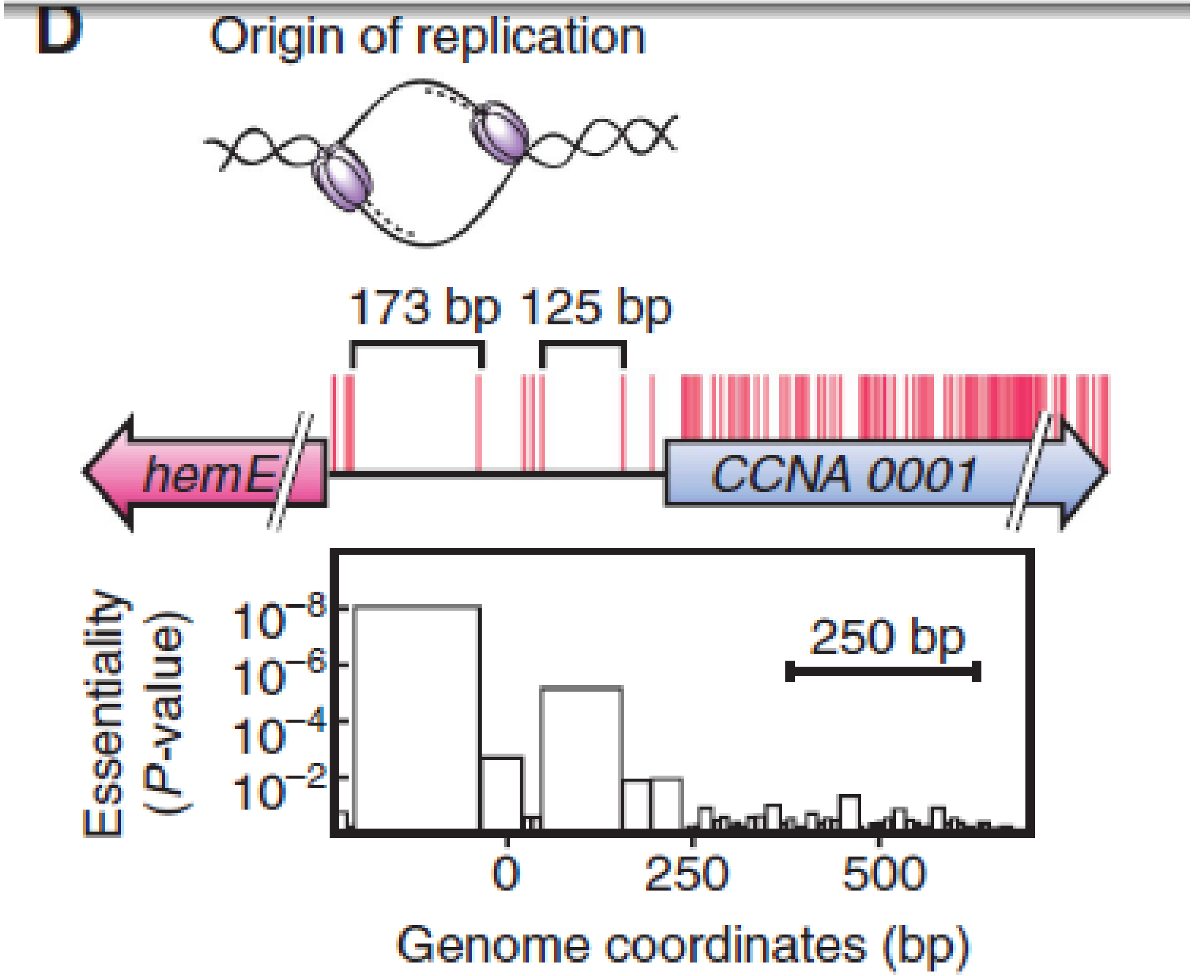


Table I The essential *Caulobacter* genome

	Quantity	Size (bp)	Fraction of genome (%)[a]
Essential non-coding elements	130	14 991	0.37
Non-coding elements, unknown function	91	10 893	0.27
tRNAs	29	2312	0.06
Small non-coding RNAs	8	1488	0.04
Genome replication elements in the *Cori*	2	298	0.01
Essential ORFs	480	444 417	11.00
Metabolism	176	160 011	3.96
Ribosome function	95	76 420	1.89
Cell wall & membrane biogenesis	54	63 393	1.57
Proteins of unknown function	49	30 469	0.75
Cell cycle, division and DNA replication	43	52 322	1.29
Other cellular processes	39	36 017	0.89
Transcription	15	16 417	0.41
Signal transduction	9	9368	0.23
Essential promoter regions	402	33 533	0.83
Contained within intergenic sequences	210	13 150	0.33
Extending into upstream ORFs	101	11 428	0.28
Driving operons	91	8955	0.22
Essential *Caulobacter* genome		492 941	12.19

Essential non-coding seq.





- Binding site for CtrA
- Binding motif for DnaA
- Modular with DNA looping

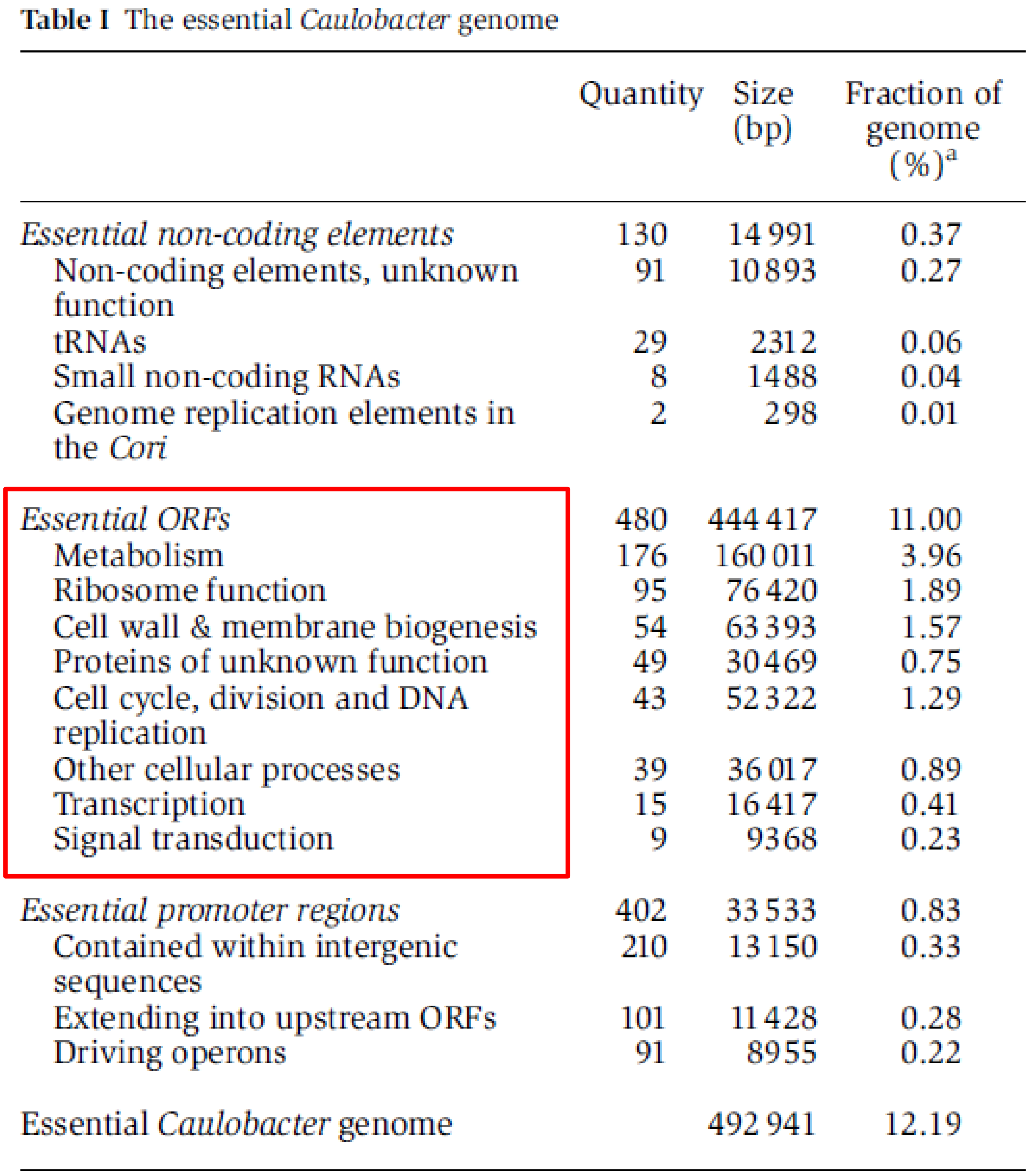
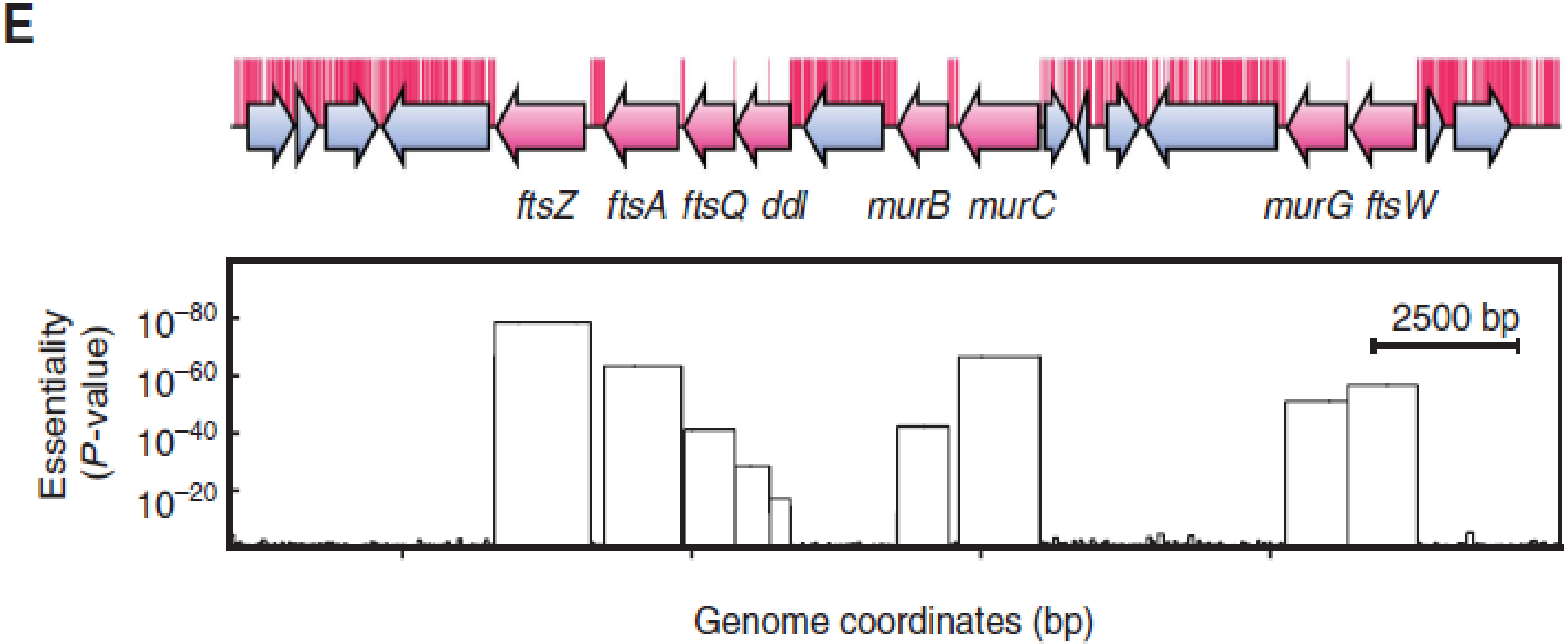
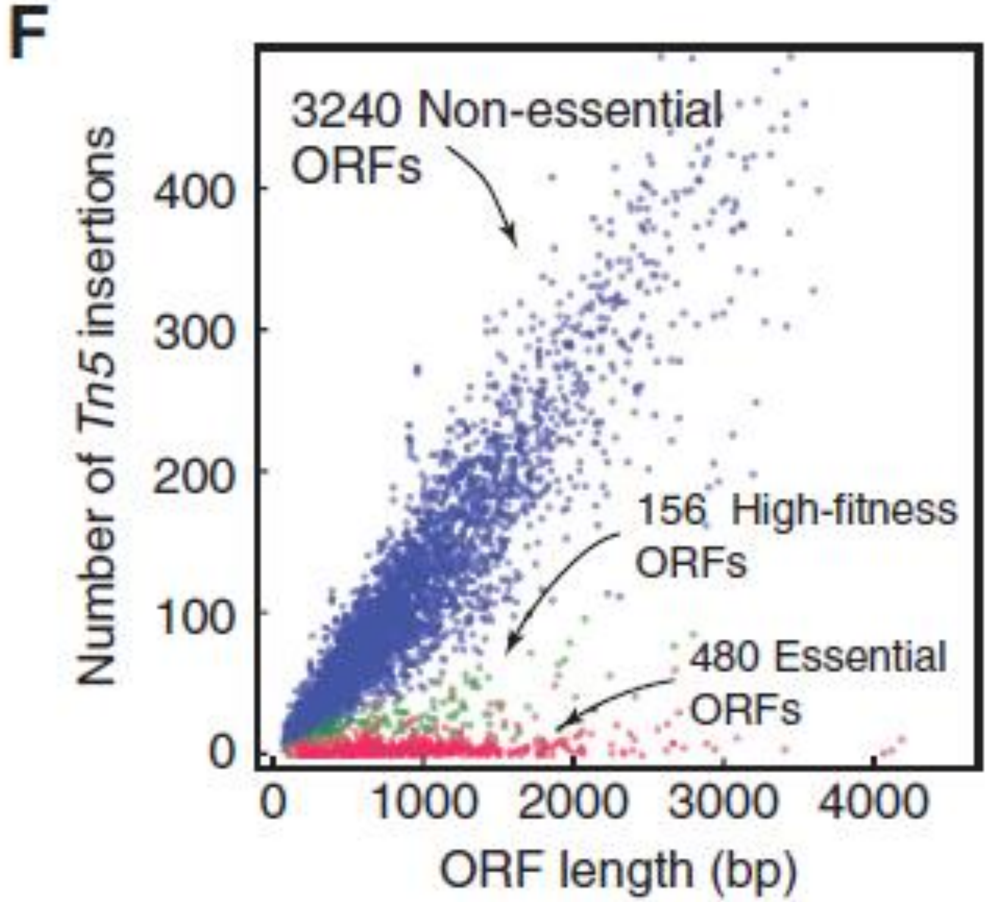


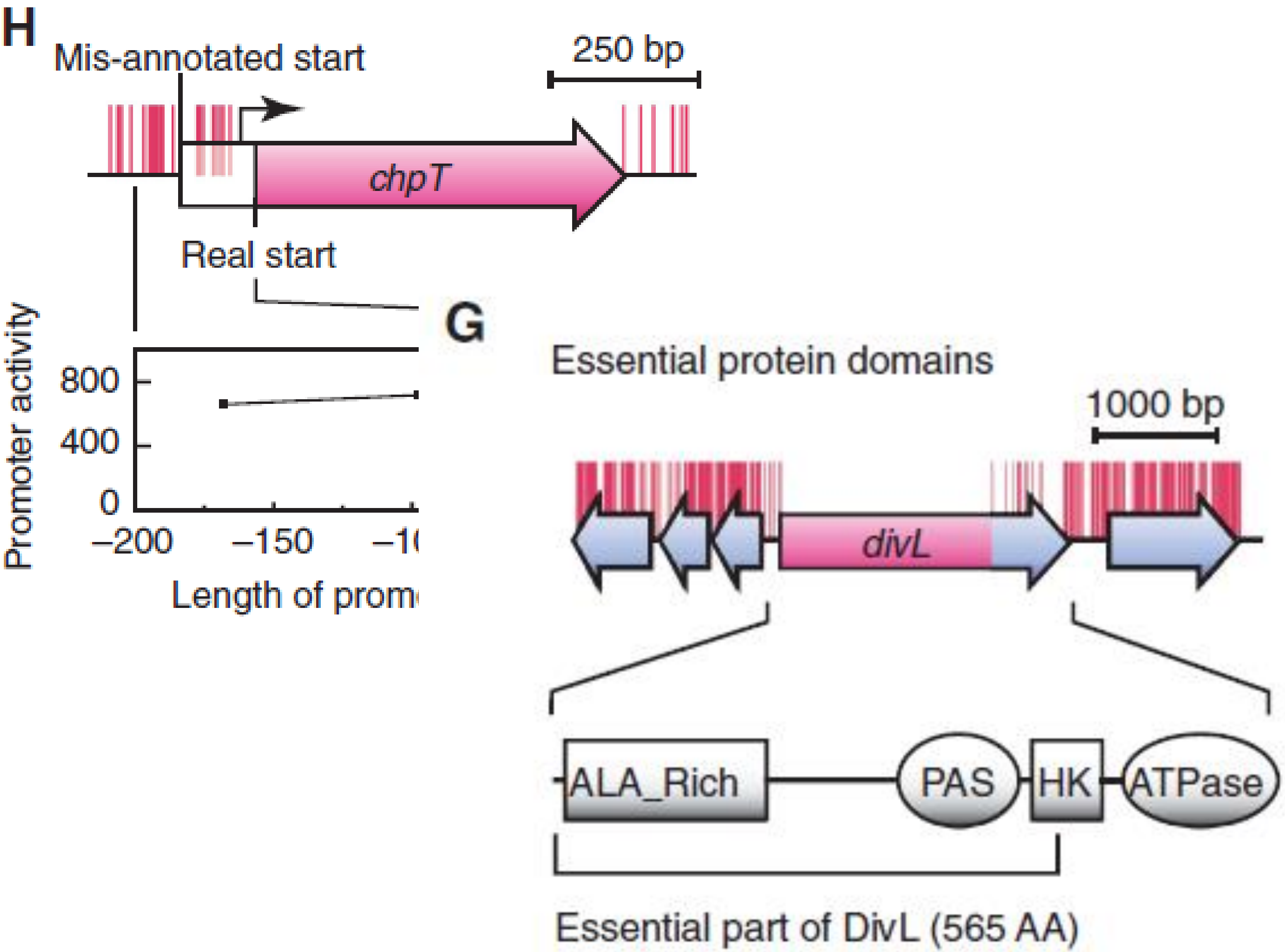
Table I The essential *Caulobacter* genome

	Quantity	Size (bp)	Fraction of genome (%)[a]
Essential non-coding elements	130	14 991	0.37
Non-coding elements, unknown function	91	10 893	0.27
tRNAs	29	2312	0.06
Small non-coding RNAs	8	1488	0.04
Genome replication elements in the *Cori*	2	298	0.01
Essential ORFs	480	444 417	11.00
Metabolism	176	160 011	3.96
Ribosome function	95	76 420	1.89
Cell wall & membrane biogenesis	54	63 393	1.57
Proteins of unknown function	49	30 469	0.75
Cell cycle, division and DNA replication	43	52 322	1.29
Other cellular processes	39	36 017	0.89
Transcription	15	16 417	0.41
Signal transduction	9	9368	0.23
Essential promoter regions	402	33 533	0.83
Contained within intergenic sequences	210	13 150	0.33
Extending into upstream ORFs	101	11 428	0.28
Driving operons	91	8955	0.22
Essential *Caulobacter* genome		492 941	12.19

ESSENTIAL PROTEIN-CODING SEQ.




F
3240 Non-essential ORFs
156 High-fitness ORFs
480 Essential ORFs
Number of *Tn5* insertions
0
100
200
300
400
0
1000
2000
3000
4000
ORF length (bp)


H
Mis-annotated start
250 bp
chpT
Real start
Promoter activity
800
400
0
–200
–150
–1
Length of prom
G
Essential protein domains
1000 bp
divL
ALA_Rich
PAS
HK
ATPase
Essential part of DivL (565 AA)

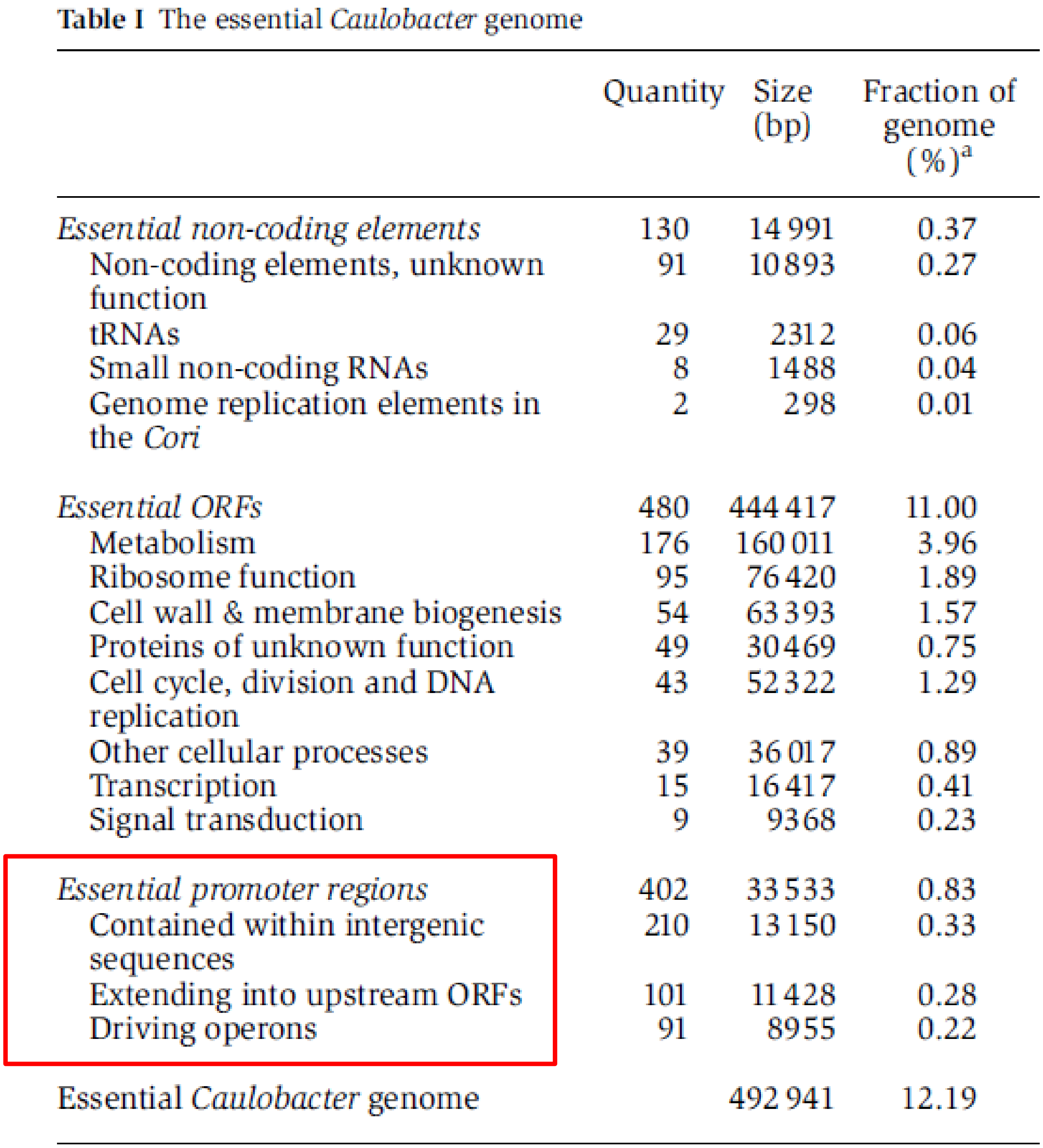
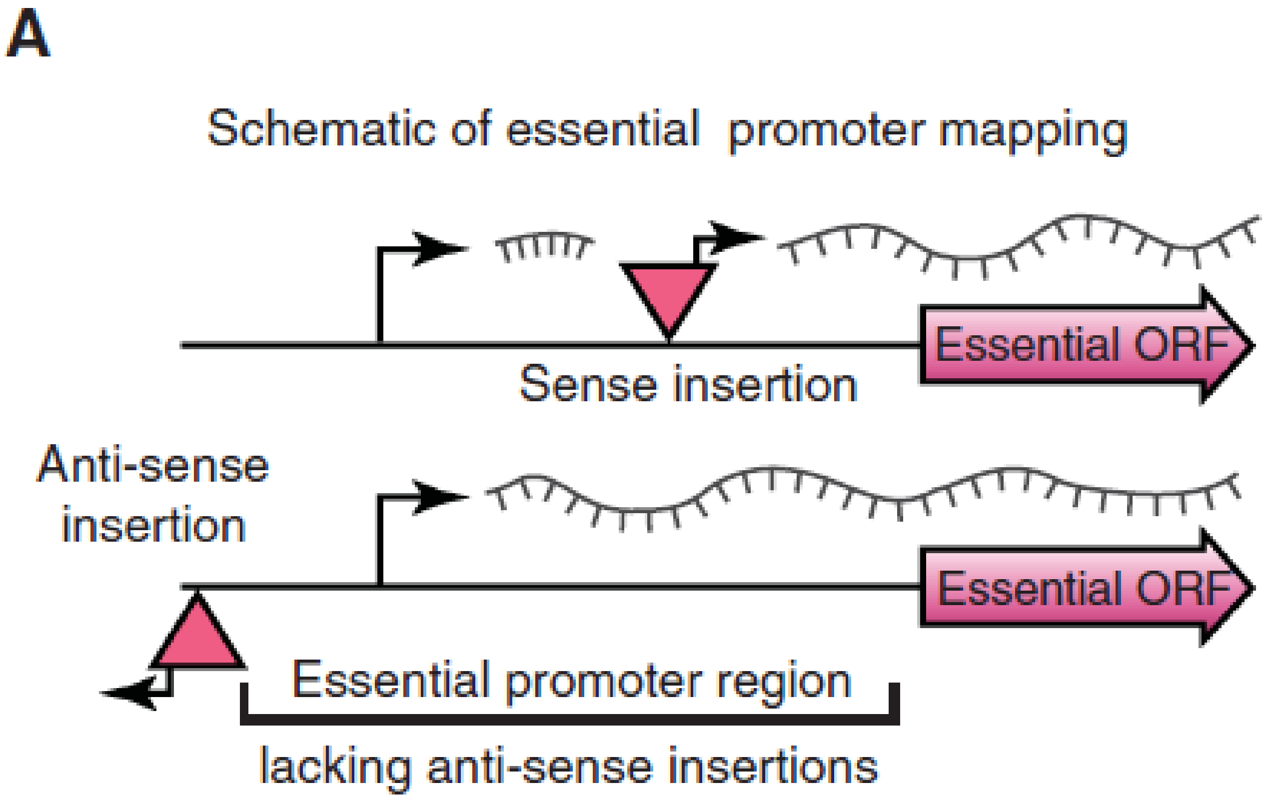
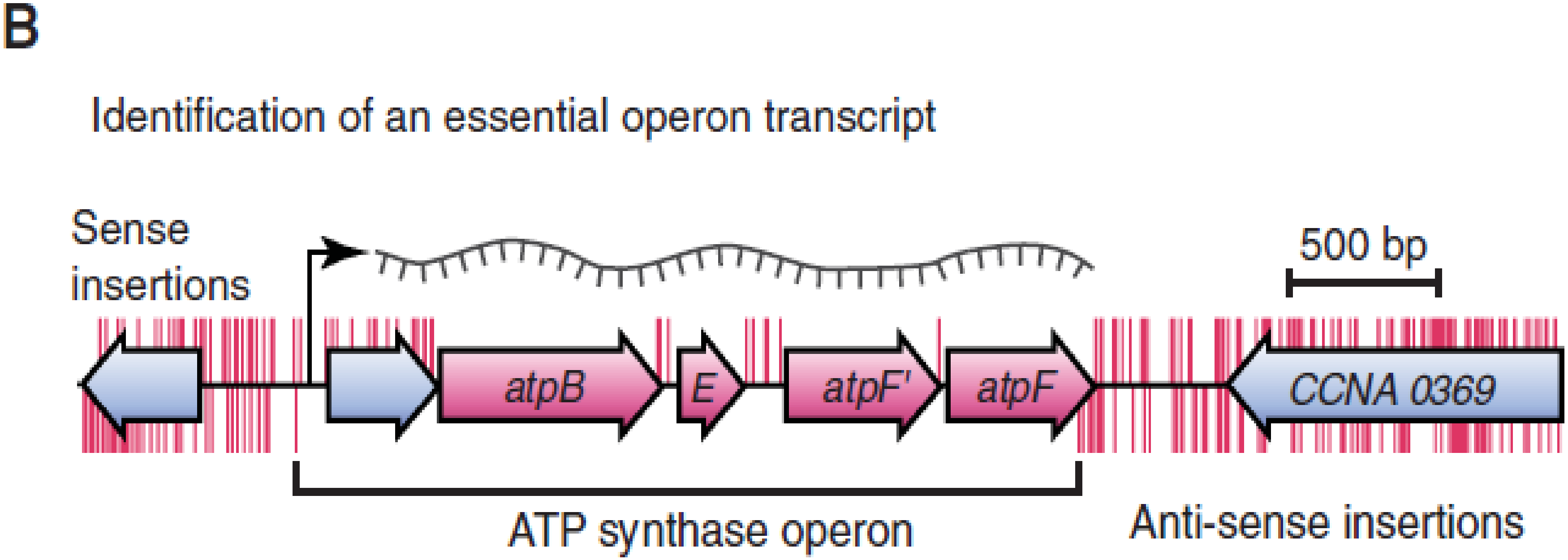


Table 1 The essential *Caulobacter* genome

	Quantity	Size (bp)	Fraction of genome (%)[a]
Essential non-coding elements	130	14 991	0.37
Non-coding elements, unknown function	91	10 893	0.27
tRNAs	29	2312	0.06
Small non-coding RNAs	8	1488	0.04
Genome replication elements in the *Cori*	2	298	0.01
Essential ORFs	480	444 417	11.00
Metabolism	176	160 011	3.96
Ribosome function	95	76 420	1.89
Cell wall & membrane biogenesis	54	63 393	1.57
Proteins of unknown function	49	30 469	0.75
Cell cycle, division and DNA replication	43	52 322	1.29
Other cellular processes	39	36 017	0.89
Transcription	15	16 417	0.41
Signal transduction	9	9368	0.23
Essential promoter regions	402	33 533	0.83
Contained within intergenic sequences	210	13 150	0.33
Extending into upstream ORFs	101	11 428	0.28
Driving operons	91	8955	0.22
Essential *Caulobacter* genome		492 941	12.19

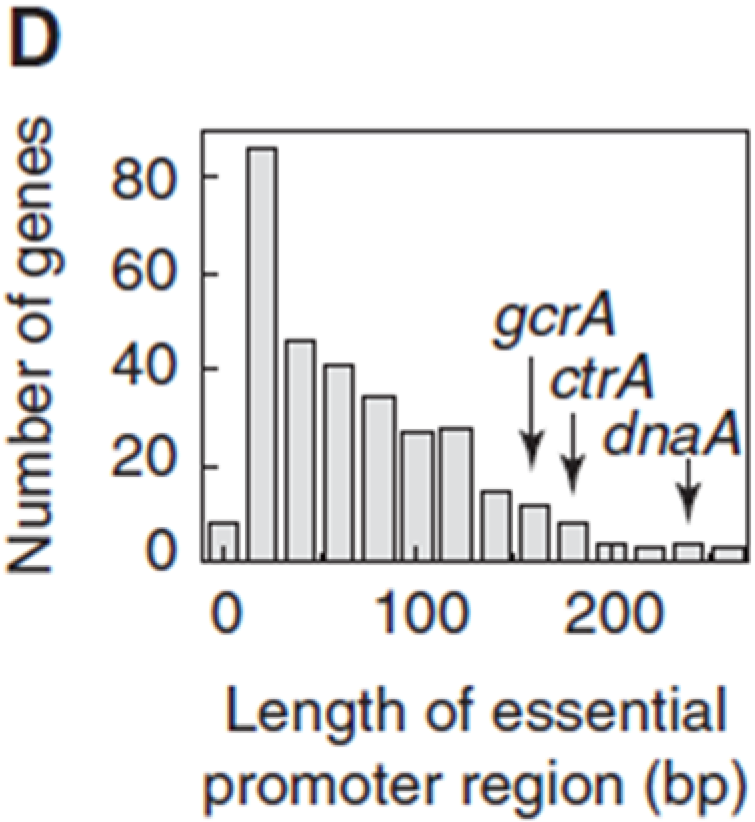
ESSENTIAL PROMOTER ELEMENTS



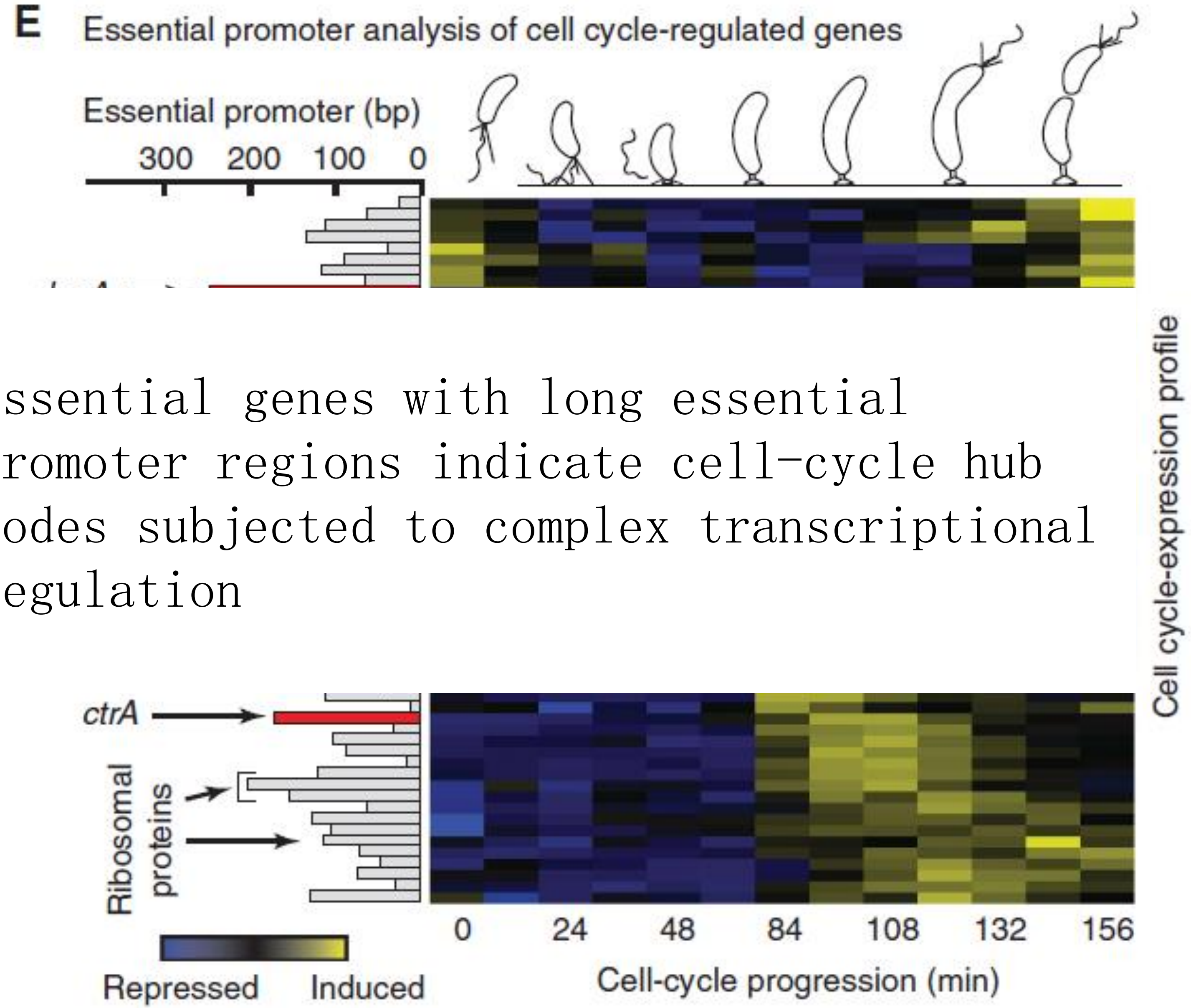


B

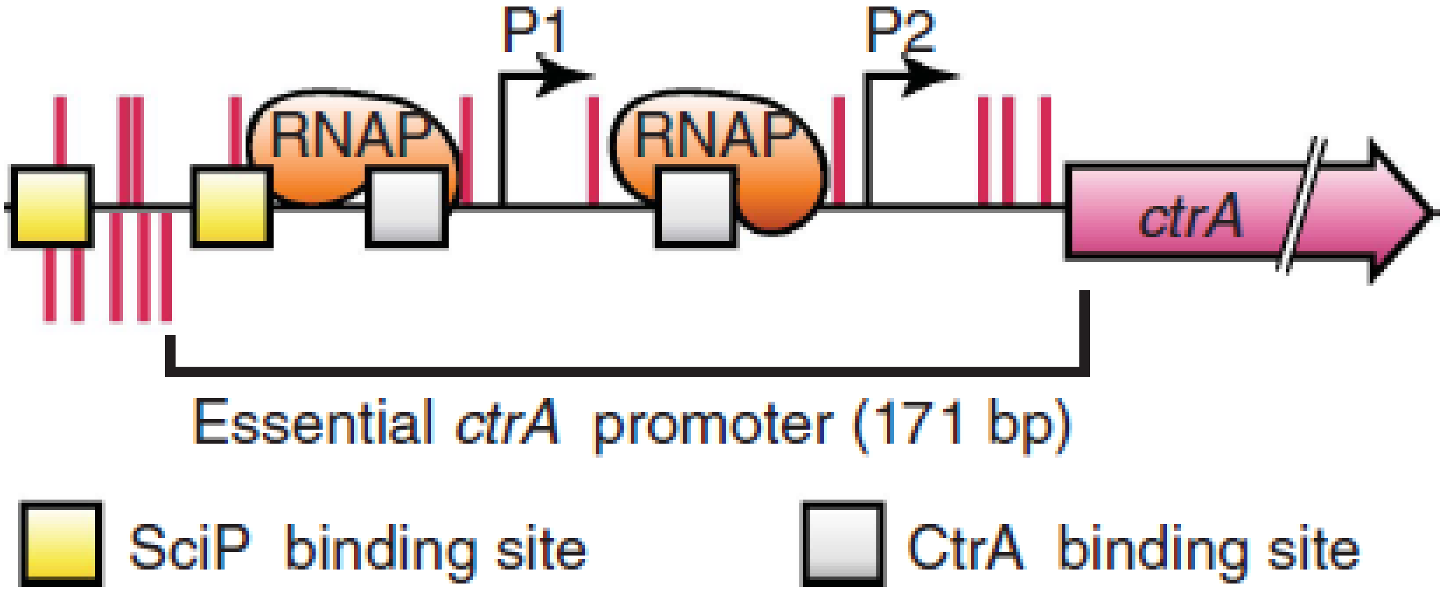
Identification of an essential operon transcript

D
Number of genes
80
60
40
20
0
0
100
200
gcrA
ctrA
dnaA
Length of essential promoter region (bp)

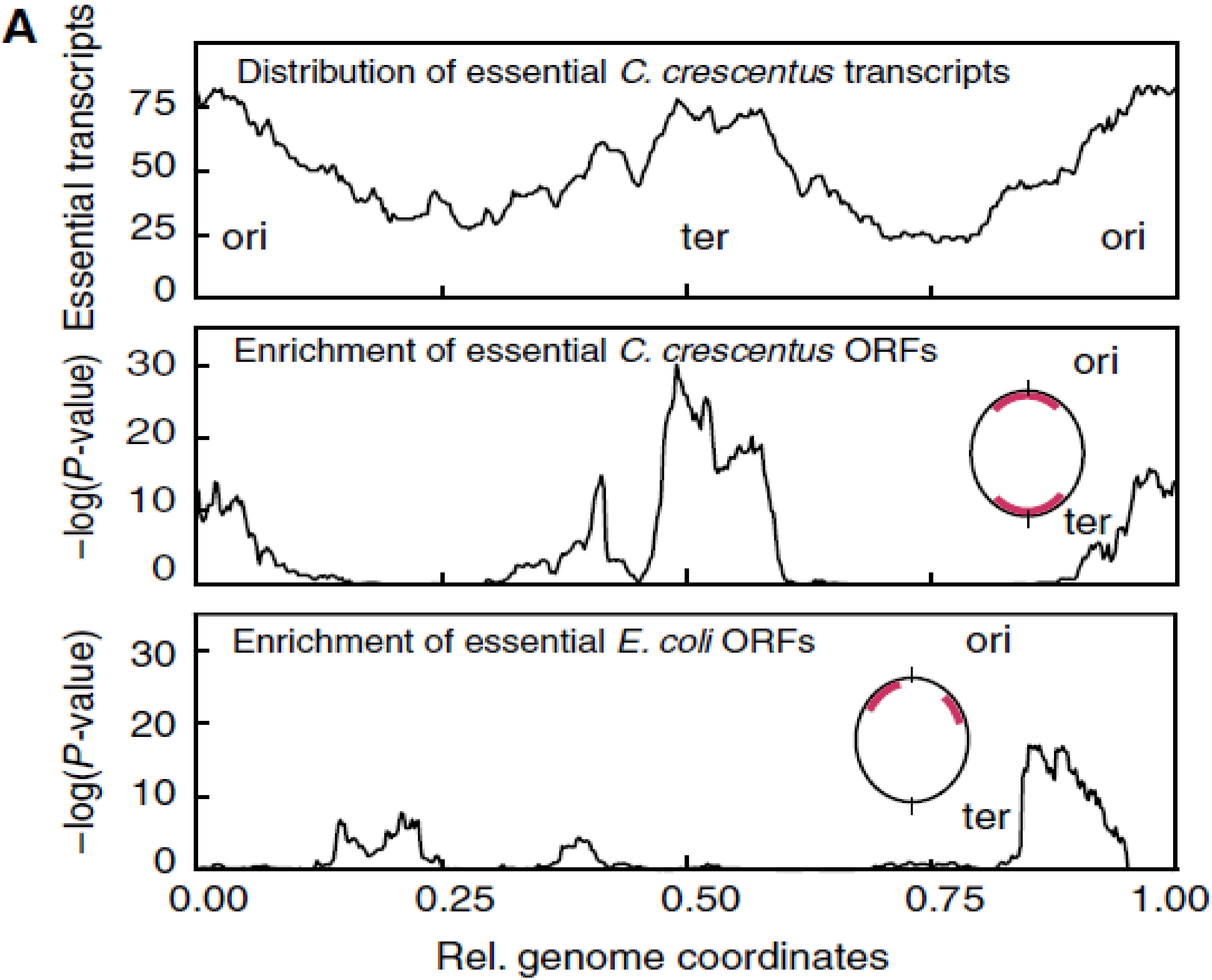
CELL CYCLE-REGULATED ESSENTIAL GENES



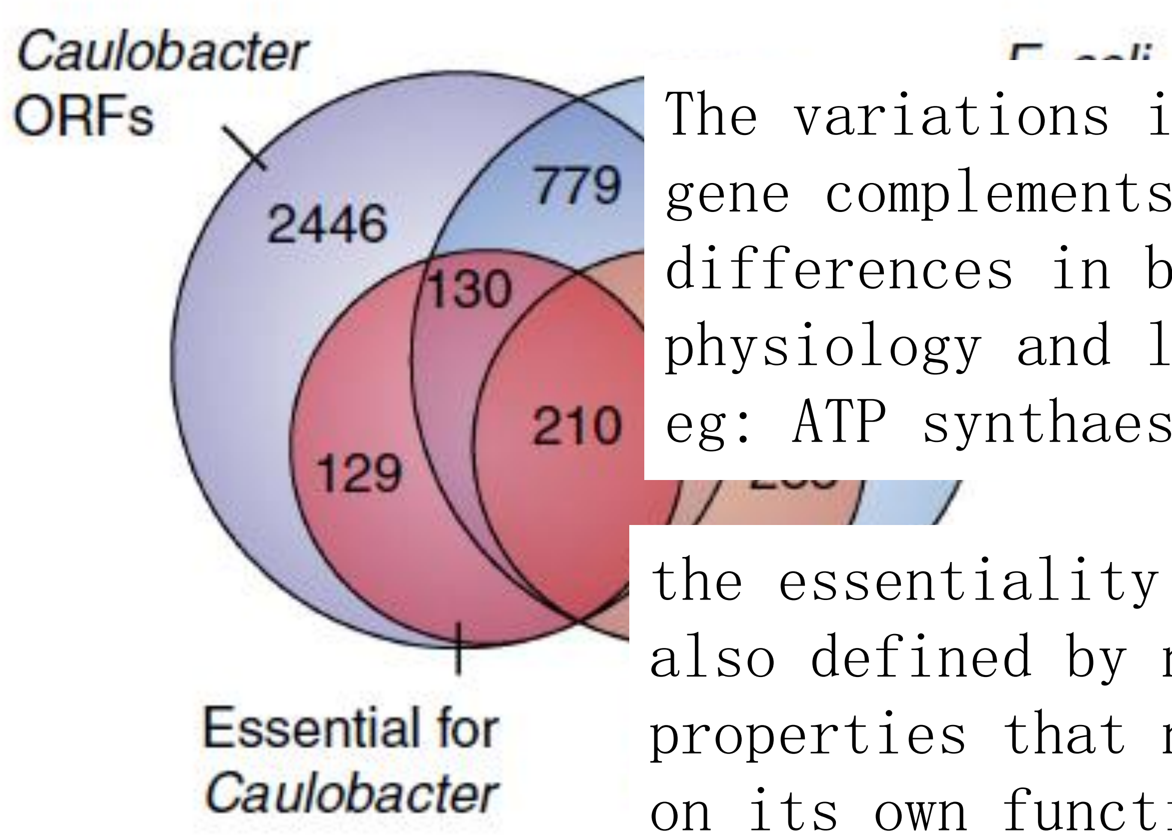
Essential genes with long essential promoter regions indicate cell-cycle hub nodes subjected to complex transcriptional regulation



F Multiple TSS within essential promoter regions


A
Distribution of essential C. crescentus transcripts
Essential transcripts
75
50
25
0
ori
ter
ori
Enrichment of essential C. crescentus ORFs
−log(P-value)
30
20
10
0
ori
ter
Enrichment of essential E. coli ORFs
−log(P-value)
30
20
10
0
ori
ter
0.00
0.25
0.50
0.75
1.00
Rel. genome coordinates

WHAT GENES CONSTITUTE THE MINIMUM SET REQUIRED FOR PROKARYOTIC LIFE?



The variations in essential gene complements relate to differences in bacterial physiology and life style.
eg: ATP synthaes

the essentiality of a gene is also defined by non-local properties that not only depend on its own function but also on the functions of all other essential elements in the genome

- The strategy described here provides a direct experimental approach that, because of its simplicity and general applicability, can be used to quickly determine the essential genome for a large class of bacterial species.

- The identification of all essential DNA elements is essential for a complete understanding of the regulatory networks that run a bacterial cell.

Thanks for your attention~