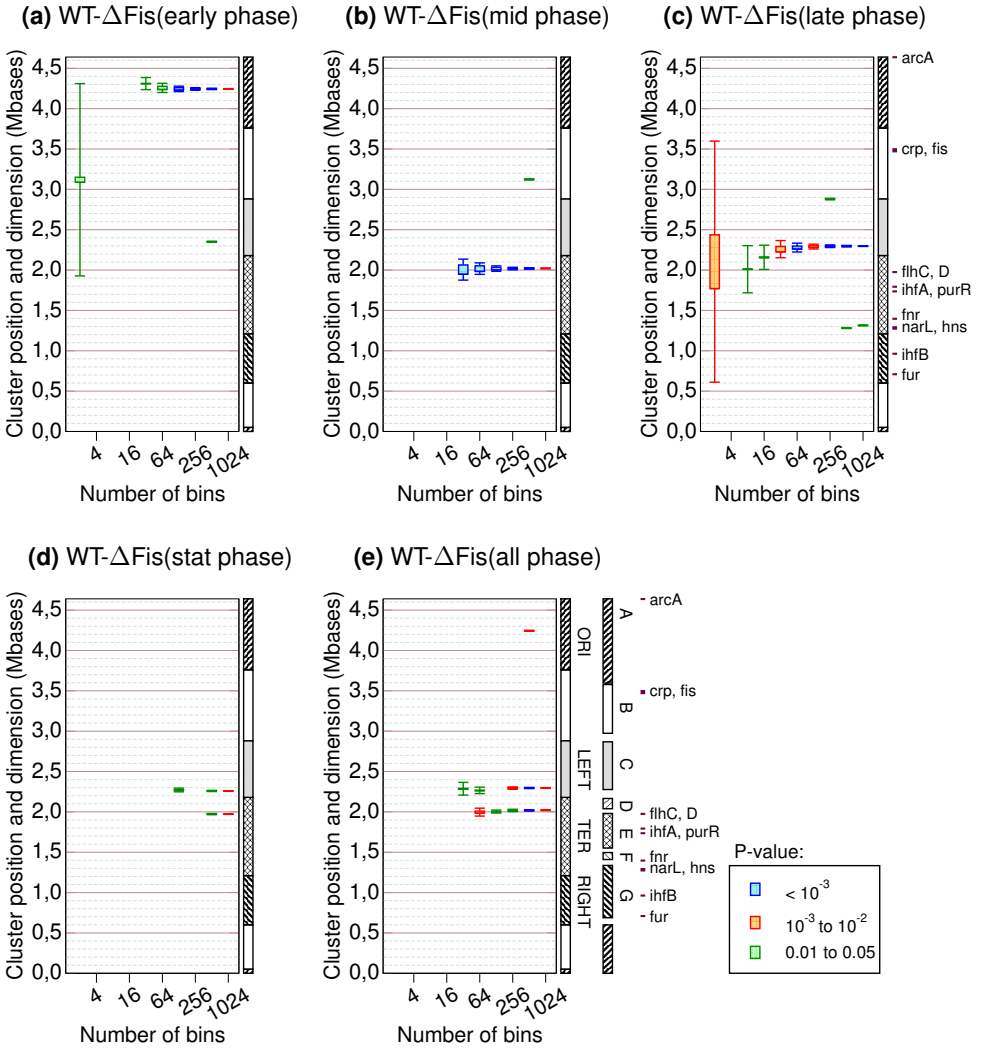
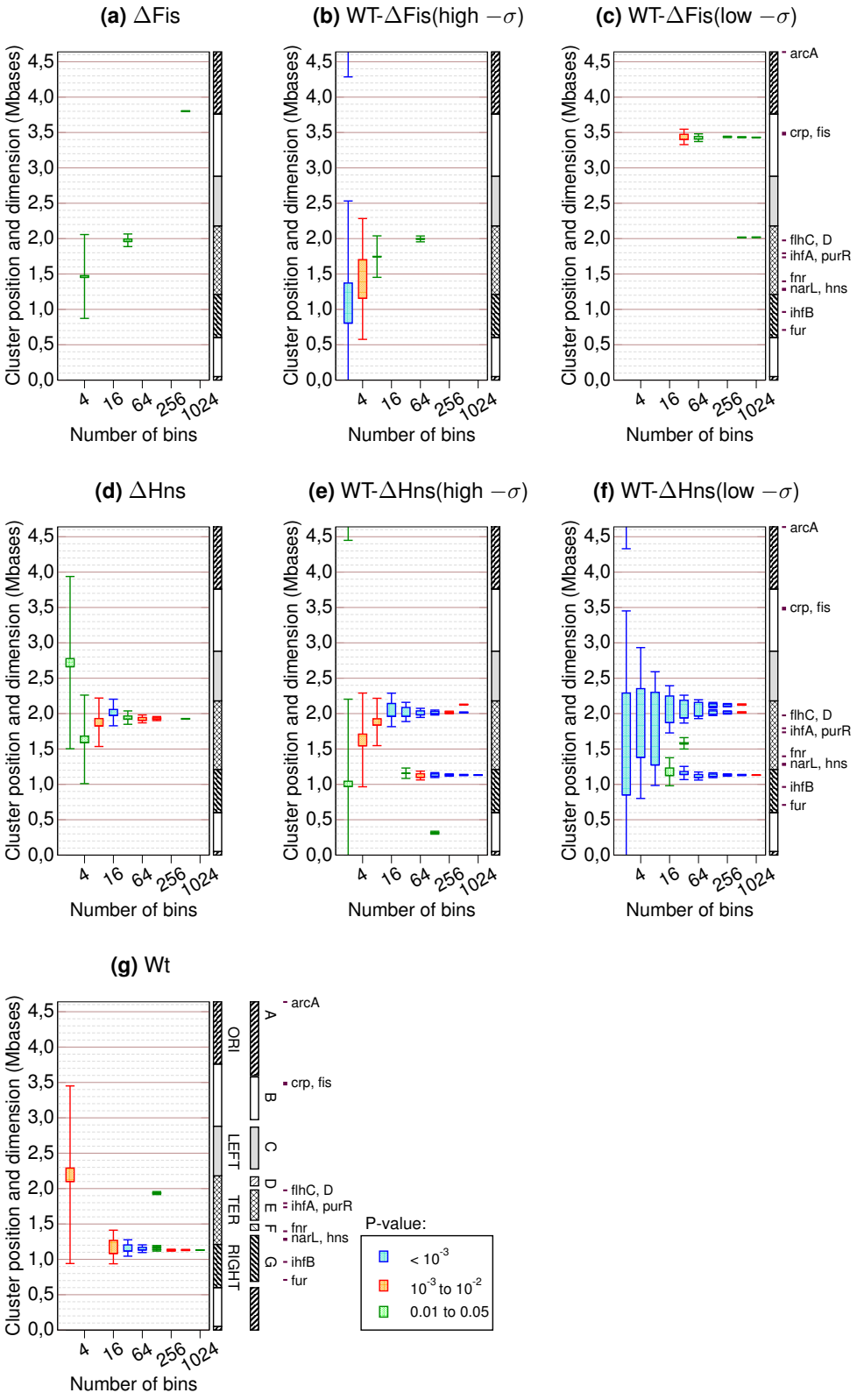
Supplementary results for Scolari *et al*

Clusters from nucleoid perturbation / transcriptomics experiments

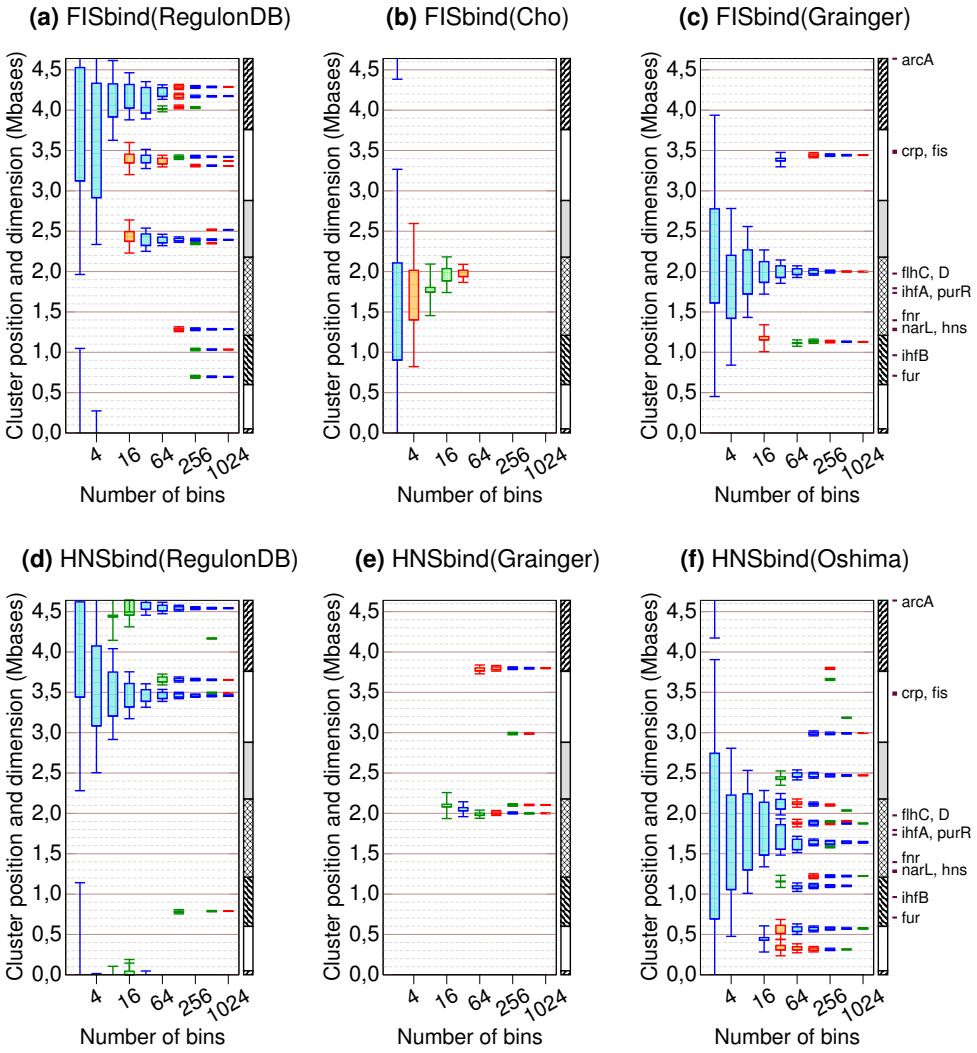


Supplementary Figure S2 Cluster diagrams for the transcription microarray Fis deletion data from ref.[32]. Different panels refer to different growth phases (early, mid-, late-exponential and stationary, in rich media), while the last panel refers to the union of all the growth phases.



Supplementary Figure S1 Cluster diagrams for the transcription microarray nucleoid perturbation data from ref. [30,31].

Clusters from protein binding data:



Supplementary Figure S3 Cluster diagrams of Fis and H-NS binding from the RegulonDB [8] database and Grainger, Cho and Oshima ChIP-chip experiments [27,28,57]. The binding sites of the Grainger data-set for the Fis and H-NS experiments show the same clusters as transcriptomics data on genes responding to supercoiling changes after H-NS deletion in microarray data from Marr *et al.* [30,31]. A hypergeometric test was performed to test the overlaps of the Grainger and Cho ChIP-chip targets with the RegulonDB database (Table SM1)

.

Summary of clusters from transcriptomics and protein binding data:

From transcriptomics:

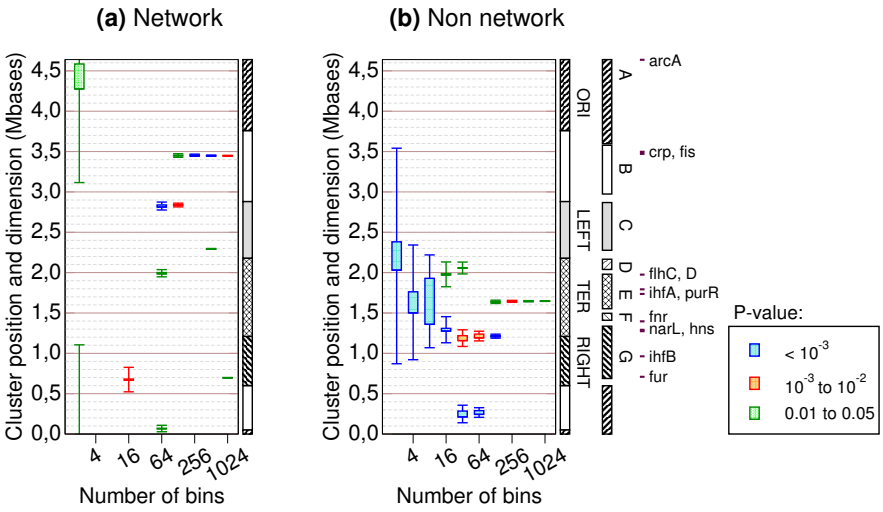
list	clust1	clust2	clust3	clust4	clust5	clust6	clust7
WT	0.03			< 0.001			
ΔFis	0.05					0.04	
ΔH-NS	0.01						
WT-ΔFis(low)	0.04	0.04			0.01		
WT-ΔFis(high)	0.05	0.06					
WT-ΔH-NS(low)	< 0.001	< 0.001	< 0.001	< 0.001			
WT-ΔH-NS(high)	< 0.001	< 0.001	0.01	< 0.001			

From protein binding data:

list	clust1	clust2	clust3	clust4	clust5	clust6	clust7
GraingerFis	< 0.001	< 0.001		< 0.001	< 0.001		
GraingerHns	< 0.001	< 0.001	0.01			< 0.001	0.01

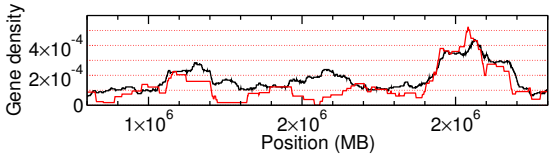
Supplementary Table S1 Summary of the most significant clusters found at all scales and their *P*-values. The coordinates (in bp) along the *E. Coli* genome are: **Cluster 1** 1929600-2195230 **Cluster 2** 1993030-2037780 **Cluster 3** 2096110-2141990 **Cluster 4** 1094210-1163310 **Cluster 5** 3428200-3447460 **Cluster 6** 3782180-3815590 **Cluster 7** 2981340-2996630. Note: **Cluster 2** and **Cluster 3** are included in **Cluster 1**.

Clusters of genes inside and outside the known transcription regulatory network:

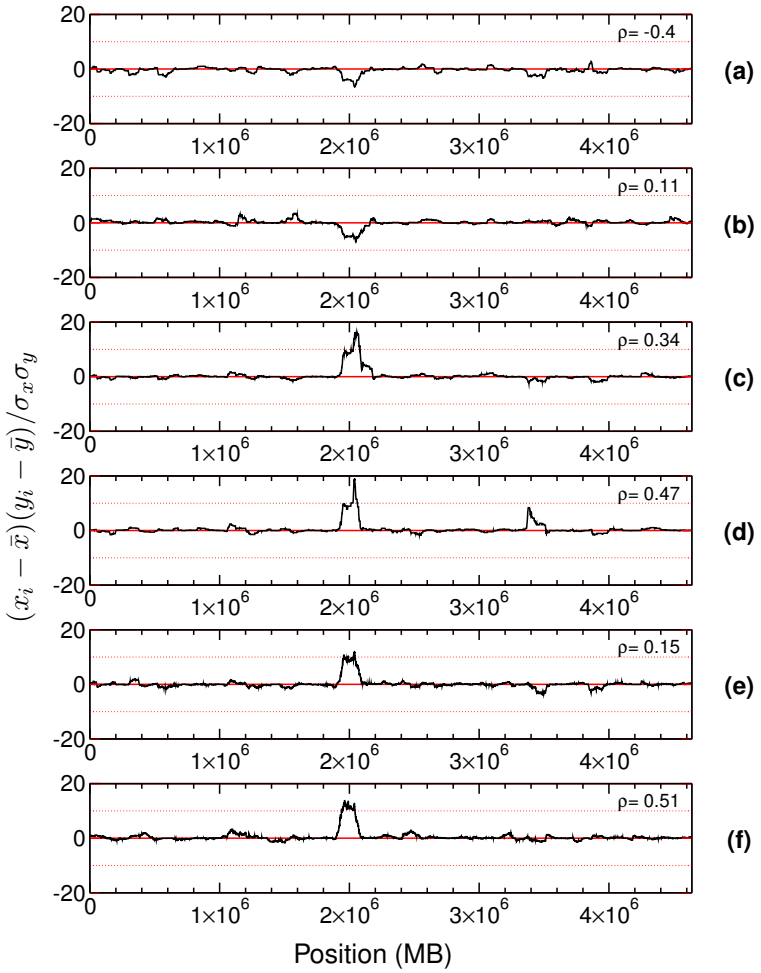


Supplementary Figure S4 Cluster diagram of the genes in the RegulonDB network and outside RegulonDB network (see ref.[36]).

Histograms of EPODs and comparison with clusters:



Supplementary Figure S5 Linear density of heEPODs along the genome (red line) compared to the density of nucleoid-perturbation sensitive genes WT-ΔH-NS(low $-\sigma$) (black line, bin size $L/32$). The x-axis spans the Ter macrodomain. Note the highly correlated regions at the border of the Ter (at $1 \cdot 10^6$ and $2 \cdot 10^6$ bases) macrodomain in accordance to the results of Figure S6.



Supplementary Figure S6 The black lines are the local contributions to the Pearson correlation coefficient along the genome coordinate (x-axis) between the normalized linear densities of he- and tsEPODS (ref. [29]) and the densities of nucleoid-perturbation sensitive genes. The densities where calculated using a sliding window of size $L/32$. **(a)** Correlation between tsEPOD and heEPOD density. **(b)** Correlation between tsEPOD density and WT-ΔH-NS(low $-\sigma$). **(c)** Correlation between heEPOD and WT-ΔH-NS(low $-\sigma$). **(d)** Correlation between heEPOD density and FISbinding(Grainger). **(e)** Correlation between heEPOD and FISbinding(Cho). **(f)** Correlation between FISbinding(Grainger) and FISbinding(Cho).

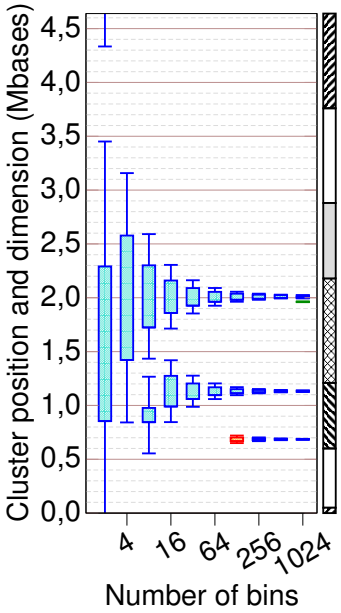
Number of genes of heEPOD in common with:			
List	Experimental value	Mean value	*P*-value
WT	19	16.58	0.300
ΔFIS	11	12.98	0.226
ΔHNS	15	9.06	0.034
WT-ΔFIS($\sigma\downarrow$)	22	13.29	0.004
WT-ΔHNS($\sigma\downarrow$)	31	15.64	< 0.002
WT-ΔFIS($\sigma\uparrow$)	18	8.69	< 0.002
WT-ΔHNS($\sigma\uparrow$)	30	11.52	< 0.002

Supplementary Table S2 Summary of the number of genes within heEPODs (strictly included) in common with the genes significantly responding to Fis and H-NS deletion, and changes in supercoiling[30].

Number of genes of tsEPOD in common with:			
List	Experimental value	Mean value	*P*-value
WT	5	6.16	0.224
ΔFIS	7	5.43	0.298
ΔHNS	6	5.65	0.496
WT-ΔFIS($\sigma\downarrow$)	5	4.68	0.496
WT-ΔHNS($\sigma\downarrow$)	13	5.70	0.002
WT-ΔFIS($\sigma\uparrow$)	4	3.20	0.402
WT-ΔHNS($\sigma\uparrow$)	13	4.35	< 0.002

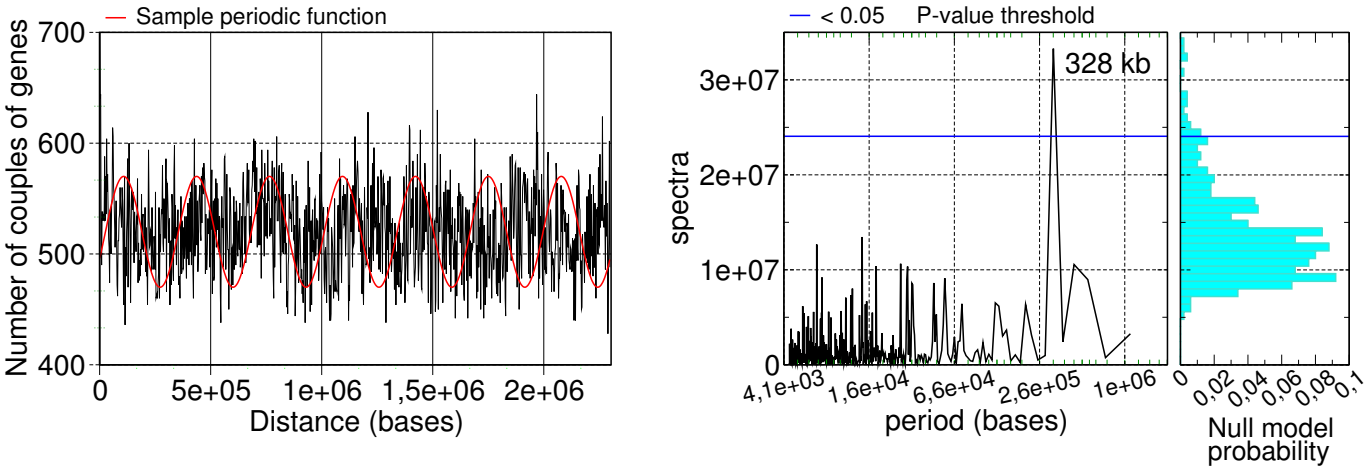
Supplementary Table S3 Summary of the number of genes within tsEPODs (strictly included) in common with the genes significantly responding to Fis and H-NS deletion, and changes in supercoiling[30].

Clusters of genes controlled by FlhC:

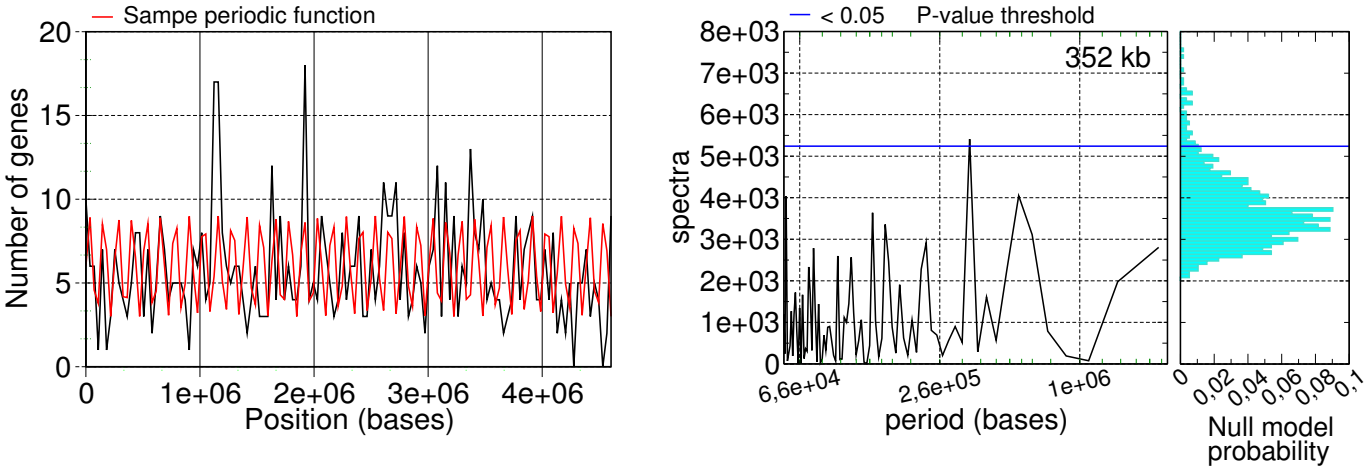


Supplementary Figure S7 Cluster diagram of the genes controlled by the FlhC transcription factor (data from RegulonDB). FlhC is a transcriptional activator that controls the operons related to assembling of the flagella. The main regions controlled by FlhC overlaps with the clusters at the border of the Ter macrodomain identified in both transcriptomics and binding sites lists. A second cluster is present in correspondence with the border between the Right macrodomain and the Right non structured zone.

Periodicity analysis:



Supplementary Figure S8 Procedure followed to detect periodicities in the gene lists. Left panel: distribution of the distance of genes in an experimental list (sliding window of bin-size $L/2048$), compared to a periodic function. Right panel: discrete Fourier transform of the distance distribution. The peaks correspond to contributions of a periodic function of a given period, reported on the x-axis. The figure shows the spectra of the intra-strain WT list (see Methods), where the peak indicates a signal for a periodicity of 328Kb. The comparison of this signal with the distribution of the maximum of the spectra found in randomized lists gives a significance score for this periodicity. The same procedure can be applied also to the sliding-window density histogram at a given bin size.



Supplementary Figure S9 Periodicity analysis for the case of a density histogram. The analysis (and the example dataset from the intra-strain WT list) coincides with that presented in Figure S8, except that the left panel is a sliding-window histogram (bin-size $L/256$) of the genes in the list, rather than a distance distribution.

list	360 ± 36 Kb	624 ± 36Kb	101 ± 36Kb	20 ± 36Kb
WT	AB			
ΔFis			AB	
ΔH-NS				B
WT-ΔFis(low)				
WT-ΔFis(high)				
WT-ΔH-NS(low)	AB	AB	B	
WT-ΔH-NS(high)	AB	AB	B	B

Supplementary Table S4 Table summarizing the significant periodicities found ($P < 0.05$). In the table the letter A indicates a significant periodicity found in the density histogram while B indicates a periodicity found in the distance pair distribution. Genes sensitive to supercoiling variation in the intra-strain WT list show a compatible periodicity of 360 ± 36Kb, this periodicity is found also in the WT-ΔH-NS lists at all supercoiling conditions. Upon Fis deletion, supercoiling sensitive genes lose the 360 ± 36 Kb periodicity, but show a new periodicity at 101 ± 36Kb, again found also in the WT-ΔH-NS lists in all supercoiling conditions. Finally, in H-NS deletion mutants, supercoiling sensitive genes lose the 360 ± 36 Kb periodicity but show a periodicity at 20 ± 36Kb also found in the inter-strain WT-ΔH-NS data in high negative supercoiling conditions. The compatibility condition of 36Kb was selected as twice the bin-size of the density distribution histogram.

Clusters and the flagellar/biofilm synthesis pathway:

Cluster 4 1094210-1163310

wt-fis low σ		wt-fis high σ		wt-hns low σ		wt-hns high σ	
rep	act	rep	act	rep	act	rep	act
csgA	bssS	flgA	bssS	csgD	flgA	bssS	flgA
ghrA	flgJ	flgE	flgH	csgE	flgB	csgA	flgB
grxB	flgM	ghrA	mdoC	yceI	flgC	csgB	flgC
hinT	flgN	grxB	plsX	yceJ	flgD	csgD	flgD
mdtH	plsX	hinT	yceK		flgE	csgE	flgE
rimJ		rimJ			flgF	csgG	flgF
ycdX		ycdX			flgG	mdoC	flgG
ycdY		ycdY			flgH	mdoG	flgH
ymdB		ycdZ			flgI	rluC	flgI
		ycfL			flgJ	ycdU	flgJ
					flgM	yceI	flgK
					flgN	yceJ	flgM
					ghrA		flgN
					mdtG		holB
					yceG		yceN
					ycfH		

wt 41-54		Hns 41-54		Fis 41-54	
hyp	rel	hyp	rel	hyp	rel
csgF	flgA	csgD		csgC	flgF
pabC	flgB	csgE		csgD	flgJ
pyrC	flgC	csgG		mdoC	flgN
rimJ	flgD	grxB		plsX	ycfH
rluC	flgE	pabC		rimJ	
ycdZ	flgF	pyrC		yceA	
yceG	flgG	rimJ		yceH	
yceH	flgH	ycfM		yceM	
yceO	flgI			ymdB	
	flgJ				
	flgM				

Cluster 1 1929600-2195230

wt-fis low σ		wt-fis high σ		wt-hns low σ		wt-hns high σ	
rep	act	rep	act	rep	act	rep	act
baeS	amyA	amn	cld	cpsB	dcm	araG	cheY
cbl	cheB	aspS	fliG	gatD	flhB	araH	dcyD
cpsB	cheZ	eda	fliJ	gatY	fliA	cpsB	flhA
dcd	flhB	fbaB	fliL	gatZ	fliC	cpsG	flhB
erfK	fliE	fliC	fliM	glf	fliD	fcl	flhE
fbaB	fliF	flu	fliN	gmd	fliE	ftnA	fliA
ftnA	fliG	ftnA	fliP	gmm	fliF	ftnB	fliC
hisI	fliL	gatC	fliQ	ogrK	fliG	gatA	fliD
ruvA	fliM	gatY	fliR	rcnR	fliH	gatC	fliE
wcaM	fliP	gmd	galF	rcsA	fliJ	gatD	fliF
yecM	fliQ	gnd	metG	thiD	fliK	glf	fliG
yedA	fliZ	hchA	rfbB	ugd	fliL	gmd	fliH
yegD	motA	hisC	znuA	uvrC	fliM	gmm	fliJ
	yecR	hisD	znuC	uvrY	fliN	gnd	fliK
	yedW	hisF		wcaA	fliQ	hchA	fliL
	yeeV	hisH		wcaC	fliS	intG	fliM
	znuA	hisI		wcaD	fliZ	mdtD	fliN
		motB		wcaE	hisH	otsA	fliO
		ogrK		wcaF	motB	rcsA	fliQ
		pykA		wcaM	tar	rfbA	fliR
		ruvA		wzc	yecR	rfbB	fliS
		tar		yedV	yeeE	rfbC	fliY
		torZ		yeeN		rfbX	fliZ
		wbbK		yegJ		rfc	motB
		wcaI		zinT		ruvA	yecR
		wcaM				shiA	yedK
		wzxC				torY	yedM
		yedF				torZ	
		yedI				ugd	
		yegD				uvrC	
		yegP				vsr	
		yegS				wbbI	
		yegV				wbbJ	
						wcaA	
						wcaB	
						wcaC	
						wcaD	
						wcaE	
						wcaF	
						wcaI	
						wcaJ	
						wcaK	
						wcaL	
						wcaM	
						wza	
						wzb	
						wzc	
						wzxC	
						yecD	
						yecT	
						yedV	
						yedW	
						yeeE	
						yeeN	
						yeeW	
						yegI	
						yegJ	
						yegR	
						yegS	
						yegW	
						yegZ	
						yodB	
						zinT	
						znuB	

wt 41-54		Hns 41-54		Fis 41-54	
hyp	rel	hyp	rel	hyp	rel
cmoA	fliJ	amn	cpsB	cmoB	fliI
cmoB	yeeN	cmoA	cpsG	cobT	fliK
cobS		dcm	rfc	dcm	fliR
flhD		flhD	tap	flhC	rcsA
hisC		fliY	wcaB	flhD	wbbI
hisD		gatC	wcaC	fliC	yeeN
insF-5		gatD	wcaE	fliD	yeeR
lpxM		gatY	wcaF	fliS	yeeT
rfbB		gatZ	wcaJ	gatY	yegS
yebK		hisC	wcaK	insA-5	
yecD		hisD	wcaL	lpxM	
yecN		hisF	wcaM	mtfA	
yedI		lpxM	wzc	pgsA	
yedR		motB	wzxC	torZ	
yeeE		sdiA	yeeN	vsr	
yeeP		shiA	yeeU	yebK	
yeeZ		torY	yegS	yecD	
yegD		torZ	yehD	yecN	
yegW		uvrY	znuB	yedA	
yegX		vsr		znuA	
znuA		yebK			
		yecN			
		yedI			
		yedL			
		yedW			
		yegI			
		znuA			

Color code legend

Flagella

O-antigen

M-antigen capsid display/biofilm

Regulator of biofilm formation

Curli

transcription factor

Supplementary Table S5 Intersection between genes found in the data sets from the transcription microarray experiments of Blot et al[30] and clusters of genes identified in this analysis. The colors indicate the gene ontology class. Genes from intra-strain experiment have been divided into rel and hyp columns corresponding to gene transcripts whose expression is associated with relaxation (rel) or high negative supercoiling (hyp). The labels act and rep indicate activation and repression in inter-strain profiles.

Flagellar gene expression cascade (by order of expression)

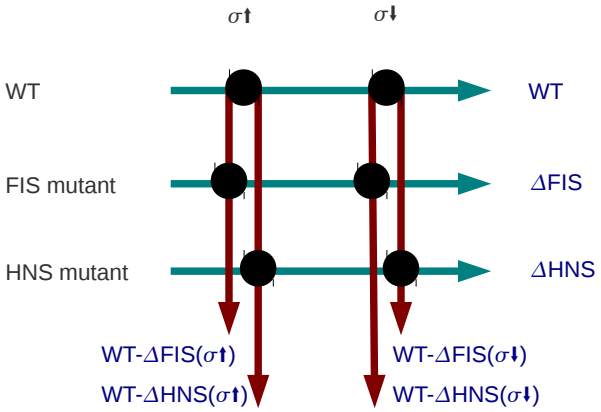
Class	Gene	Cluster	Pos Regulator	Neg Regulator	Hns CC	predicted Hns site	Function
I	flhCD	1	Crp, Hns, s70, s54, s28	Fur, OmpR, RcsAB, IHF, LrhA			Master transcirptional regulator
II	fliLMNOPQR	2	s70 FlhCD				Basal body hook
II	fliE	2	s70 FlhCD				
II	fliFGHIJK	2	s70 FlhCD		Y	Y	
II	flgA	4	s70 FlhCD				
II	flgBCDEFGHIJ	4	s70 FlhCD			Y	
II	flhBAE	1	s70 FlhCD				
II	fliA (s28)	2	s28, s70 FlhCD	NsrR	Y	Y	Sigma 28
II	fliZ		s28, s70 FlhCD		Y	Y	Inhibitor of curli
II	fliY		ss				
III	flgKL	4	s28, s70 FlhCD				Hook
III	fliDST	2	s28, s70 FlhCD		Y		Hook
IIIb	fliC	2			Y	Y	Filament
II III	flgMN	4	s28, s70 FlhCD		Y	Y	Anti-sigma 28
IIIb	motAB	1	s28				motive force chemosensor
IIIb	cheAW	1	s28				
IIIb	tar	1	s28				
IIIb	tap	1	s28				
IIIb	cheRBYZ	1	s28				

Genes involved in biofilm formation (in order of chromosome position)

MD	Gene	Cluster	Regulates	Regulated by	Hns CC	predicted Hns site	Process
right	*yadCKLM*			Hns, Crp			criptic fimbriae
	sfmACDHF			Hns, Crp			criptic fimbriae
	ycbQRSTUVS			Hns, Crp	Y	Y	criptic fimbriae
	pgaABCD	~ 4		CsrA, NhaR			PGA, adhesin
	csgD	4	csg operon, bcs operon	OmpR, CpxR, Hns, IHF, RstA, Fis, ss, YdaM		Y	curli synthesis
	bssS	4					biofilm
ter	*ycgv*				Y	Y	criptic adhesin
	ydaM						curli inhibitor
	tqsA						AI-2 transport, quorum sensing
	uvrY/uvrC	1	csrB	LexA, SdiA		Y	biofilm
	sdiA	2	uvrY, fliE				biofilm, cell division
	rcsA	2	wca, fliPQR et al				osmolarity, membrane perturbations
	yedQ	2*	c-di-GMP	ss			curli regulation
	yeej	1*					adhesin
	flu	1		OxyR, Dam			antigen 43
	rfb operon	3			Y	Y	O-antigen
	wca operon	3		RcsC/S/B/A		Y	colanic acid, M-antigen
left	*yegE*	1*	c-di-GMP	ss			curli regulation
	yehABCD	1		Hns, Crp	Y		criptic fimbriae
	yfaI						adhesin
	yfcOPQRSTUV						criptic fimbriae
	yfjR						
	ypja						adhesin
	gutq						
	ttda						
ori	*yraHIJK*			Hns, Crp			criptic fimbriae
	bcsABZC, EFG			CsgD			cellulose
	cysE						
	waaG	6		Hns ?	Y	Y	LipoPolySaccaride synthesis
	tnaA					Y	
	rfah						transcriptional antiterminator
	yihr						
	cpxAR		motAB, cheAW, mdtA(1)				envelope stress
	yjbe					Y	
	fumb						
	fimAICDFGH			Hns ?		Y	fimbriae
	yjip						

Supplementary Table S6 Genes involved in flagellar expression and in biofilm formation, data from refs [46,63,74–80]. Flagellar and chemotaxis genes are ordered according to the sequence of expression. Biofilm genes are ordered according to their position on the chromosome, the first column (MD) show the macrodomain in which the gene is located on the chromosome. The known factors regulating gene expression are indicated, as well as the targets of the transcription factors in the list, the abbreviation ss stands for sigma s. The presence of H-NS binding sites in the promoter region is shown in different columns whether it was determined by ChIP-chip [27] (Hns CC) or by prediction from bioinformatic sequence analysis [33] (Hns site).

Supplementary methods for Scolari *et al*



Supplementary Methods Figure SM1 Summary of microarray data of Fis, H-NS, and supercoil sensitive genes.

(a)

	RegulonDB	Cho *et al*	Grainger *et al*
RegulonDB	200	58 ($P = 10^{-4}$)	11 ($P = 0.26$)
Cho		894	71 ($P = 10^{-9}$)
Grainger			220

(b)

	RegulonDB	Grainger *et al*
RegulonDB	71	1 ($P = 0.44$)
Grainger		96

Supplementary Methods Table SM1 The table compares reported Fis **(a)** and H-NS **(b)** binding sites from different data sources. The table reports the number of genes in the overlap between the lists, and the *P*-value in parentheses.

(a)

Macro-	**Ori**		**Right**		**Ter**		**Left**	
domain:	start	end	start	end	start	end	start	end
Position (*Mb*)	3.76	0.05	0.60	1.21	1.21	2.18	2.18	2.88

(b)

Chromosomal	**A**		**G**		**F**		**E**		**D**		**C**		**B**	
sector:	start	end	start	end	start	end	start	end	start	end	start	end	start	end
Position (*Mb*)	3.59	0.59	0.68	1.33	1.40	1.49	1.54	1.97	2.03	2.16	2.27	2.86	2.97	3.57

Supplementary Methods Table SM2 (a) Start and end positions in *Mb* of the macrodomains defined by Boccard and coworkers (ref. [10,13]), and **(b)** chromosomal sectors defined by Mathelier and Carbone (ref. [37])