MixNet results on E. Coli Regulatory network

F. Picard, V. Miele, J.-J. Daudin, S. Robin

June 27, 2008

Transcriptional regulatory networks constitute one important instance of biological networks that are studied from the structural point of view. Indeed, they were shown to share some important properties, such as a relative sparseness, a very low number of feed back circuits, and a hierarchical organization [2], and Grouping operons based on their connectivity structure appears essential to understand the wiring diagram of such complex networks. In this paragraph, we consider the transcriptional regulation network of E. Coli as published by [3]. Nodes of the network correspond to operons, and two operons are linked in the network if one operon encodes a transcription factor that directly regulated another operon. We considered the latest version of the network (version 1.1 [1]), and we consider the connex component of the network only.

Summarizing regulatory structure. The clustering results with 5 classes (given by the ICL criterion) gives a rough picture of the functioning of the network: the majority of operons are regulated by very few nodes. Figure 1 displays the whole network, each color corresponding to a Mixnet class. Another representation of the results is presented in Figure 2, which gives the parameters of the Mixnet model (probabilities of connection between-within groups, and α, the proportion of the groups). We also propose a graphical representation of matrix π on the same Figure.

At this resolution level, the network is summarized into regulated operons (groups 1 and 4), which receive edges only. These two groups are distinguished based on their regulatory elements: operons of group 4 are regulated by `crp` only (which makes its own group), whereas operons of group 1 are regulated by many cross-talking elements (group 2, 3, and 5).

Meta Motifs of regulation. It has been show that some motifs like the popular Feed forward loop constituted the core structures of E. Coli regulatory network [3]. When looking at Figure 2, it appears that Mixnet exhibits the same global structures at the group level. For instance, when considering relations between group 5 and 4, this structure is known as single input module : one TF regulating other operons that do not communicate ($\hat{\pi}_{4,4} < 1\%$). Another example of meta motif is the "meta" Feed-Forward loop for groups 2-3-1 and 2-5-1. In both cases the effector group is group 1, and groups 2 and 3 can be viewed as information relays. This structure is in coherence with known biological information. In Table 2 we provide a list of Transcriptional factors which are known to be global [1]. We can see that 60% of nodes of group 2 are known global TF (9/17), which also means that the other nodes of this group (`fliAZY`, `FruR`,`himA`,`nlpD_rpoS`,`rpoE_rseABC`, `rpoH`,`rpoN`) behave like global TF from the connectivity

---

[1] `http://www.weizmann.ac.il/mcb/UriAlon/Network_motifs_in_coli/ColiNet-1.1/`

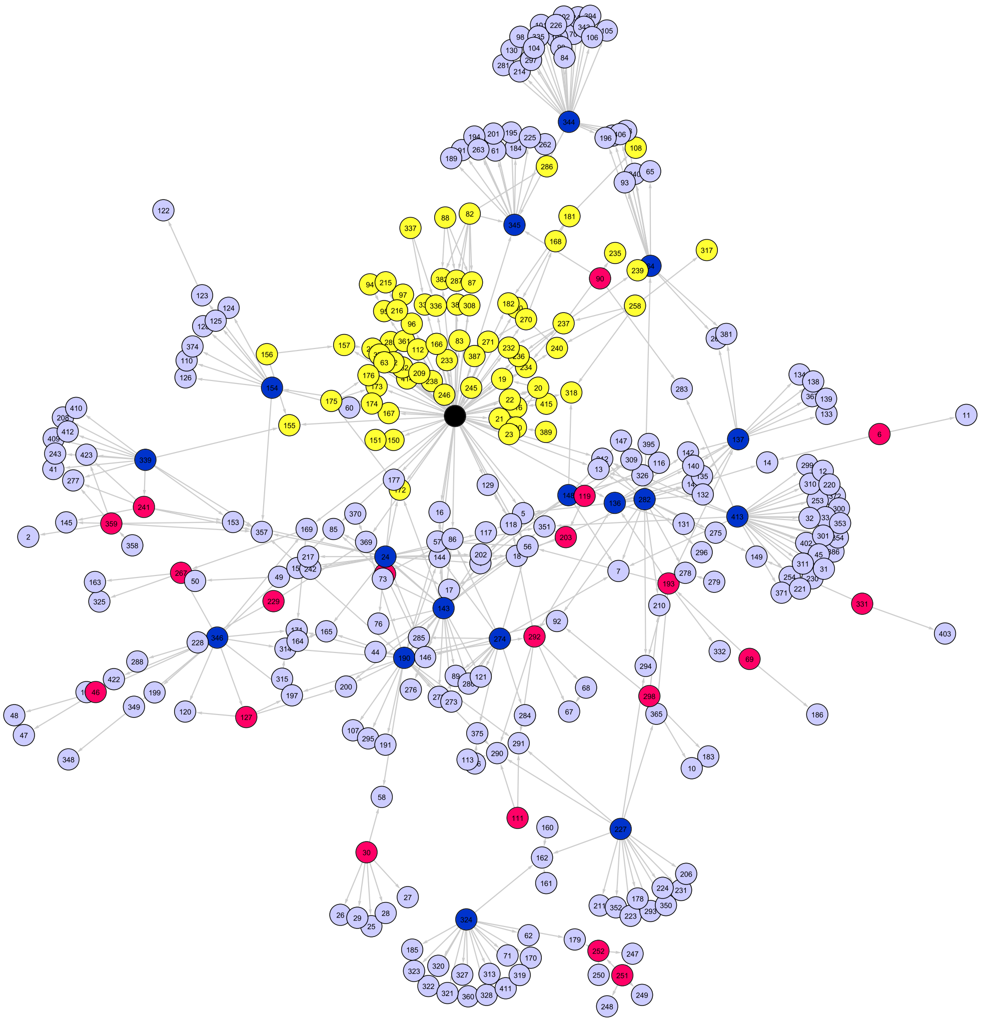

Figure 1: Regulatory network of E. Coli displayed with colors for each MixNet class (5 groups).

	1	2	3	4	5
1	.	.	.	.	.
2	6.40	1.50	1.34	.	.
3	1.21	.	.	.	.
4	.	.	.	.	.
5	8.64	17.65	.	72.87	11.01
alpha	65.49	5.18	7.92	21.10	0.30

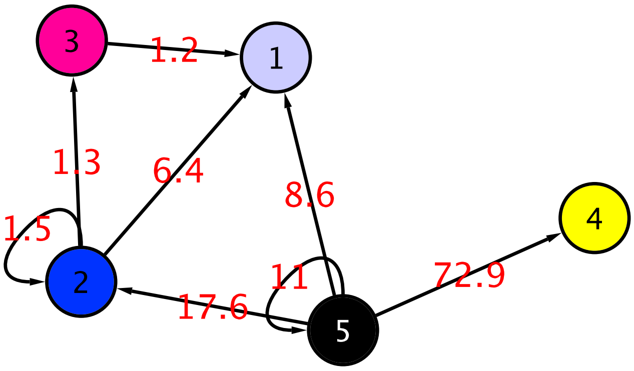

Figure 2: MixNet parameters for the EColi regulatory network network with $Q = 5$ classes. Connections (×100) lower than 1% are not represented.

	1	2	3	4	5
1	22	0	0	0	0
2	0	1	0	0	0
3	112	0	12	1	0
4	0	8	8	1	0
5	55	0	0	0	0
6	0	5	0	0	0
7	0	1	0	0	0
8	23	0	0	1	0
9	0	2	0	0	0
10	7	0	0	50	0
11	0	0	0	18	0
12	0	0	0	0	1

Table 1: confusion matrix between clusters with Q=5 and Q=12 groups for E Coli network.

point of view. Even if this is not surprising for `rpo` operons involved in the σ factor of the RNA polymerase, the question can be asked for `fliAZY`, `FruR` and for `himA`.

No Community structure in TRN. The global connectivity structure of the network can be seen on matrix π (Figure 2) which summarizes the intensity of connection between and within groups. This matrix shows empty rows, which corresponds to groups that are made of strictly regulated operons (nodes that receive edges only). When studying the structure of this matrix, we can see that there is no community structure, *i.e.* there is no group which is heavily intra-connected and poorly inter-connected (the diagonal elements of π are small). This is an interesting feature, since it means that in this regulatory network, nodes do not share modularity patterns. This observation is coherent with the structure of regulatory circuits which form cascades of regulations without feedback.

Getting a more detailed picture. It is obvious that summarizing a network with only 5 classes only give a crude view of the structure of the data, this is why it may be more relevant to study other configurations with a higher number of groups. If the number of groups is selected using the adaptive strategy, 12 groups are selected. Looking at empty rows of matrix π, we observe that groups 1,3,5,8,10 are strict regulated operons (with a 1% threshold), which corresponds to the biggest clusters. Then regulators are distinguished based on their connectivity patterns and on their targets. For instance operon `yhdG_fis` (group 2) regulates operons of groups 1 and 8, operons of group 9 (`fnr`,`narL`) regulate operons of group 8.

Overall, regulator operons are clustered into groups with a very small number of nodes, which emphasizes the hierarchical structure of the regulation wiring diagram. Interestingly, operon `rpoE_rseABC` forms a group on its own (group 7), which regulates operons of group 5 and 6 (`cpxAR`,`flhDC`,`fliAZY`, `fur`,`rpoH`), whereas it has not been referenced as a global TF. This result is not surprising, as `rpo` operons are involved in the σ unit of the RNA polymerase.

Meta motifs are also present in this representation: a Meta Feed Forward Loop is to be found with group 5-6-7. Single Input Modules (SIM) are present with 12-10, 12-11, 12-8, 2-8

	1	2	3	4	5	6	7	8	9	10	11	12
1	.	.	.	.	.	.	.	.	.	.	.	.
2	100.00	3.98	.	.	.	.	.	8.64	.	3.14	.	.
3	.	.	.	.	.	.	.	.	.	.	.	.
4	.	.	5.86	.	.	5.70	.	4.65	.	1.35	.	.
5	.	.	.	.	.	.	.	.	.	.	.	.
6	.	.	.	.	15.54	5.00	.	.	.	.	.	.
7	.	.	.	.	42.22	20.00	3.38	.	.	.	.	.
8	.	.	.	.	.	.	.	.	.	.	.	.
9	.	.	.	2.85	.	.	.	60.33	.	.	.	.
10	.	.	.	.	.	.	.	.	.	.	.	.
11	.	.	.	.	.	.	.	.	.	3.25	.	.
12	.	.	1.87	.	.	60.00	.	33.24	.	89.78	40.87	11.01
alpha	6.66	0.30	37.10	5.35	16.61	1.52	0.30	8.59	0.61	16.84	5.81	0.30
nb. nodes	22	1	125	17	57	5	1	24	2	57	18	1

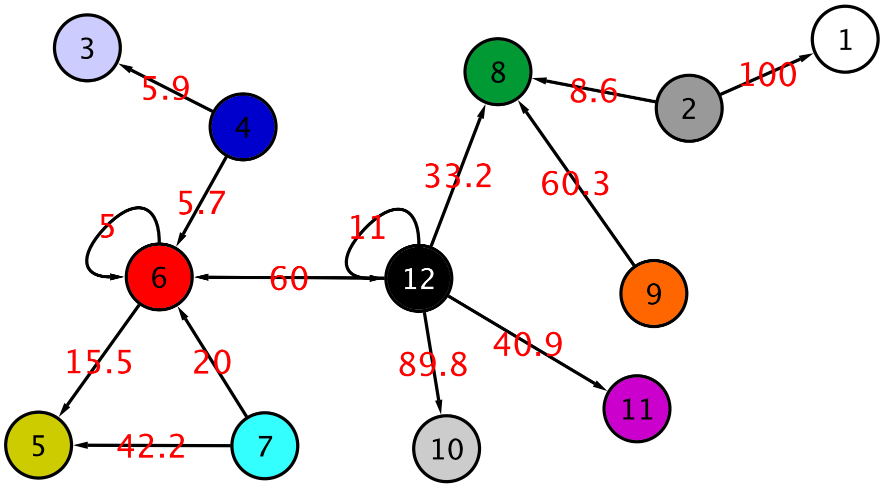

Figure 3: MixNet parameters for the EColi regulatory network network with $Q = 12$ classes. Connections ($\times 100$) lower than 1% are not represented in the table, and connections lower than 5% are not displayed in the summary graph.

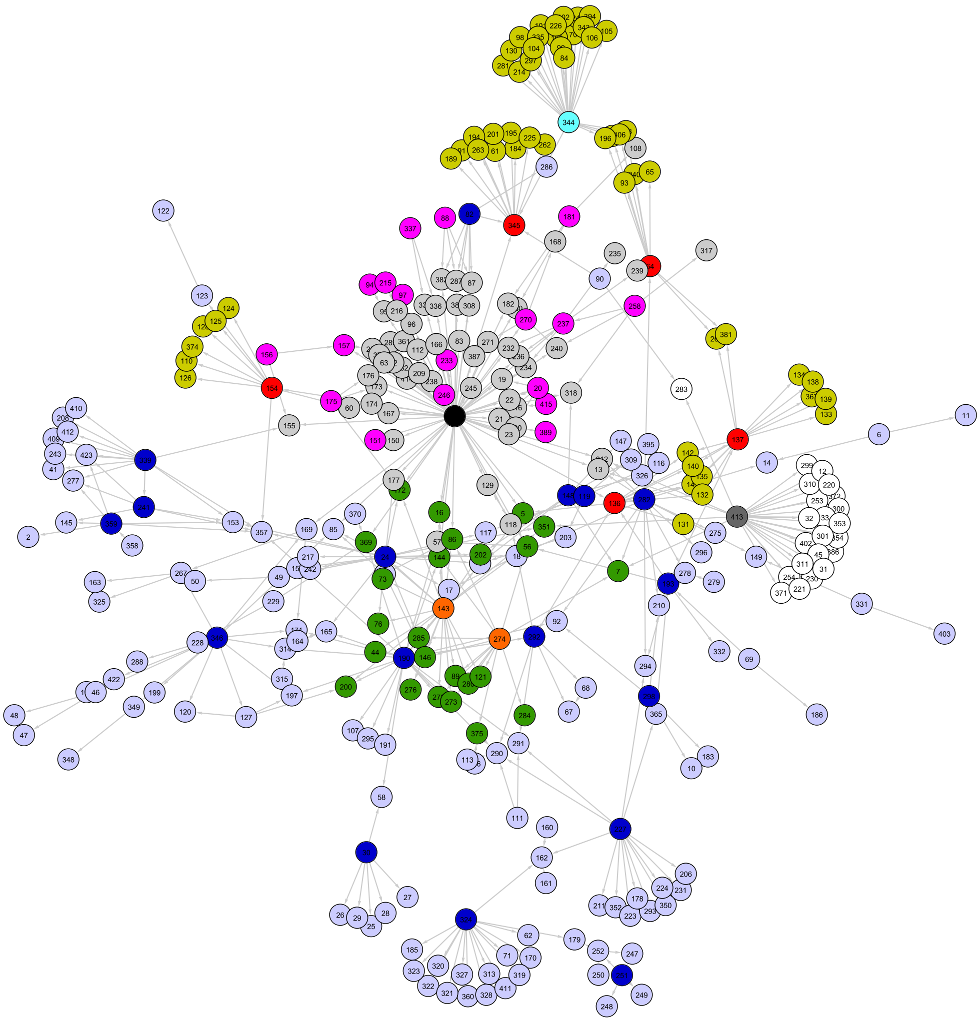

Figure 4: Regulatory network of E. Coli displayed with colors for each MixNet class (12 groups).

TF	Mixnet_Q5	Mixnet_Q12
rob	2	4
purR	2	4
ArcA	2	4
Lrp	2	4
FNR	2	9
NarL	2	9
Fur	2	6
CpxR (CpxAR)	2	6
FIS	2	2
CspA	3	3
soxRS	3	3
Hns	3	4
OmpR	3	4
Mlc	4	11
CRP	5	12
pdhR_aceEF_lpdA	1	1
flhBAE	1	5
flhDC	2	6
oxyR	3	4
DnaA	3	3
IHF	NID	NID
PhoB	NIC	NID
CysB	NIC	NIC
modE	NIC	NIC
BolA	NID	NID
IciA	NID	NID

Table 2: Global regulators from Martinez-Antonio et al. (2003). NID = not in dataset, NIC = not in connex component.

and 2-1. Their formation is due to groups 12 and 2 which are composed of one operon only (`crp` and `yhdG_fis` respectively). Another meta motif is the Dense Overlapping Regulon with groups 4-3. Group 4 is made of 7 operons which have been referenced as global operons , and two `rpo` operons (`nlpD_rpo`, `rpoN`, Table 2). It would be interesting to check that the other operons of this group can be considered as major regulatory actors. Even if they are not globalm regulators, Mixnet reveals that they act like global regulators from the topological point of view.

References

[1] Martinez-Antonio A. and Collado-Vides. Identifying global regulators in transcriptional regulatory networks in bacteria. *Current Opinion in Microbiology*, 2003(6):482–489, 2003.

[2] Balazsi G., Barabasi A.-L., and Oltvai Z.N. Topological units of environmental signal processign in the transcriptional network of escherichia coli. *PNAS*, 102(22):7841–7846, 2005.

operon	group_Q12	out-deg	in-deg
yhdG_fis	2	26	0
arcA	4	20	1
argR	4	6	0
cytR	4	7	0
fadR	4	5	0
FruR	4	7	0
himA	4	21	0
hns	4	7	1
lrp	4	14	0
marRAB	4	5	1
metJ	4	4	0
nlpD_rpoS	4	14	0
ompR_envZ	4	6	1
oxyR	4	4	0
purR	4	16	0
rob	4	12	0
rpoN	4	13	0
soxS	4	6	1
cpxAR	6	9	1
flhDC	6	7	3
fliAZY	6	12	2
fur	6	9	1
rpoH	6	10	4
rpoE_rseABC	7	24	0
fnr	9	22	0
narL	9	13	0
crp	12	72	0

Table 3: list of operons which correspond to regulated operons in the Coli regulation network with $Q = 12$ groups.

[3] Shen-Orr S., Milo R., Mangan S., and Alon U. Network motifs in the transcriptional regulation network of escherichia coli. *Nature genetics*, 31:64–68, 2002.