EvoEvo Deliverable 1.2

Analysis of robustness in TEV and E. coli strains

Due date:	M20 (July 2015)
Person in charge:	D. Schneider
Partner in charge:	UJF
Workpackage:	WP1 (Experimental observation of EvoEvo in action)
Deliverable description:	Analysis of robustness in TEV and *E. coli*: identification of molecular bases in both viral and bacterial models; end of identification of combinations strains-environments affecting bacterial robustness; relationship between gene order, genome architecture and bacterial physiology.

Revisions:

Revision no.	Revision description	Date	Person in charge
1.0	First version; TEV part	25/11/15	S. Elena (CSIC)
1.1	Add data for *E. coli* experiments	30/11/15	D. Schneider (UJF)
1.2	Corrections by G. Beslon	14/12/15	G. Beslon (INRIA)
1.3	Add Intro and Conclusions to TEV part	14/03/16	S. Elena (CSIC)
1.4	Corrections by O. Lamrabet	24/03/16	O. Lamrabet (UJF)
1.5	Corrections and validation	25/03/16	G. Beslon (INRIA)

Table of Contents

1. Introduction

Evidence has accumulated during recent years that organisms can maintain their performance in the face of a broad range of perturbations (de Visser *et al*. 2003; Wagner 2005). This includes the tolerance of proteins to amino acid replacements (Sinha & Nussinov 2001), the ability of genetic networks to withstand alterations (Aldana *et al*. 2007), the stability of cellular processes to stochastic variations of gene expression levels (Batada & Hurst 2007), or the resilience of embryonic development to environmental or genetic changes (von Dassow *et al*. 2000). In general, the term robustness is used to describe this behavior and genetic robustness or mutational robustness when mutations are the cause of perturbations. Many issues related to genetic robustness remain unsolved. For example, asserting the elevated robustness is a fundamental property of living organisms is problematic because we often ignore what normal robustness should be (Ciliberti *et al*. 2007). Still, we can try to identify the genetic and ecological factors associated with differences in robustness between species or genotypes (Krakauer & Plotkin 2002, Sanjuán & Elena 2006). Also, it remains unclear whether the evolutionary transition to a robust state occurs as a direct product of selection (Wagner *et al*. 1997, van Nimwegen *et al*. 1999, Proulx & Phillips 2005, Proulx *et al*. 2007) or merely as a byproduct of selection acting on correlated traits (Stearns *et al*. 1995, Ancel & Fontana 2000, Stearns 2002).

Task 1.1 addresses robustness at three different levels using the most appropriate experimental model. It is divided in three sections devoted to the study of robustness at the population level in the TEV model, at the regulatory level in the *E. coli* model, and at the genome level in both models.

2. Task 1.1 Robustness at the population, regulatory network and genome levels

2.1. Section 1: Robustness at the population level in the TEV experimental system

It has been postulated that natural selection may have shaped viral populations as a whole to be robust against the effect of mutations (Krakauer & Plotkin 2002). This type of robustness will pop up as an emerging property of viral populations, due to the inherent mutational coupling of quasispecies populations. In other words, it will be a population property rather than a characteristic of individuals. The theory predicts than such population robustness would evolve at large effective population sizes and high mutation rates, very much alike the conditions of RNA virus replication. To test this theory and its applicability to an RNA virus, we have performed the experiments described in Task 1.1 Section 1 of the proposal. In short, four different experimental treatments were designed (Elena *et al*. 2007): small effective population size and the standard mutation rate, small effective population size and increased mutation rate, large effective population size and standard mutation rate, and large effective population size and increased mutation rate. Small effective population sizes were obtained by isolating individual clones from the local lesion host *Chenopodium quinoa*; it is between 1 and 10 infectious units. Large population sizes were obtained by 1:10 dilution of the population resulting from an infected *Nicotiana tabacum* plant and it ranges between 10^4 – 10^5 infectious units. Mutation rate was artificially increased by submitting infected tobacco plants to one pulse of 10 min/d with a Phillips G36T8 germicidal lamp (maximal output at 253.7 nm; dose 2 J/cm^2) (Codoñer *et al*. 2006).

Twenty-five cycles of infection were performed for each treatment, 10 independent evolution lineages per treatment. At the end of the evolution experiments, the sensitivity of the evolved lineages to mutation was evaluated by means of three passages of random mutation-accumulation, as described in Sanjuán *et al*. (2007). If a viral population is robust against mutational effects, then its fitness would be poorly affected by the accumulation of a limited number of mutant individuals. By contrast, if the population is very sensitive to mutation, its fitness will be affected by a larger extent. Mathematically, these two propositions are equivalent to say that the slope of a log-linear regression of fitness on the number of bottleneck transfers will be different for each type of viral population (Elena *et al*. 2007): steeper slopes will be typical of sensitive populations whereas flatter slopes will be typical of more robust populations.

The slopes from the log-linear regressions estimated for each population were fitted to a GLM with a Normal distribution and identity link function, using “population size” and “mutation rate” as orthogonal factors and “lineage” as a factor nested within the interaction of “population size” and “mutation rate”. Population size had a significant effect of robustness, with populations evolved in the large effective population size regime being more robust on average than populations evolved in the small effective population size (P = 0.036). By contrast, mutation rate had no net significant effect on robustness (P = 0.366), although it has a highly significant effect in combination with the effective population size (interaction term, P < 0.001). Finally, lineages submitted to the same treatment were significantly heterogeneous in their fitness (P < 0.001). A graphical summary of these results can be found in Figure 1.

In conclusion, we found that the predictions of the Krakauer & Plotkin (2002) model for the evolution of population-level robustness were partially fulfilled for TEV: viral populations evolved under conditions of large population size become more robust than populations evolved at small effective population sizes. However, we failed to observe the predicted synergistic effect of mutation rate: large populations evolved at high mutation rates were not more robust than their counterparts evolved at standard mutation rate. One possible explanation for this discrepancy may be that TEV mutation rate is already large enough (Tromas & Elena 2010) and that the extra increase due to UVC resulted in genomes with a too high mutational load that were lethal.

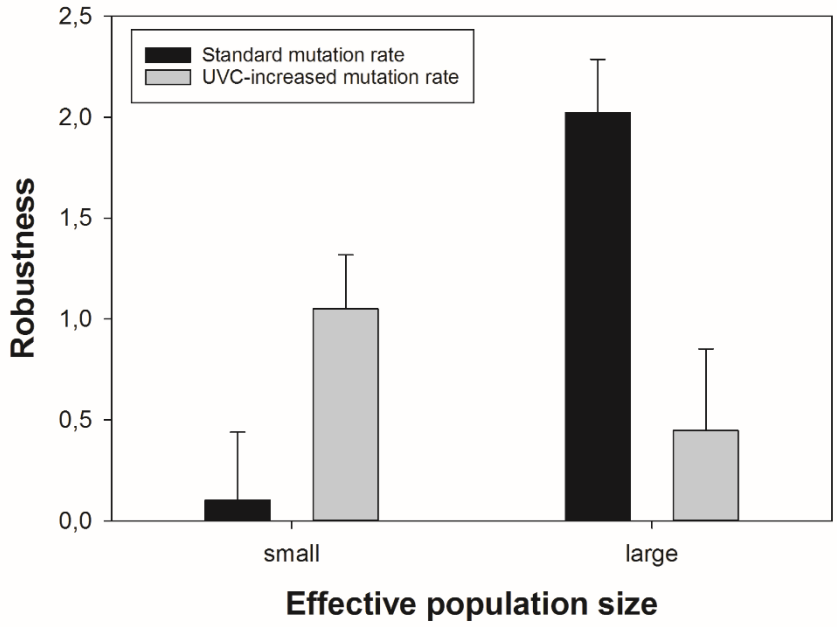

Figure 1 – Dependence of robustness on effective population size and mutation rate. Error bars represent ±1 SEM.

2.2. Section 2: Robustness at the regulatory network level in the *E. coli* experimental model

2.2.1. Introduction

Long-term evolution of *E. coli* in glucose minimal medium has been characterized by beneficial mutations in genes encoding global regulatory genes (Philippe *et al*., 2007) and changes in epistatic interactions between global regulators (Cooper *et al*., 2008). Therefore, long-term adaptation in this environment was achieved by substantial rewiring of global regulatory networks. We investigate whether and how these newly organized networks affected bacterial physiology in alternative environments, and therefore whether long-term adaptation resulted in changes in the robustness at the level of regulatory networks.

CRP is a key hub in the *E. coli* transcriptional network, involved in more than 200 regulatory interactions (Gosset et al., 2004; Zheng et al., 2004).The CRP-controlled regulon has been shown to be increasingly important during evolution in the LTEE. Deletions of the gene encoding *crp* have been introduced in the LTEE ancestor and in two independent evolved clones, one sampled from each of two of the twelve evolving populations after 20,000 generations (Cooper *et al*., 2008). Deleting *crp* had a much more dramatic effect on the growth in the evolution environment and on the global transcription profile of the two evolved clones than on the ancestor. Because the sequence of the *crp* gene was unchanged during evolution, these differences indicated epistatic interactions between *crp* and mutations at other loci that accumulated during evolution (Cooper *et al*., 2008). Therefore, epistasis has been important in the adaptive evolution of these bacterial populations, and they provided new insight into the types of genetic changes through which epistasis can evolve. Indeed, we identified a number of regulatory genes (*spoT*, *fis*) harbouring beneficial mutations that accounted for these changes in epistatic interactions with the CRP regulon. We address whether these changes in the interactions between global regulators within the regulatory network affected the robustness of the evolved clones.

This deliverable included the construction of the *crp* deletions in evolved strains from all twelve populations (D1.1, due and finished at M12, not detailed anymore here), the investigation of the effect of these deletions in different environments (D1.2, due at M20), and the identification of the mutations that interfere with the *crp* deletion during the long-term evolution experiment (D1.2, due at M20). Therefore, we can investigate the impact of rewiring regulatory network on robustness.

2.2.2. Growth traits in various environments

The effect of the *crp* deletion on growth was investigated in all populations by direct plating on rich medium plates. In these conditions, deleting *crp* had a drastic effect on growth in all evolved clones. Indeed, all deleted strains produced small colonies, indicative of growth defects. We investigated growth defects in more details in two populations, called Ara-1 and Ara+1, by growing strains in liquid media, including the minimal glucose medium used in the evolution experiment and alternative environments.

Deleting *crp* in evolved clones sampled after 20,000 generations from each of the Ara-1 and Ara+1 populations severely affected the growth rates in minimal glucose medium DM1000 (Figure 2). These growth defects were much more severe in the evolved genetic backgrounds than in the ancestor, implying that mutations substituted during evolution in both populations interacted with the *crp* deletion. To investigate when these epistatic interactions occurred during evolution, the *crp* deletion was constructed in evolved clones sampled earlier during the evolution of the two populations. The results shown in Figure 2 revealed that the severe growth defects were already detected after 1500 and 2000 generations in Ara+1 and Ara-1, respectively. Therefore, in both populations, mutations that were substituted early during evolution, before 2000 generations, interacted epistatically with the *crp* deletion. These results revealed a high dynamics of the CRP-controlled regulon, and therefore of the regulatory network, during long-term evolution.

To extend these analyses to other environments, we measured the growth abilities of parental and *crp*-deleted strains using GN2 Biolog plates (AWELinternational, BLAIN, France) that contain 95 different carbon sources (Figure 3). We measured the effect of the *crp* deletion on the growth in these alternative environments in the ancestor and two evolved clones sampled after 2000 and 20,000 generations from population Ara-1. Deleting *crp* drastically reduced the catabolic breadth of each of the three strains (Table 1). However, this effect was similar in all three genetic backgrounds, ancestral and evolved. Therefore, at least for population Ara-1, the *crp* deletion affected growth more severely in the evolved clones than in the ancestor only in the minimal glucose medium in which the evolution occurred. This suggests that evolution in the glucose environment strongly selected a particular structure of the regulatory network and that disturbing this structure (here by deleting *crp*) results in lower robustness of the evolved clones specifically in that environment.

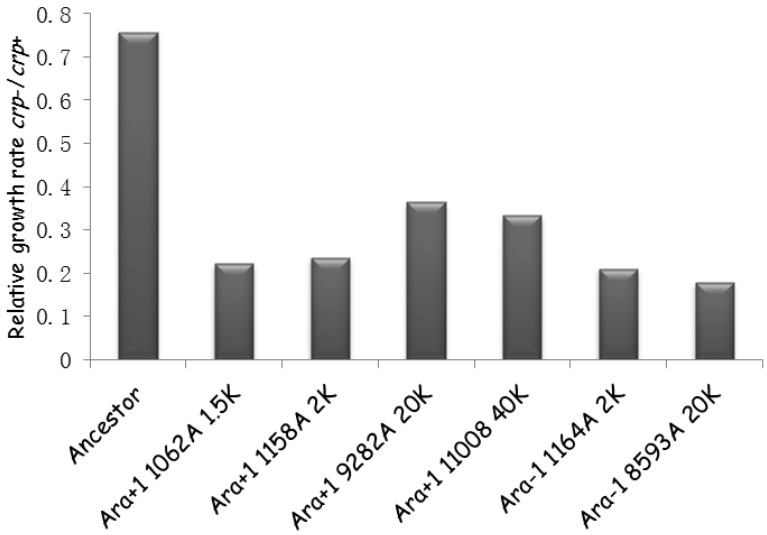

Figure 2 – Effect of the *crp* deletion on the growth abilities in evolved clones from populations Ara-1 and Ara+1 in DM1000 medium (K: x1000 generations). The relative growth rate of each *crp*-deleted strain compared to the corresponding parental crp^+ strain is given.

A1 eau	A2 α-cyclodextrine	A3 dextrine	A4 glycogène	A5 Tween 40	A6 Tween 80	A7 N-acétyl-D-galactosamine	A8 N-acétyl-D-glucosamine	A9 adonitol	A10 L-arabinose	A11 D-arabitol	A12 cellobiose
B1 i-érythritol	B2 D-fructose	B3 L-fucose	B4 D-galactose	B5 gentiobiose	B6 α-D-glucose	B7 m-inositol	B8 α-lactose	B9 α-D-lactose lactulose	B10 maltose	B11 D-mannitol	B12 D-mannose
C1 D-melibiose	C2 B-méthyl D-glucoside	C3 psicose	C4 D-raffinose	C5 L-rhamnose	C6 D-sorbitol	C7 sucrose	C8 D-tréhalose	C9 turanose	C10 xylitol	C11 méthyl pyruvate	C12 mono-méthyl succinate
D1 acide acétique	D2 acide cis-aconitique	D3 acide citrique	D4 acide formique	D5 acide D-galactonique lactone	D6 acide D-galacturonique	D7 acide D-gluconique	D8 acide D-glucosaminique	D9 acide D-glucuronique	D10 acide α-hydroxy butyrique	D11 acide β-hydroxy butyrique	D12 acide γ-hydroxy butyrique
E1 acide p-hydroxy phényl-acétique	E2 acide itaconique	E3 acide α-kéto butyrique	E4 acide α-kéto glutarique	E5 acide α-kéto valérique	E6 acide D,L-lactique	E7 acide malonique	E8 acide propionique	E9 acide quinique	E10 acide D-saccharique	E11 acide sebacique	E12 acide succinique
F1 acide bromo succinique	F2 acide succinamique	F3 glucuronamide	F4 alaninamide	F5 D-alanine	F6 L-alanine	F7 L-alanyl-glycine	F8 L-asparagine	F9 acide L-aspartique	F10 acide L-glutamique	F11 acide glycyl-L-aspartique	F12 acide glycyl-L-glutamique
G1 L-histidine	G2 hydroxy L-proline	G3 L-leucine	G4 L-ornithine	G5 L-phényl alanine	G6 L-proline	G7 acide L-pyro glutamique	G8 D-serine	G9 L-serine	G10 L-thréonine	G11 D,L-camitine	G12 acide γ-amino butyrique
H1 acide urocanique	H2 inosine	H3 uridine	H4 thymidine	H5 phényl éthylamine	H6 putrescine	H7 2-amino éthanol	H8 2,3-butanédiol	H9 glycérol	H10 D,L-α-glycérol phosphate	H11 glucose-1-phosphate	H12 glucose-6-phosphate

Figure 3 – Representation of a GN2 Biolog plate containing 95 carbon sources.

Table 1 – Growth abilities of bacterial strains with (*crp*$^+$) and without (*crp*$^-$) *crp* in alternative environments. The number of environments allowing growth of each strain is given.

		Ancestor	**Ara-1**	
		606	**2000**	**20,000**
Biolog plate (95 carbon sources)	***crp*$^+$**	41	45	48
	***crp*$^-$**	11	14	15

2.2.3. Identification of mutations that interfere with the *crp* deletion

We showed above that mutations that were substituted before 2000 generations during evolution interacted epistatically with the *crp* deletion in both populations Ara+1 and Ara-1. Indeed, deleting *crp* in evolved clones sampled after 2000 generations had a much more drastic effect on growth rates and global transcription profiles than in the ancestor.

As we have the genome sequences of the evolved clones, all beneficial mutations substituted during the first 2000 generations are known. By moving each of these mutations from population Ara-1 with the *crp* deletion in the ancestor, we were able to identify which one interfered with the *crp* deletion. We found that the *crp* deletion had a more drastic effect both on growth rates and global transcription profiles when the level of DNA supercoiling was modified by one of the beneficial mutations. These results are currently being written as a manuscript.

2.3. Section 3: Robustness at the genome level in both TEV and *E. coli* experimental models

2.3.1. Robustness at the genome level in the TEV experimental model

A different way of attaining mutational robustness is by evolving genomic architectures that are robust against the perturbation effect of mutations, including large genomes with duplicated genes (Krakauer & Plotkin 2002). Such redundancy-based robustness will operate at the individual level, in contraposition to what was described in Task 1.1. Section 1 above. The mutational robustness of the 11 TEV strains engineered for Deliverable 1.1 was evaluated using the experimental protocol also described in Deliverable 1.1. In short, we have developed a method based in chemical mutagenesis with HNO_2 and evaluation of the infectivity of mutagenized genomes. The logic of this assay is as follows: the more robust a genotype, the less affected would it be by the treatment with HNO_2. Infected tissue was collected and analyzed after 5, 6 and 7 days post-inoculation (dpi). As an example, Figure 4 summarizes the results of the method for the wildtype TEV and four of the engineered strains (TEV-alkB, TEV-2NIb2, TEV-2b and TEV-eGFP). TEV-alkB and TEV-2b genomes encode for additional functional genes: the alkB domain involved in removing alkylation damage from RNA and the 2b suppressor of RNA silencing from *Cucumber mosaic virus*. These two genes clearly provide a fitness benefit for the virus, either in terms of reduced mutations or in terms of interfering with the host defenses. In addition, carrying the 2b gene adds functional redundancy (*i.e.*, a second suppressor of RNA silencing in addition to the normal one HC-Pro) without adding genetic redundancy. TEV-eGFP encodes for an additional gene, the eGFP which does not provide any fitness benefit to the virus. Finally, TEV-2NIb2 carries a duplication of the viral replicase gene NIb cloned in the second proteolytic position of the genome. This duplication generates both genetic and functional redundancy. In all cases, mutagenic treatments lapsed between 1 and 3.5 h, and the infectivity of each virus was evaluated by inoculating batches of 4-week old *N. tabacum* plants. Control infections with non-mutagenized viruses were done in all cases (solid dots). The difference between the areas under the control and treatment curves was taken as a measure of robustness and plotted in Figure 5. Statistically significant differences in robustness exists between genomic architectures, being the wildtype virus the most robust one and TEV-alkB the less robust one (1-way ANOVA, $P < 0.01$).

From these experiments we conclude that the wildtype genome architecture of TEV is more robust to mutational effects than any engineered architecture we created.

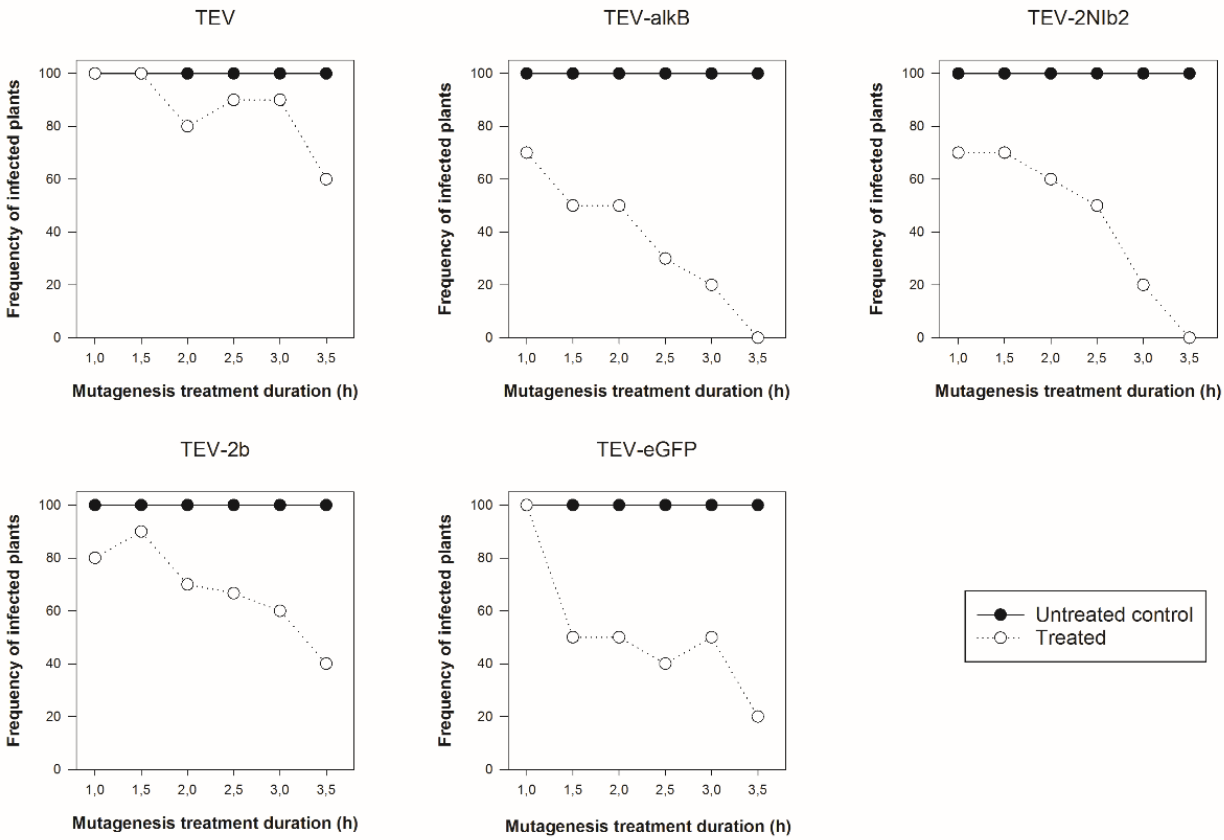

Figure 4 – Response curves to mutagenic treatments on some of the different genomic organizations of TEV. Mutagenic conditions with HNO_2 consisted in pH = 5.4 and 26 ºC for the number of hours indicated in the abscissa of each graph.

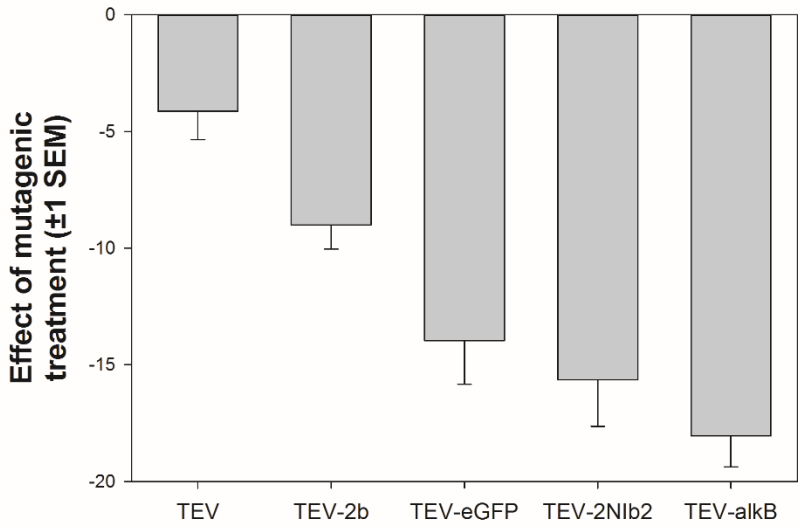

Figure 5 – Effect of mutagenesis treatment. The smaller the effect, the more robust the viral genomic architecture.

2.3.2. Robustness at the genome level in the *E. coli* experimental model

This deliverable included the characterization of all large chromosomal rearrangements in the 12 populations of the LTEE and the investigation of the effect of duplications and inversions on fitness and global transcription profiles (D1.2, due at M20 and finished at M12).

We found more than 100 large chromosomal rearrangements (deletions, duplications, amplifications, inversions) in all 12 populations after 40,000 generations. The fitness effects of some rearrangements were deduced from their level of parallelism (*i.e.* whether they occurred several times independently within different populations). These results have been published (Raeside *et al*. 2014).

Since we finished this deliverable well in advance, we performed additional experiments to investigate the effect of increased mutation rates on the robustness of the genomes during evolution. This has been performed and finished between M12 and M20. Six of the twelve populations evolved a hypermutator phenotype during evolution, thereby revealing an increased mutation rate owing to defects in DNA repair. While this is known to increase the rate of beneficial mutations, it also increases the genetic load (the deleterious mutation rate). Using one of the hypermutator population as a model, we showed that the mutation rate increase had a large impact on genome evolution. Indeed, we found deleterious mutations that were substituted due to the increased mutation rates. We showed that these deleterious effects were compensated by the activity of RNA chaperones and this was the first proof of this buffering effect of RNA chaperones. These results have been published (Rudan *et al*. 2015).

3. Conclusion

As stated in the introduction, robustness can be achieved at different levels. From the results presented here we can draw some general conclusions, first at the population level our experiments with the TEV model have shown that:

- Population size had a net significant effect of the evolution of TEV robustness, with populations evolved under a large effective population size regime being more robust, on average, than populations evolved in the small effective population size.
- Mutation rate has a highly significant effect only in combination with the effective population size in the evolution of TEV robustness.
- Hence, the predictions of the Krakauer & Plotkin (2002) model for the evolution of population-level robustness were partially fulfilled for TEV: we failed to observe the predicted synergistic effect of mutation rate as large populations evolved at high mutation rates were not more robust than their counterparts evolved at standard mutation rate.

Then, at the regulation level our experiments on *E. coli* have shown that:

- Perturbing the regulatory network of *crp* drastically reduced the catabolic breadth of *E. coli*. This effect was larger in the minimal glucose medium in which the evolution of the 12 populations of the LTEE occurred.
- Evolution therefore strongly selected a particular structure of the regulatory network. Disturbing this structure resulted in lower robustness of the evolved clones specifically in the evolution environment.

- The physiological effect in the evolution environment of perturbing the *crp* network was larger in the evolved clones than in their ancestor. We identified the mutations responsible for this larger effect and showed that they affected the level of DNA supercoiling. A link between catabolic repression and DNA topology is therefore responsible for maintaining robustness at the level of regulatory networks in *E. coli*.

Finally, when robustness is achieved at the genome level, we can conclude that:

- The wildtype genome architecture of TEV is more robust to mutational effects than any alternative architecture that we have created in the laboratory.
- Adding functional redundancy does not pays off for the cost of replicating the additional genetic material inserted, at least in the case of RNA virus for which selection for fast replication is strong.
- The large chromosomal rearrangements that have been identified during evolution of the 12 populations of the LTEE heavily restructured the chromosome but had no dramatic effect on robustness. They even in some cases increased the fitness of the evolved bacterial clones.

4. References

Aldana M, Balleza S, Kauffman S, Resendiz O. 2007. Robustness and evolvability in genetic regulatory networks. *J. Theor. Biol*. **245**: 433-448.

Ancel LW, Fontana W. 2000. Plasticity, evolvability and modularity in RNA. *J. Exp. Zool*. **288**: 242-283.

Batada NN, Hurst LD. 2007. Evolution of chromosome organization driven by selection for reduced gene expression noise. *Nat. Genet*. **39**: 945-949.

Codoñer FM, Daròs JA, Solé RV, Elena SF. 2006. The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. *PLoS Pathog*. **2**: 1187-1193.

Ciliberti S, Martin OC, Wagner A. 2007. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput. Biol*. **3**: e15.

Cooper, T. F., Remold, S. K., Lenski, R. E., Schneider. D. 2008. Expression profiles reveal parallel evolution of epistatic interactions involving the CRP regulon in *Escherichia coli*. *PLoS Genet.* **4**: e35.

De Visser JA, *et al*. 2003. Evolution and detection of genetic robustness. *Evolution* **57**: 1959-1972.

Elena SF, Wilke CO, Ofria C, Lenski RE. 2007. Effects of population size and mutation rate on the evolution of mutational robustness. *Evolution* **61**:666-674.

Gosset G, Zhang Z, Nayyar S, Cuevas WA, Saier Jr MH. 2004. Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *J. Bacteriol.* **186(11)**:3516-3524.

Krakauer DC, Plotkin JB. 2002. Redundancy, antiredundancy, and the robustness of genomes. *Proc. Natl. Acad. Sci. USA* **99**: 1405-1409.

Philippe N, Crozat E,Lenski RE, Schneider D. 2007. Evolution of global regulatory networks during a long-term experiment with *Escherichia coli. BioEssays* **29(9)**:846-860.

Proulx SR, Nuzhdin S, Promislow DE. 2007. Direct selection on genetic robustness revealed in the yeast transcriptome. *PLoS ONE* **2**: e911.

Proulx SR, Phillips PC. 2005. The opportunity for canalization and the evolution of genetic networks. *Am. Nat*. **165**: 147-162.

Raeside C, Gaffé J, Deatherage DE, Tenaillon O, Briska AM, Ptashkin RN, Cruveiller S, Médigue C, Lenski RE, Barrick JE, Schneider D. 2014. Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. *mBio* **5**: e01377-14

Rudan M, Schneider D, Warnecke T, Krisko A. 2015. RNA chaperones buffer deleterious mutations in *E. coli*. *eLife* **4**: e04745

Sanjuán R, Cuevas JM, Furió V, Holmes EC, Moya A. 2007. Selection for robustness in mutagenized RNA viruses. *PLoS Genet*. **3**: 939-946.

Sanjuán R, Elena SF. 2006. Epistasis correlates to genomic complexity. *Proc. Natl. Acad. Sci. USA* **103**: 336-342.

Sinha N, Nussinov R. 2001. Point mutations and sequence variability in proteins: redistributions of preexisting populations. *Proc. Natl. Acad. Sci. USA* **98**: 3139-3144.

Stearns SC. 2002. Progress on canalization. *Proc. Natl. Acad. Sci. USA* **99**: 539-557.

Stearns SC, Kaiser M, Kawecki TJ. 1995. The differential genetic and environmental canalization of fitness components in *Drosophila melanogaster*. *J. Evol. Biol*. **8**: 539-557.

Tromas N, Elena SF. 2010. The rate and spectrum of spontaneous mutations in a plant RNA virus. *Genetics* **185**: 983-989.

Van Nimwegen, E, Crutchfield JP, Huynen M. 1999. Neutral evolution of mutational robustness. *Proc. Natl. Acad. Sci. USA* **96**: 9716-9720.

Von Dassow G, Meir E, Munro EM, Odell GM. 2000. The segment polarity network is a robust developmental module. *Nature* **406**: 188-192.

Wagner A. 2005. Robustness, evolvability and neutrality. *FEBS Lett*. **579**: 1772-1778.

Wagner GP, Booth G, Bagheri-Chaichian H. 1997. A population genetic theory of canalization. *Evolution* **51**: 329-347.

Zheng D, Constantinidou C, Hobman JL, Minchin SD. 2004. Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucl. Acids Res* , **32(19)**:5874-5893.