Grant agreement for: Collaborative project*

Annex I - "Description of Work"

Project acronym: EvoEvo

Project full title: " Evolution of Evolution "

Grant agreement no: 610427

Version date:

Table of Contents

Part A

Workplan Tables

A1:
Project summary

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One form per project

General information

Project title [3]	Evolution of Evolution		
Starting date [4]	01/11/2013		
Duration in months [5]	36		
Call (part) identifier [6]	FP7-ICT-2013-10		
Activity code(s) most relevant to your topic [7]	:		
Free keywords [8]	Computational evolution; artificial chemistry; evolvability; microorganisms; smart-buildings; personal companion.		

Abstract [9]

Evolution is the major source of complexity on Earth, at the origin of all the species we can observe, interact with or breed. On a smaller scale, evolution is at the heart of the adaptation process for many species, in particular micro-organisms (e.g. bacteria, viruses…). Microbial evolution results in the emergence of the species itself, and it also contributes to the organisms' adaptation to perturbations or environmental changes. These organisms are not only organised by evolution, they are also organised to evolve.

The EvoEvo project will develop new evolutionary approaches in information science and will produce algorithms based on the latest understanding of molecular and evolutionary biology. Our ultimate goal is to address open-ended problems, where the specifications are either unknown or too complicated to express, and to produce software able to operate in unpredictable, varying conditions.

We will start from experimental observations of micro-organism evolution, and abstract this to reproduce EvoEvo, in biological models, in computational models, and in application software. Our aim is to observe EvoEvo in action, to model EvoEvo, to understand EvoEvo and, ultimately, to implement and exploit EvoEvo in software and computational systems.

The EvoEvo project will have impact in ICT, through the development of new technologies. It will also have impact in biology and public health, by providing a better understanding of micro-organism adaptation (such as the emergence of new pathogens or the development of antibiotic resistances).

A2:
List of Beneficiaries

Project Number [1]	610427	Project Acronym [2]	EvoEvo

List of Beneficiaries

No	Name	Short name	Country	Project entry month[10]	Project exit month
1	INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE	INRIA	France	1	36
2	UNIVERSITE JOSEPH FOURIER GRENOBLE 1	UJF	France	1	36
3	UNIVERSITEIT UTRECHT	UNIVERSITEIT UTRECHT	Netherlands	1	36
4	UNIVERSITY OF YORK	UNIVERSITY OF YORK	United Kingdom	1	36
5	AGENCIA ESTATAL CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS	CSIC	Spain	1	36

A3:
Budget Breakdown

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One Form per Project

Participant number in this project [11]	Participant short name	Fund. % [12]	Ind. costs [13]	Estimated eligible costs (whole duration of the project)					Requested EU contribution
				RTD / Innovation (A)	Demonstration (B)	Management (C)	Other (D)	Total A+B+C+D	
1	INRIA	75.0	S	694,776.00	0.00	117,182.00	0.00	811,958.00	638,263.00
2	UJF	75.0	T	606,027.00	0.00	33,724.00	0.00	639,751.00	488,244.00
3	UNIVERSITEIT UTRECHT	75.0	A	576,000.00	0.00	26,000.00	0.00	602,000.00	458,000.00
4	UNIVERSITY OF YORK	75.0	T	710,708.00	0.00	26,407.00	0.00	737,115.00	559,438.00
5	CSIC	75.0	A	628,199.00	0.00	13,906.00	0.00	642,105.00	485,055.00
Total				3,215,710.00	0.00	217,219.00	0.00	3,432,929.00	2,629,000.00

Note that the budget mentioned in this table is the total budget requested by the Beneficiary and associated Third Parties.

*** The following funding schemes are distinguished**

Collaborative Project (if a distinction is made in the call please state which type of Collaborative project is referred to: (i) Small of medium-scale focused research project, (ii) Large-scale integrating project, (iii) Project targeted to special groups such as SMEs and other smaller actors), Network of Excellence, Coordination Action, Support Action.

1. Project number

The project number has been assigned by the Commission as the unique identifier for your project, and it cannot be changed. The project number **should appear on each page of the grant agreement preparation documents** to prevent errors during its handling.

2. Project acronym

Use the project acronym as indicated in the submitted proposal. It cannot be changed, unless agreed during the negotiations. The same acronym **should appear on each page of the grant agreement preparation documents** to prevent errors during its handling.

3. Project title

Use the title (preferably no longer than 200 characters) as indicated in the submitted proposal. Minor corrections are possible if agreed during the preparation of the grant agreement.

4. Starting date

Unless a specific (fixed) starting date is duly justified and agreed upon during the preparation of the Grant Agreement, the project will start on the first day of the month following the entry info force of the Grant Agreement (NB : entry into force = signature by the Commission). Please note that if a fixed starting date is used, you will be required to provide a detailed justification on a separate note.

5. Duration

Insert the duration of the project in full months.

6. Call (part) identifier

The Call (part) identifier is the reference number given in the call or part of the call you were addressing, as indicated in the publication of the call in the Official Journal of the European Union. You have to use the identifier given by the Commission in the letter inviting to prepare the grant agreement.

7. Activity code

Select the activity code from the drop-down menu.

8. Free keywords

Use the free keywords from your original proposal; changes and additions are possible.

9. Abstract

10. The month at which the participant joined the consortium, month 1 marking the start date of the project, and all other start dates being relative to this start date.

11. The number allocated by the Consortium to the participant for this project.

12. Include the funding % for RTD/Innovation – either 50% or 75%

13. Indirect cost model

A: Actual Costs
S: Actual Costs Simplified Method
T: Transitional Flat rate
F :Flat Rate

Workplan Tables

Project number

610427

Project title

EvoEvo—Evolution of Evolution

Call (part) identifier

FP7-ICT-2013-10

Funding scheme

Collaborative project

WT1
List of work packages

Project Number [1]	610427	Project Acronym [2]	EvoEvo

LIST OF WORK PACKAGES (WP)

WP Number [53]	WP Title	Type of activity [54]	Lead beneficiary number [55]	Person-months [56]	Start month [57]	End month [58]
WP 1	Experimental observation of EvoEvo in action	RTD	2	85.00	1	36
WP 2	Development of an integrated modelling platform	RTD	1	52.00	1	18
WP 3	In silico experimental study of Evoevo	RTD	3	51.00	1	36
WP 4	A computational EvoEvo framework	RTD	4	61.70	6	36
WP 5	EvoEvo applications	RTD	1	45.00	18	36
WP 6	Project management	MGT	1	21.30	1	36
			Total	316.00		

WT2:
List of Deliverables

Project Number [1]	610427	Project Acronym [2]	EvoEvo

List of Deliverables - to be submitted for review to EC

Delive-rable Number [61]	Deliverable Title	WP number [53]	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D1.1	TEV and E. coli strains for robustness analysis	1	5	15.00	R	PU	12
D1.2	Analysis of robustness in TEV and E. coli	1	2	18.00	R	PU	20
D1.3	Analysis of evolvability (part 1)	1	5	16.00	R	PU	22
D1.4	Analysis of phenotypic innovation (part 1)	1	5	10.00	R	PU	24
D1.5	Analysis of evolvability (part 2)	1	2	18.00	R	PU	36
D1.6	Analysis of phenotypic innovation (part 2)	1	2	8.00	R	PU	36
D2.1	Specifications of the genome-network model	2	1	2.00	R	PU	3
D2.2	Genome-network model	2	1	6.00	P	PU	12
D2.3	Specifications of the population model	2	1	3.00	R	PU	6
D2.4	Population model	2	1	8.00	P	PU	14
D2.5	Specifications of the realistic network model	2	1	3.00	R	PU	6
D2.6	Network model	2	1	8.00	P	PU	14
D2.7	Specification of the integrated evolutionary model	2	1	3.00	R	PU	12
D2.8	Integrated model	2	1	8.00	P	PU	18

WT2:
List of Deliverables

Delive-rable Number [61]	Deliverable Title	WP number [53]	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D3.1	Evolution of variability; Mechanisms and consequences	3	3	12.00	R	PU	24
D3.2	Evolution of robustness; Mechanisms and consequences	3	3	12.00	R	PU	28
D3.3	Evolution of evolvability; Mechanisms and consequences	3	3	12.00	R	PU	28
D3.4	Evolution of open-endedness; Mechanisms and consequences	3	3	12.00	R	PU	30
D4.1	Computational meta-model definition	4	4	13.00	R	PU	18
D4.2	Computational model requirements specification	4	4	9.00	R	PU	18
D4.3	Computational run-time platform	4	4	14.00	O	PU	30
D4.4	Computational reflective run-time platform	4	4	18.00	O	PU	36
D4.5	Reflective application	4	4	6.00	D	PU	36
D5.1	Impact obtained from EvoEvo mechanisms on data stream cluster analysis	5	1	20.00	R	PU	36
D5.2	Impact obtained from EvoEvo mechanisms on evolution of a hardware personal companion	5	1	25.00	R	PU	36
D6.1	Project website	6	1	2.00	O	PU	1

WT2:
List of Deliverables

Delive-rable Number [61]	Deliverable Title	WP number [53]	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D6.2	Project communication media	6	1	1.00	R	PU	3
D6.3	Report of the kickoff meeting	6	1	2.00	R	CO	3
D6.4	first review report	6	1	3.00	R	CO	12
D6.5	Mid-term dissemination report	6	1	2.00	R	PU	18
D6.6	Program of interdisciplinary dissemination workshop	6	1	0.50	R	PU	32
D6.7	second review report	6	1	5.00	R	CO	36
D6.8	Final report	6	1	5.00	R	PU	36
			Total	299.50			

WT3:
Work package description

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One form per Work Package

Work package number [53]	WP1	Type of activity [54]	RTD
Work package title	Experimental observation of EvoEvo in action		
Start month	1		
End month	36		
Lead beneficiary number [55]	2		

Objectives

WP1 will explore EvoEvo properties in two microorganisms (the model bacterium Escherichia coli and the RNA virus Tobacco etch virus). Both organisms evolve at a high pace but their molecular structures are very different. The comparison of their evolutionary dynamics will thus likely result in the identification of the common (or different) traits that confer them their high evolutionary potential. These traits will then be feeded in the computational models and tested in the computational experiments, thus constituting the first step in the biological-to-application scheme of EvoEvo.

In WP1, we will address experimentally the pace of evolution of microorganisms and relate it to their robustness (task 1.1), evolvability (task 1.2) and open-endedness (task 1.3). In particular, we will address the relationship between robustness and evolvability by directly testing whether more robust genotypes are also more evolvable or, by contrast, whether they adapt in a slower pace to new environmental conditions. We will tackle these issues using in vivo experimental evolution (Figure 7), which consists in propagating living organisms for hundreds to tens of thousands of generations in defined environments [Hindré et al., 2012]. It provides a powerful methodology to analyse the molecular basis of adaptation and to draw a rigorous phenotype-to-genotype map. Here, we will use two different experimental systems, an RNA virus (TEV) and a bacterium (E. coli), both of which have become classic models in experimental evolution that allow detailed genetic manipulations and analyses. Furthermore, given their short generation times, large population sizes and, in the case of the RNA virus high mutation rates, relevant evolutionary changes take place after short periods of time, allowing us to observe evolution in action.

Short introduction to TEV Experimental Evolution:

Tobacco etch virus (TEV) belongs to genus Potyvirus within the family Potyviridae and is phylogenetically related to the Picorna-like supergroup of (+) strand RNA viruses. TEV is a plant virus and can thus be manipulated at low risks without any animal experimentation. TEV consists of a genomic RNA strand of ~9.5 Kb linked at the 5' end to a viral protein (VPg) and with a poly(A) tail at the 3' end. It encodes eleven mature gene products that result from the processing of a large polyprotein by three viral proteases, and a second polypeptide derived from a translational read-through process (relevant for the present study are the NIb cistron that encodes the viral RNA-dependent RNA polymerase, the CP cistron encoding the coat protein, the NIa-Pro cistron encoding the main viral proteinase involved in the proteolytic maturation of the polyprotein, and the HC-Pro cistron that encodes a multifunctional protein involved in aphid transmission, and a proteinase involved in genome amplification and suppression of plant RNA silencing defense). Owing to its huge economic impact, TEV is a well-studied model system, and CSIC has tailored it for studying evolution. Main advantages of this system are:
• The small and well-characterized viral genome.
• The ease with which recombinant viruses can be generated; this efficiency allows for experiments on relatively large scales (moreover, CSIC has generated a TEV infectious clone with enhanced stability in E. coli¸ pMTEV).
• The tools necessary to perform and interpret evolutionary experiments are already developed in this system (e.g. quantitative RT-qPCR assays for determining viral titers and fitness and expression profiling for the analysis of host-virus interactions).
• The host species which, while being a complex eukaryote, is simple enough to generate highly homogeneous host populations, allowing precise and reproducible quantitative analyses. Moreover, since TEV is a virus infecting common plants, the host species are easily and cheaply available.

In the context of EvoEvo, TEV is a particularly well-suited model. Since CSIC already focuses on evolutionary experiments in this pathosystem, recombinant viruses are regularly generated and already available. Another advantage is that many potyviral genomes are available, allowing for comparisons. What is particularly striking about comparisons between Potyviridae genomes is that gene order has been strictly conserved in all seven genera within the family. Thus, we are working in a system where the genome architecture appears to be rigid and under purifying selection. This is a prerequisite for better understanding why a given architecture has evolved (i.e., if changes to architecture were anticipated to have only minor fitness effects, this could make a model system less suitable). Finally, the short replication time and high mutation rates of TEV enables evolutionary experiments to be performed relatively rapidly.

Short introduction to the E. coli long-term evolution experiment:

The E. coli bacterium, one of the most thoroughly studied of all life forms, is a common inhabitant of the human gut but can also live freely as far as nutrients are available. The complete sequence of the genome of the harmless laboratory strain K-12 was published in 1997. It consists of a single DNA molecule of 4.6 Mb encoding ~4300 proteins, many of these having well-characterized functions. Systems Biology approaches provided network maps at the regulatory and metabolic levels. This bacterium is therefore an ideal system for experimental evolution since mutations can be traced at the genetic, transcriptomic, proteomic, metabolic and ultimately phenotypic levels. Moreover, many sequences of other E. coli strains are now available, enabling diversity studies. Owing to the huge information available, E. coli has been used in many evolution experiments, including the longest-running one, called here the “long-term evolutionary experiment” (LTEE) that was initiated in 1988 by Richard Lenski (Michigan State University). In the LTEE, twelve populations of E. coli B are propagated by daily serial transfer in a minimal glucose-limited medium for more than 50,000 generations [Lenski, 2004]. All twelve populations achieved large fitness gains. Several other phenotypic traits evolved in parallel in most or all populations, including cell size, growth parameters, catabolic functions, and global gene expression. Moreover, half of the populations evolved increased mutation rates.
Genetic and genomic analyses revealed that adaptation through these parallel phenotypic changes involves global modifications of both the structure and expression of the bacterial genome. First, mutation rates and large chromosomal rearrangements have been shown to be highly dynamic over evolutionary time [Wielgoss et al., 2013; UJF unpublished data], implying a strong restructuration of the evolved genomes and a large effect on bacterial evolutionary fate. Second, long-term evolution is associated with substantial rewiring of global regulatory networks including many beneficial mutations that affected genes encoding the most global regulators of transcription [Philippe et al., 2007]. Moreover, changes in epistatic interactions between global regulators strongly influenced the fate of bacterial cells during this long-term evolution [Cooper et al., 2008].
The LTEE therefore provides an ideal biological system to understand how the dynamic architecture of both global regulatory networks and genomes affects the overall adaptation properties, including robustness, evolvability and phenotypic innovation, at the entire and complex level of bacterial populations.

Description of work and role of partners

Task 1.1 Robustness at the population, regulatory network and genome levels (UJF, M01 - M20):

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, 2005a]. 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 behaviour and “genetic robustness” or “mutational robustness” when mutations are the cause of perturbations. Many issues related to genetic robustness remain unresolved. For example, asserting that elevated robustness is a fundamental property of living organisms is problematic because we often ignore what normal robustness should be [Ciliberti et al., 2007a]. 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 [Proulx & Phillips, 2005; Proulx et al., 2007; van Nimwegen et al., 1999; Wagner et al., 1997] or merely as a by-product of selection acting on correlated traits [Ancel & Fontana, 2000; Stearns et al., 1995; 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 network level in the E. coli model, and at the genome level in both models.

Section 1: Robustness at the population level in the TEV experimental model
As a consequence of high mutation rates, natural selection may have favoured the evolution of genomic robustness mechanisms in riboviruses [Elena et al., 2006; Elena 2012]. These mechanisms can be either intrinsic (inherent to the genetic architecture and replication mode) or extrinsic (the result of interacting with cellular factors). Previous results suggest that RNA viruses adopt an anti-robustness strategy based on individual hypersensitivity to mutations that allows the average population fitness to be maintained high due to the efficiency of selection in purging deleterious genomes from the population [Elena et al., 2006; Elena, 2012; Krakauer & Plotkin, 2002; Lalic & Elena, 2012; Sanjuán et al., 2004]. Whether mutational robustness can evolve in such conditions is an open question. Krakauer and Plotkin theoretically established under which demographic and mutational conditions individual genomes would evolve robustness or hypersensitivity (i.e., population robustness) [Krakauer & Plotkin, 2002]. By simulating these conditions in a context of experimental evolution, would it be possible to change TEV hypersensitivity?
The following experimental design is based on the theoretical study done by Krakauer & Plotkin [2002] about the conditions for the evolution of individual versus population-level robustness. It will specifically test whether at high mutation rate and large population size mutational robustness emerges as a population-level property, whereas at low mutation rate and small population sizes robustness evolves at the individual level.
The starting point will be two mutator TEV genotypes and two anti-mutator TEV genotypes previously isolated by CSIC. Each genotype will be evolved under two different demographic conditions: large and small population sizes. Population size will be manipulated by intercalating bottlenecks of size one (the effective population size of a population fluctuating in size is the harmonic mean of the population sizes at each stage. Hence bottlenecks enable to highly reduce the effective population size) through inoculation of Chenopodium quinoa leaves between consecutive Nicotiana tabacum passages.

a) Experimental evolution phase
1. Ten tobacco plants will be inoculated with each of the four chosen genotypes. Seven days post-inoculation (dpi), purify virions.
2. Split these extracts into two aliquots. Use each one for initiating experimental protocols that differ in the imposed effective population size.
3. For those lineages evolved at large population size, inoculate new tobacco plants. For those evolved at small population size, prepare serial dilutions and inoculate C. quinoa leaves. After lesions appear, isolate one of them for each lineage and use the resulting virion preparations for inoculating (without dilution) tobacco plants.
4. Repeat #3 25 – 30 times. In the case of the experiment at small population size, by “passage” we mean the entire cycle C. quinoa – N. tabacum.
b) Estimating mutational robustness
At the end of the experimental evolution phase we will estimate mutational robustness both for the numerically dominant genotype as well as for the entire evolved population for each lineage. Robustness measures will be done with five-fold replication.
1. Isolation of the dominant genotype from a local lesion taken from C. quinoa leaves.
2. Robustness estimation through 2 – 3 passages mutation-accumulation experiments (The logic behind using mutation-accumulation experiments for estimating mutational robustness is as follows. If an organism is robust against mutational effects, then its fitness would be poorly affected by the accumulation of a limited number of them. By contrast, if an organism is very sensitive to mutational effects, its fitness will be affected in a larger extent. In precise mathematical terms, 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 organism. Indeed, the slope measures the average sensitivity to mutational effects and, therefore, its inverse is a measure of mutational robustness [Elena et al., 2006]).

According to the Krakauer & Plotkin [2002] model, the following results are expected:
• For large populations and high mutation rates: increase in population-level robustness and evolution of individual hypersensitivity; that is, a negative correlation between the estimates obtained for heterogeneous populations and numerically dominant genotypes.
• For small populations at low mutation rate: increase in individual robustness without a precise prediction about the directionality in the population-level robustness.
The factorial design of the experiment will allow analysing data using a model II ANOVA in which population size and mutation rate will be treated as orthogonal random factors and genotype nested within the interaction of the

two main factors. Specific comparisons to test for differences between the two scenarios will be done by means of non-parametric tests.

Section 2: Robustness at the regulatory network level in the E. coli experimental model
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. UJF will 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. Mutations resulting in such changes will be identified to understand how robustness can be affected through the re-structuring of regulatory network. These questions will be addressed as a function of the environment, of the genetic background, and over evolutionary time.
The CRP-controlled regulon has been shown to be increasingly important during evolution in the LTEE (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]). 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, UJF identified a number of regulatory genes (spoT, fis) harbouring beneficial mutations that accounted for these changes in epistatic interactions with the CRP regulon. We will address whether these changes in the interactions between global regulators within the regulatory network affected the robustness of the evolved clones. This will be done in three steps:
1. Inactivation of crp in additional evolved clones: One sampled from each of the ten other populations at the same 20,000-generation time point, and one from each of the twelve populations at 40,000 generations. This will be performed by allelic exchange using a suicide-plasmid methodology routinely used by UJF [Cooper et al., 2008].
2. Fitness assays in various environments: The strains deleted for crp will be analysed for their growth abilities in alternative environments. Carbon-utilization profiles will be determined using Biolog plates, as well as abilities to cope with different stresses (osmotic, oxidative, temperature, pH, presence of antibiotics and drugs). Growth traits will include the measures of lag time, growth rate, and maximal optical density (in the environments in which growth will occur) and survival rates (for stress responses; All values will be given for each evolved background relative to the ancestral one), We will therefore identify combinations of E. coli crp-deleted evolved clones and alternative environments for which the robustness of the regulatory network has been strongly affected compared to the crp-deleted ancestral clone. Therefore, robustness to the crp deletion will be assessed as a function of environment, genetic background and evolutionary time. Moreover, as an additional parameter potentially influencing robustness after long-term adaptation, the effect of mutation rates will be addressed. Indeed, four of the twelve evolving populations became mutators owing to mutations in DNA repair genes at early time points during evolution (well before 20,000 generations), which resulted in a strong increase in mutation rates [Sniegowski et al., 1997]. Two additional populations evolved a mutator status after 20,000 generations and had elevated mutation rates by 40,000 generations [Blount et al., 2012; Wielgoss et al., 2013]. Therefore, the growth abilities in alternative environments for mutator versus non-mutator crp-deleted clones will be addressed.
3. Identify mutations that interfere with the crp deletion and with the robustness of the regulatory network: Based on the genome sequences that are available for all the clones that will be used here and on the phenotypes that we will measure in the different clones sampled from all populations at two different time points, we will identify mutations that were substituted during evolution and that may interact with the crp deletion to produce changes in epistatic interactions within the regulatory network. A set of relevant mutations will then be introduced in the ancestral chromosome together with the crp deletion to reproduce the network restructuring. We therefore hope to understand how robustness can be altered, and if possible improved, by regulatory changes inside the network of interactions among global regulators in an E. coli cell.

Section 3: Robustness at the genome level in both TEV and E. coli experimental models
Evolution of genome architecture can be grossly divided into three sorts of processes:

1. Those that increase genome complexity. An example would be horizontal gene transfer (HGT), the movement of coding sequences between lineages, followed by evolution to accommodate the new genes.
2. Those that decrease genome complexity. An example might be the deletion of a redundant gene or regulatory sequences.
3. Those involving a reshuffling of existing elements, without duplication events, and do not appear to represent increases or decreases in complexity.

For many real-world examples, it may be problematic to categorize the evolution of genome complexity. But we make this division as a conceptual tool to illustrate the sorts of process we are interested in, and to stress that our research program will consider scenarios that lead to different outcomes in terms of overall genome complexity. We intend to study seven scenarios (a-g below) pertaining to all three categories of genome evolution in either the TEV system or the E. coli system (since the genome architecture is strongly different in both systems, they can be used to test similar scenarios in conditions of highly stable genomes – for TEV – or fluid genomes – for E. coli). By generating new genomic organization, we will test whether any of them turns out to be more robust to the effect of mutational and/or environmental perturbations. For the TEV model, the robustness of these new genome constructions will be evaluated by mutation accumulation experiments as described in Section 1.

a) Complexity increase by genetic redundancy in TEV
One of the striking features of many eukaryotic genomes is the apparent amount of redundancy in coding and non-coding elements of the genome. There are fewer examples of redundant sequences in viral genomes, although there are clear examples of apparent redundancy in large dsDNA viruses. We will consider two cases in which we hypothesize it might be beneficial for TEV to evolve genetic redundancy (i.e., the existence of genetic elements performing the same function): the duplication of the CP and protease NIa-Pro cistrons. Both will increase genetic redundancy but the consequences of these duplications are unclear. We speculate that the duplication of these proteins may have widely different impacts in TEV fitness. At the one side, higher levels of CP expression could allow for the encapsidation of more genomic RNA molecules without affecting the accumulation of all other mature peptides. However, completion of the infectious cycle would still depend on the cytoplasmic amount of other limiting proteins (e.g. replicase NIb or silencing suppressor HC-Pro). Therefore, we postulate that overexpression of CP may have a minor fitness effect in TEV. At the other side, higher levels of NIa-Pro may result in a more efficient processing of the polyprotein, making more mature viral proteins available faster for the replication process. Since potyviruses have only a limited number of post-translational mechanisms for regulating expression levels of the different viral proteins, as all viral proteins originate from the polyprotein, we predict that the overproduction of NIa-Pro will alter the equilibrium concentrations of all the different mature peptides and thus have a major impact in TEV fitness. Both duplications will be performed and the robustness of the modified strains will be measured. The consequences of the duplication events will help illustrate how the virus genome accommodates potential beneficial gene duplications, and whether they lead to higher viral fitness.

b) Complexity increase by genetic redundancy in E. coli
Long-term evolution of E. coli in glucose minimal medium is associated with large chromosomal rearrangements that have been detected by genome sequencing [Wielgoss et al., 2013; D. Schneider unpublished data] and optical maps [www.opgen.com; Anantharaman et al., 2005; D. Schneider unpublished data]. More than 100 rearrangement events have been substituted in all twelve evolving populations after 40,000 generations, including large inversions of more than one third of the chromosome, deletions and duplications. All duplications involve one particular chromosomal region that includes the rpoS gene, which is the master regulator of stress responses in E. coli [Klauck et al., 2007]. Owing to their large sizes, these rearrangements are likely to strongly influence gene order, genome architecture and dosage of regulatory proteins within networks, and may affect growth traits and global transcription profiles. This makes these strains ideally suited to study evolution of genetic robustness. In scenario b, two representative duplications will be analysed by UJF to measure their effect on bacterial physiology in both the selection and alternative environments (see section 2) and on global transcription profiles.
1. As a first step, we will select two evolved clones sampled from each of two populations at a given time point, one bearing each specific duplication and one without, for a total of four evolved clones (to avoid any complication associated with high mutation numbers in the mutator populations, we will focus on duplications that occurred in populations that retained the low ancestral mutation rate).
2. The second step will consist in determining the effect of the duplications on the physiological traits of bacterial cells. Each pair of contemporary clones, with and without the given duplication, will be used in competition assays, each clone being competed with the other in the same conditions as prevailing during the evolution experiment (a neutral phenotypic marker will be introduced in each competitor to distinguish them). The fitness

associated with each duplication will therefore be calculated to evaluate whether the duplications were directly associated with adaptation or merely hitchhiked to fixation with other beneficial mutations. As for section 2, the robustness associated to these large rearrangements will be evaluated by comparing the growth and stress response traits of each clone pair in the same alternative environments.
3. Finally, the global transcription profiles of the four evolved clones will be compared to the one of the ancestor in the conditions prevailing during the evolution experiment, with six replicates for each condition, resulting in a total of 30 profiles (5 clones x 6 replicates). Analyses of these global transcription profiles will be performed by UJF as previously described [Le Gac et al., 2012].
All these experiments will therefore allow us to establish the link between adaptation, robustness, gene redundancy, and global gene expression.

c) Complexity increase by acquisition of new genes through HGT in TEV
There is strong evidence that HGT is an important mechanism in viral evolution. However, such events are rare and as such have not been studied by experimental evolution. Here we propose to add both functional and non-functional elements to the TEV genome, and study how they are accommodated by subsequent evolution. To study accommodation of a non-functional element, we will insert the enhanced green fluorescent protein (eGFP), a commonly used marker protein, into the TEV genome. To study the incorporation of a functional element, we will insert the Cucumber mosaic virus 2b silencing suppressor protein (CMV2b). We do not expect eGFP to play any function in TEV, and it is therefore a suitable model for the insertion of a non-functional sequence. An added advantage of using eGFP is that looking for fluorescence in plants during passaging can easily monitor loss of functional eGFP expression. TEV achieves suppression of silencing through HC-Pro, a multifunctional protein. By introducing a second, dedicated silencing suppressor protein into the TEV genome, we anticipate that HC-Pro may further optimize any of its other functions, while losing activity as a silencing suppressor. We can then study how the genome accommodates CMV2b.

d) Complexity decrease by functional redundancy in TEV
Functional redundancy of a particular element in a viral genome could be a consequence of a number of changes in the host environment. The function performed by an element could not be required any longer (e.g. because changes in receptor usage) or an analogue could be provided in trans by the host (e.g. by transferring genetic material from the virus to the host). Here we will study the evolution of the genome when there is functional redundancy for a given locus. In order to do so, we will make use of N. tabacum plants expressing TEV NIb protein. These transgenic plants support replication of a TEV NIb knockout (TEV-ΔNIb). The notion that microorganisms, and viruses in particular, optimize their genome size and remove functionally-redundant sequences is widespread, but not supported by much empirical evidence. We will first evolve a wildtype TEV in transgenic plants. This situation allows us to study the evolution of the redundant sequences, in particular whether the redundant sequences are removed from TEV genome and what other evolution occurs in the genome to compensate for such elimination. Then, we will evolve TEV-ΔNIb in the transgenic plants. This situation allows us to focus on the evolution of the remainder of the genome after a functionally redundant element has already been removed.

e) Structural shuffling by variation of gene order in TEV (constant complexity)
There are many instances in which evolution has conserved the meta-structure of the genome, and in many of these cases it is not obvious why genome organization has been conserved. In the case of the Potyviridae, gene order has been strictly conserved. Even in the member of the genus Bymovirus, which have bipartite genomes (i.e., two segments which are separately encapsulated), the order of genes is conserved on the two segments. Why has gene order been conserved in the family? One reason might be the processing of the polyprotein, as three TEV genes have protease activity (P1, HC-Pro and NIa). However, P1 and HC-Pro cleave the first two sites, while NIa cleaves the remaining seven. Given that there is no variation in gene order, cleaving of the polyprotein is not a sufficient reason, even in principle. We therefore ask whether conserved potyviral gene order has been selected for, or whether it is accidental.
To gain insight into this matter, we will experimentally manipulate gene order and test whether viruses with alternative gene orders can be viable, and if so what their fitness and virulence is. Finally, we will evolve viable viruses with an alternative gene order and then characterize evolved strains. We chose to alternate the position of NIb, because (i) deletion of the gene is lethal, allowing for a simple readout in infectivity assays, and (ii) transgenic, NIb-expressing tobacco plants are available (see above), allowing us to determine which viruses are still viable when NIb is provided in trans. The latter might give insight into whether losses of viability are due to NIb expression, or disruption of the expression of other genes. It should be noted that our strategy for generating constructs is amenable to other genes, allowing us to expand the experiment to test the effects of variation in gene order for other genes, or even multiple genes.

f) Structural shuffling by variation of gene order in E. coli (constant complexity)
The scenario (b) above will be repeated with two other clones bearing large chromosomal inversions instead of large duplications. In scenario f, two pairs, each isolated from a different evolving population of the LTEE, of two contemporary evolved clones (1 with an inversion and 1 without) will be studied for their fitness, growth and stress response traits, and global transcription profiles (24 in total: 4 clones x 6 replicates; the transcription profile of the ancestor being reused from scenario b). This will allow us to establish the link between adaptation, robustness, gene order, genome architecture and global gene expression in the context of a constant genomic complexity.

g) Structural shuffling by generation of bicistronic genomes in TEV (constant complexity)
Why have organisms evolved genomes that are distributed over multiple DNA or RNA molecules (i.e., multiple chromosomes, or in the case of viruses, multi-partite genomes)? All genera within the Potyviridae family are monopartite, with the exception of the Bymovirus genus, which are bipartite. We want to investigate why a segmented genome has evolved in this genus, in order to experimentally address the larger issue of why such organization evolves. What makes the evolution of genome segmentation in this virus particularly interesting is that it comes at an obvious cost: not only is the genome segmented, but the segments are also packaged into separate virions [unlike some other viruses, where multiple genome segments are included within the same virion (e.g. Orthomyxoviridae)]. Only when both virions infect that same cell are they infectious. Given this clear limitation, why has this one potyviridae genus evolved a segmented genome? In order to address this question, we will generate a TEV-based bipartite virus by splitting the TEV genome in the HC-Pro/P3 junction, mimicking the situation of the bymoviruses. We will then assess the fitness and virulence of this bicistronic TEV. Subsequently, we will evolve this virus in N. tabacum plants. Our experiments will therefore help elucidate what the cost of genome segmentation is, and how evolution will proceed upon the introduction of segmentation in the genome.

Roles in task 1.1:
• CSIC will perform all experiments on TEV (sections 1 and 3),
• UJF will perform all experiments on E. coli (sections 2 and 3),
• INRIA and UU will follow all the experimental process in order to be able to model it in WP3,
• CSIC and UJF will collectively produce deliverable 1.1 under the responsibility of CSIC,
• UJF and CSIC will collectively produce deliverable 1.2 under the responsibility of UJF.

Task 1.2 Evolvability at the population and regulatory network levels (UJF, M12 - M36):

A system is said to be evolvable if it can be modified through genetic change in a way that enhances survival and reproduction. For natural selection to act, the system must show heritable phenotypic variation. Yet, genetic robustness implies that the system produces little phenotypic variation in response to genetic variation. Therefore, robustness might limit evolutionary optimization and innovation [Lenski et al., 2006]. In this vein, theoretical work has postulated that buffering mechanisms can lead to maladaptation compared to what would be achieved in their absence [Frank, 2007]. Also, the analysis of gene expression noise in yeast suggests that noise control may indirectly increase mutational robustness, which might in turn hamper evolvability at the level of gene expression [Lehner, 2008]. On the other hand, genetic robustness facilitates the accumulation of neutral or nearly neutral variation by relaxing the intensity of natural selection. This accumulated diversity can become visible to selection upon changes in the environment or genetic background, and thus be the source of evolutionary innovation. Computer simulations on simple population genetics models predict that genetic robustness can sometimes facilitate access to new adaptive peaks provided that occasional failures of robustness mechanisms occur [Kim, 2007]. The view that robustness can foster evolvability has also been supported by lattice protein models and PCR-based mutagenesis experiments showing that protein variants with increased thermodynamic stability have increased genetic robustness and are more likely to evolve new catalytic capabilities [Bloom et al., 2006]. Task 1.2 will build on the preliminary results of task 1.1 to study evolvability and its interactions with robustness. It is divided in two sections that will address evolvability at two different levels using the most appropriate experimental model: evolvability at the population level in the TEV model, and at the regulatory network level in the E. coli model.

Section 1: Evolvability at the population level in the TEV experimental model
The evolvability of the different TEV genotypes constructed in Sections 1 and 3 of Task 1.1 will be evaluated. We assume that differences in robustness will exist among this large collection of genotypes, thus allowing us to test whether a correlation exists between robustness and evolvability both in the short- and long-term. To do so, we will proceed to evaluate the rate of adaptation of the different genotypes to new hosts. This may represent either

a case of soft (if the virus is already able of infecting although it does inefficiently and replicates and accumulates at low rates) or hard selection (if the virus is not able of replicating in the new host unless host-range mutants exist in the standing population variation). We will use hosts for which CSIC already has extensive experience. TEV primary host is N. tabacum, although it infects efficiently other members of the Solanaceae family [Lalic et al., 2011]. By contrast, its ability to infect hosts from other botanical families is limited [Lalic et al., 2011]. To cover a wide spectrum of possible hosts, we will use Nicotiana benthamiana and Datura stramonium (both belonging to the Solanaceae thus resulting in soft selection) and Helianthus annuus and Spinaca oleracea (that belong to other botanical families thus resulting in hard selection). We will investigate whether the highly perturbed genomes generated in Section 3 of task 1.1 will evolve back the ancestral gene order on each new alternative host or whether new mutations will appear to accommodate and re-optimize their levels of protein production.
Evolution experiments will consist of undiluted serial passages on each host. At least 10 independent evolution lineages will be founded with each of the genotypes (or with a subset of the most representative ones) on each of the four hosts. Infectious dosages will be equivalent in all transmission events and the impact of genetic drift will be minimized to make mutation and selection become the dominant evolutionary forces operating in these experiments. This will be achieved by using highly concentrated virus preparations and maximizing the number of infection foci per plant by mechanical inoculation thus minimizing transmission bottlenecks [Zwart et al., 2011]. Absolute fitness will be evaluated from growth curves as described elsewhere [Lalic & Elena, 2012]. A minimum of 25 consecutive passages will be performed, which is equivalent to ca. 100 viral generations (and to millions of replication events). Fitness trajectories for each TEV genotype and host will be analysed. To explore the genetic basis of adaptation, full genome consensus sequence will be characterized for each viral population at each experimental passage using a multiplexing approach with the Illumina HiSeq 2500 technology. We hypothesize that more robust genotypes will be less evolvable at the short-term whereas they will produce more innovation at the long-term in the more permissive hosts. By contrast, we also hypothesize that the more robust a TEV genotype, the faster it will adapt to the less permissive hosts (H. annuus and S. oleracea) since they will retain higher fitness than their more sensitive relatives.
To relate evolvability to the shape of the fitness landscape, we will evaluate the evolvability of TEV by testing how the structure of fitness landscapes conditions the existence and accessibility of adaptive pathways [Franke et al., 2011]. To do so, we propose to construct the empirical adaptive landscape for TEV. We will use our surrogated wildtype TEV as non-adapted genotype and its derivative TEV-At17, which has been experimentally adapted to Arabidopsis thaliana ecotype Ler-0 in previous work from CSIC [Agudelo-Romero et al., 2008]. TEV-At17 differs from TEV in six point mutations (three synonymous and three nonsynonymous) and phenotypically in ~2-logs increase in infectious viral load per gram of tissue and 10-fold increase in infectivity [Agudelo-Romero et al., 2008]. All 62 intermediate genotypes carrying 1 to 5 mutations will be constructed and their fitness evaluated in A. thaliana Ler-0. The accessibility properties and the topological characteristics of the resulting empirical landscape will be analysed using the statistical tests developed by Franke et al. [2011]. After describing the landscape, we will test whether the presence of a given mutation(s) conditions the future evolutionary steps to be taken by the evolving populations (i.e., reproducibility of the evolutionary pathway). To do so, we will let intermediate genotypes (between TEV and TEV-At17) evolve in A. thaliana Ler-0. After several serial passages in this host, we will characterize the consensus sequences at each passage and confirm whether or not the mutations fixed afterwards are those that characterized the TEV-At17 isolate.

Section 2: Evolvability at the regulatory network level in the E. coli experimental model
The objective will be to analyse the relationships between robustness and evolvability by genotype and environment. The results obtained during Task 1.1 section 2 will identify combinations of E. coli crp-deleted evolved clones and alternative environments for which the robustness of the regulatory network has been strongly affected after long-term adaptation, compared to the crp-deleted ancestral clone. We will choose two such combinations. To investigate the evolvability of the E. coli crp-deleted clones, we will let the two chosen crp-deleted evolved clones evolve in the two new alternative environments as well as in the original minimal glucose medium of the LTEE. The results will be compared with the evolution of the crp-deleted ancestral strain in the same environments. We will propagate the three E. coli crp-deleted clones in each of the three environments for 200 generations (less than one month), by serial transfers and with three replicates for each population. These 27 populations (3 clones x 3 environments x 3 replicates) will be propagated in microtiter plates, and samples will be frozen at 50-generation intervals at -80°C.
The effect of the new structures of the regulatory network will be evaluated on the adaptive ability of the bacteria in the various environments. The fitness of each evolving population will be estimated over evolutionary time both within the environment in which it was propagated and in the two other alternative environments. We expect the fitness to increase in the initial environment but it will be more difficult to predict in the alternative ones.

The genome of two evolved clones sampled after 100 and 200 generations and from each population will be sequenced (2 clones x 2 time points x 27 populations = 108 genomes). UJF will then identify mutations that compensated for the deletion of crp in the different clones, resulting in an increased fitness in their respective evolution environment. The relevant mutations will then be introduced by allelic exchange into the crp-deleted clone that was used as ancestor for the corresponding propagated population. These reconstructed clones will be used in competition experiments to confirm that the identified mutations were indeed responsible for adaptation and therefore for the emergence of a new regulatory network structure. Relevant biochemical and molecular analyses will be also performed to understand the new network structure (the type of analyses will be adapted to the identified mutations, on the basis of the information available for E. coli).
These experiments will allow us to understand how regulatory networks can be rewired after perturbation to provide new adaptive abilities to bacterial cells. The precise mechanisms underlying the plasticity of these regulatory networks will therefore be investigated as a function of both the genetic background (and therefore of a given state of the regulatory network) and environment. In particular, pleiotropic effects will be identified since the restructuration of the network in one particular evolution environment will undoubtedly affect the organismal phenotypic traits in the alternative environments. Moreover, newly emerging epistatic interactions will be identified inside the network as well as their translation into the phenotype map of the cells.

Roles in task 1.2:
• CSIC will perform all experiments on TEV (section 1),
• UJF will perform all experiments on E. coli (section 2),
• INRIA and UU will follow all the experimental process in order to be able to model it in WP3,
• CSIC and UJF will collectively produce deliverable 1.3 under the responsibility of CSIC,
• UJF and CSIC will collectively produce deliverable 1.5 under the responsibility of UJF.

Task 1.3 Phenotypic innovation at the population and regulatory network levels (UJF, M19 - M36):

This task will be divided in two sections that will address phenotypic innovation at two different levels using the most appropriate experimental model: phenotypic innovation at the population level in the TEV model, and at the regulatory network level in the E. coli model.

Section 1: Phenotypic innovation at the population level in the TEV experimental model
What is phenotypic innovation for an RNA virus? The phenotype of a virus is the effect it causes on its host, that is, the ability to infect a host (infectivity) and the severity of symptoms generated (virulence). How likely is a phenotypic innovation to occur in viral populations? Or, in other words, how likely is a viral population to contain a new genotype that successfully infect a new host species or that produces dramatically different symptoms in an already susceptible host species? What determines its spread in the population? What differences exist between naïve and well-adapted viruses in the way they interact with the cellular components of the host? Tackling these questions will enable us to shed light on the open-endedness properties of evolution in viruses. They also have profound implications in the emergence of new viral diseases.
We first sought to evaluate a first phenotypic innovation, namely, the ability to infect a new host. We propose to evaluate the fraction of all possible mutations that may confer TEV the ability to infect a set of new hosts. We will evaluate the infectivity and fitness of a collection of 66 single-nucleotide substitution mutants of TEV already available in CSIC across a panel of eight different hosts (infectious RNAs will be produced by in vitro transcription of the corresponding plasmids, and used to inoculate at least 10 plants from each species with equal amounts of RNA). These hosts (N. benthamiana, Capsicum annuum, D. stramonium, Solanum lycopersicum, H. annuus, Gomphrena globosa, S. oleracea, and A. thaliana) differ in their degree of genetic relatedness to the primary one (N. tabacum). The development of symptoms will be followed (no symptomatic infections will be determined by one-step RT-PCR). These experiments will identify whether certain mutations are more prone to broad TEV host range than others (i.e., to be associated with the capacity of infecting several of the new hosts) or whether significant interactions exist between mutations and environments (i.e., the outcome of infection is unpredictable and depends on each exact combination of virus genotype and host species).
Then, we will seek to evaluate a second level of phenotypic innovation, namely, the severity of symptoms induced by a relevant sub-set of TEV genotypes. First, we will characterize the symptomatology of all the genotypes described in previous sections. Second, we will chose a sub-set of 10 genotypes that induce a wide range of symptoms, from the normal etching pattern induced by the wildtype virus in tobacco to the dwarfism and necrosis induced by some of the single-nucleotide substitution mutants in N. benthamiana. Next, we will proceed to characterize the interaction between the virus' proteins and the host transcriptome by means of RNAseq (using the Illumina HiSeq 2500 technology). To do so, we will compare the transcriptomic profiles of infected

and non-infected plants (at least three biological replicates and two technical replicates per viral genotype). This analysis will identify lists of differentially expressed genes that will be used for functional annotation using the TAIR database (www.tair.org). Finally, we will attempt to contextualize the altered genes in the transcriptional regulatory network (TRN) and protein-protein interaction network (PPIN) models available for A. thaliana. With these analyses, we expect to uncover mechanisms by which different TEV genotypes induce symptoms in the host plants upon manipulation of key elements in the TRN and PPIN networks. Based on previous studies [Elena & Rodrigo 2012; Rodrigo et al., 2012], we hypothesize that stronger symptoms will be associated with the manipulation of key hub transcription factors that amplify the perturbation in the networks.

Section 2: Phenotypic innovation at the regulatory network level in the E. coli experimental model
Bacterial evolution experiments have shown that adaptive diversification associated with phenotypic innovation may occur in almost all tested environments. This evolutionary outcome is expected and even predictable when environments are heterogeneous (presence of spatial structure, different carbon sources…), owing to the availability of different ecological niches [Kassen & Rainey, 2004]. Moreover, and less expected, adaptive diversification also emerged in more homogeneous environments [Le Gac et al., 2012; Rosenzweig et al., 1994], owing to niche construction whereby bacteria generate themselves new ecological opportunities, for instance by secreting metabolites [Laland et al., 1999]. When known, the mutations resulting in such outcomes affect regulatory genes or sequences [Bantinaki et al., 2007; Spencer et al., 2007; Treves et al., 1998]. The objective here will be to investigate the relationships between the structure of global regulatory networks and the bacterial ability to produce phenotypic innovation. In particular, clones with different regulatory network structures will be propagated under conditions known to promote adaptive diversification and investigated for their ability to produce co-existing lineages of bacterial cells with differential phenotypic abilities.
A total of 26 strains will be used, all obtained from task 1.1 section 2: the ancestor and one evolved clone sampled at 40,000 generations from each of the 12 long-term populations (these 13 clones are available), together with each of their crp-deleted counterpart (three of them are already available [Cooper et al., 2008] and the ten others will be constructed during task 1.1 section 2). The state of the regulatory network will be known from the data that will be obtained during task 1.1 section 2. Each of these clones will serve as ancestor to initiate three replicate populations (for a total of 78 populations) that will be propagated in minimal medium containing two sugars. The two sugars will be chosen on the basis of the results of the growth measurements that will be performed during the second step of task 1.1 section 2. All populations will be propagated for 500 generations under these conditions (2.5 months) and plated at regular time intervals. We will then screen for mixtures of small and large colonies typical of the differential consumption of sugars [Le Gac et al., 2008; Spencer et al., 2007], and investigate the characteristics of these polymorphisms. In particular, we will investigate their stability over the 500 generations. For polymorphic populations, one evolved clone from each co-existing lineage will be physiologically analysed by measuring its growth traits with each sugar alone or both of them. Moreover, their interactions will be characterized by competing clones from different lineages at different initial frequencies to investigate whether negative frequency-dependent selection emerged [Le Gac et al., 2012]. Finally, genomes from relevant clones will be re-sequenced to identify the molecular bases of adaptive diversification as a function of the status of the regulatory networks.

Roles in task 1.3:
• CSIC will perform all experiments on TEV (section 1),
• UJF will perform all experiments on E. coli (section 2),
• INRIA and UU will follow all the experimental process in order to be able to model it in WP3,
• CSIC and UJF will collectively produce deliverable 1.4 under the responsibility of CSIC,
• UJF and CSIC will collectively produce deliverable 1.6 under the responsibility of UJF.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
2	UJF	45.00
5	CSIC	40.00
	Total	85.00

WT3:
Work package description

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D1.1	TEV and E. coli strains for robustness analysis	5	15.00	R	PU	12
D1.2	Analysis of robustness in TEV and E. coli	2	18.00	R	PU	20
D1.3	Analysis of evolvability (part 1)	5	16.00	R	PU	22
D1.4	Analysis of phenotypic innovation (part 1)	5	10.00	R	PU	24
D1.5	Analysis of evolvability (part 2)	2	18.00	R	PU	36
D1.6	Analysis of phenotypic innovation (part 2)	2	8.00	R	PU	36
		Total	85.00			

Description of deliverables

D1.1) TEV and E. coli strains for robustness analysis: Production of TEV populations with diverse degrees of mutational robustness and of different genomic architectures; production of crp-deleted bacterial strains and identification of combinations strains-environments affecting bacterial robustness. [month 12]

D1.2) 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. [month 20]

D1.3) Analysis of evolvability (part 1): Evolvability: correlation between viral evolvability and robustness; correlation between rates of virus evolution and genomic architecture; correlation between bacterial evolvability and restructuration of regulatory networks. [month 22]

D1.4) Analysis of phenotypic innovation (part 1): Phenotypic innovation: viral host transcriptome and bacterial evolution experiments with changes in regulatory networks. [month 24]

D1.5) Analysis of evolvability (part 2): Evolvability: determination of viral adaptive constraints; genetic characterization of the restructuration of bacterial regulatory networks. [month 36]

D1.6) Analysis of phenotypic innovation (part 2): Phenotypic innovation: identification of hub genes that may be targets of viral infections and that result in symptoms; relationships between bacterial regulatory networks and ability to produce stable polymorphisms. [month 36]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS6	Production of innovative strains	2	14	Evolution lineages have been produced in both the viral and bacterial models.
MS12	production of strains to study robustness	2	17	The viral and bacterial strains have been

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
				correctly produced and isolated.
MS13	Analysis of innovative strains	2	17	Innovative phenotype are listed and relevant mutations are identified and validated.
MS15	Characterization of the robustness strains	2	20	The strains have been correctly characterized at the molecular and phenotypic levels.
MS16	Characterization of evolvability	2	24	Mutation have been identified and validated
MS17	Evolvability: evolution experiments	2	25	Check that evolution lineages and reconstructed mutants have been produced.

WT3:
Work package description

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One form per Work Package

Work package number [53]	WP2	Type of activity [54]	RTD
Work package title	Development of an integrated modelling platform		
Start month	1		
End month	18		
Lead beneficiary number [55]	1		

Objectives

WP1 is studying EvoEvo in real microorganisms through experimental evolution. The objective of WP2 and WP3 are, respectively, to develop models able to tackle the different aspects of EvoEvo (variability, robustness, evolvability and open-endedness) at the different levels of organization explored in the project (genomes, networks and populations) and to reproduce in silico the experiments of WP1. They represent the second step in the route from the biological domain to the application domain.

In WP2 models will be built following the “digital genetics” approach [Adami, 2006]. In digital genetics, organisms are modelled by data-structures in a computer. The kind of structure used depends on the studied level of organization. It can be sequences, lists, networks, numeric vectors or programs depending on the formalism used to develop the model (for a review of the main formalisms, see [Mozhayskiy & Tagkopoulos, 2012; Hindré et al., 2012]). The essence of digital genetics lies in an evolutionary engine that enables the structure to reproduce, mutate and that selects it depending on a fitness criterion. Then, as the simulation goes on, the data-structures can change and acquire specific properties that can be studied latter on. In the context of the EvoEvo project, digital genetics has two decisive advantages. First, it enables to mimic not only the organisms but also the experimental framework of WP1 (thus being the direct in silico pendant of the in vivo experiments). Second, from an algorithmic view, it is very close to evolutionary computation. The construction of an integrated model in which all levels of organization will be interacting will thus be an important step towards the computational framework developed in WP4.
Many models have been proposed in the literature, Avida being the most well known [Wilke et al., 2001; Adami, 2006]. However, only few models are able to efficiently address questions related to evolution of evolution, in particular because most formalisms impose the organism’s structure to be stable. For example, in the “program formalism” used in Avida, structural modifications of the program, though possible, have a high probability to be deleterious and cannot be tested in practice. An indirect consequence is that most formalisms cannot afford chromosomal rearrangements that however crucial to reorganize the organism structure in order to increase robustness or evolvability. INRIA and UU have developed independently two formalisms that are specifically dedicated to the study of indirect selection. INRIA used the “sequence of nucleotides” formalism to develop the aevol model [Knibbe, 2007a; Knibbe, 2007b]. Using this model Inria showed that indirect selection could select specific genetic and network structure depending on the mutational and selective pressure [Knibbe, 2007b; Beslon, 2010a; Beslon, 2010b]. UU has proposed the “pearls on a string” formalism and used it to show that, in time-varying environments, regulation networks, metabolic networks and species networks can acquire structures that increase the evolvability of the organisms [Crombach & Hogeweg, 2008]. However, both models are restricted to specific levels of organization. In WP2, we will progressively merge the two formalisms in order to produce an integrated model able to decipher the contribution of the different levels of organization that interact to in fine “produce” a microorganism. To do so, the integrated model needs at least to contain (1) a sequence, to be able to precisely mimic the mutational events; (2) a regulation network, to change dynamically the activity of the genes; (3) a metabolic network, to consume and produce elements from its environment; (4) a trophic network, to enable different species to evolve in interaction.
It would be unrealistic to try to develop directly such a complex model. That is why WP2 is organized in four tasks. Tasks 2.1, 2.2 and 2.3 are parallel tasks in which two or three different components will be merged together. They have no or few interdependencies. The resulting codes and knowhow will then be used in task 2.4 in order to conceive and develop the integrated model. WP2 will thus result in a panoply of models that could be used to tackle specific evolutionary questions. It is important here to remind that, to be efficient, models must remain simple enough. The integrated model will be used as a reference and to tackle the most complex

questions but it must be considered as the last resort in the panoply. It is however also – in the context of the global project – the ultimate modelling step that will enable to project to fall into the computational framework development phase (WP4).

Description of work and role of partners

Task 2.1 Integration of sequence and network levels (Inria, M01 - M10):

In task 2.1, the two formalisms developed by Inria (aevol) and UU (pearls-on-a-string) will be merged together in order to propose a model in which the organism's genetic and metabolic networks is encoded at the sequence level. The resulting model will thus enable to study the effect of realistic genetic structure and realistic mutational operators (i.e. operators acting at the nucleotide sequence level) on the structure of the networks and on its EvoEvo properties.
By sharing knowhow of INRIA and UU we will identify the decisive elements of both formalisms and select the ones to conserve, the ones to change and the ones to remove. This first subtask will result in a model sketch that will be implemented and tested. As in every model development, both partners will care to maintain the model complexity low enough to enable its practical use. This includes computational complexity (in both time and space) but also experimental complexity. Indeed, the number of parameters to explore must be kept low to allow their practical exploration in a realistic computational time.
The model conception will be a joint work of partners INRIA and UU and it is not possible to detail it here. However, two main directions could be followed.
1. The first one is to extend the aevol model by changing its artificial chemistry model (see [Dittrich et al., 2001] for a review on artificial chemistries) from today's functional space model (in aevol, the function of a protein is modelled as a fuzzy set in an abstract functional space [Knibbe et al., 2007a; Knibbe et al., 2007b; Knibbe et al., 2008]) to a network model similar to the one used in the pearl-on-a-string formalism (metabolic elements are represented by integers and enzymatic functions are numerical functions in the natural integer space). The resulting model will be relatively simple and could strongly benefit from the expertise of both partners. However, it may be difficult to model genetic regulation with this extended formalism since genetic regulation involves interaction between the genome level and the protein level (both being independent elements here, encoded by two different formalisms: a sequence of nucleotides and a network of integer functions).
2. The second possibility is to start from the integers chemistry developed by UU (that enables to efficiently model both the genetic network and the metabolic network) and to develop a nucleotide sequence encoding for it. This will enable to increase the realism of the pearls-on-a-string model, in particular for the DNA binding of transcription factors. This approach could benefit from the knowhow of Inria that developed a prototype of regulation model in the aevol formalism by computing sequence alignments between the proteins primary sequences and the mRNA promoter sequences ([Beslon et al., 2010a; Beslon et al., 2010b]). Other formalisms proposed in the literature can be used as [Kuo et al., 2006; Leier et al., 2007; Mattiussi & Floreano, 2007; Marbach et al., 2009; Tagkopoulos et al., 2008]

Roles for task 2.1:
Following the model conception, Inria will implement the model and Inria and UU will test it in parallel. UJF and CSIC will validate the main conception choices in the light of microbiology and evolutionary biology knowledge. UoY will bring its software development expertise. As in all the development tasks of the EvoEvo project, we will use the best practice software engineering approach suitable for developing research software. This will include an agile software development methodology [Beck, 2000] and a test driven development approach [Freeman & Pryce, 2010]. The final documented stable release will be made Open Source, for potential 3rd party use.

Task 2.2 Modelling regulation and metabolism (Inria, M01 - M18):

In task T2.1, metabolism and regulation will be modelled by mean of network structures in which each node will be represented by a transfer function (typically a Hill-like function or a Michaelis-Menten function). However, this description is a huge simplification of the biochemistry reality that cannot account for complex situations such as cooperative/competitive binding of transcription factors, stochastic effects due to small enzyme or transcription factors concentrations, competitive effects between different binding sites, cycling between different transcription factors or energy-dependent binding/unbinding...
In many cases, such complex effects have been shown to interact - sometimes positively - with the evolutionary process. For instance, it has been shown both theoretically and experimentally that stochasticity of gene expression can speed-up evolution [Kaneko, 2011] or drive the emergence of non-evolutive strategies based

on stochastic switch between two phenotypes [Veening et al., 2008; Beaumont et al., 2009], a strategy often compared to the bet-hedging strategy that is well known in plants.
The objective of task 2.2 is to propose a realistic model of biological networks that account for all these phenomenon. In particular, this model will include a realistic-enough model of stochasticity in gene expression that will eventually interact with the network structure (e.g. networks motifs [Alon, 2007]) to propagate (or not) the noise up the phenotypic level. Such a model will enable us to test the cooperativity of heritable and non-heritable phenotypic variations and their positive or negative effect on evolution. The main difficulty here will be to propose a realistic enough modelling scheme and simultaneously to keep the computational complexity low enough for the model to be computable in a reasonable time. One on the possibilities is to use biochemistry simulation algorithms such as the Gillespie SSA algorithm [Gillespie, 1977] in order to compute the biochemical reactions in the model. However, this will probably results in high computational loads. Another possibility is to use stochastic equation in place of the deterministic ones to model the nodes of the network. Obviously less powerful, such a solution would however be more realistic in terms of computational load.

Roles for task 2.2:
Task 2.2 will be divided in three subtasks:
• The model will be jointly designed by Inria and UU. It will benefit from the knowhow of both partners and, in particular, from the cooperation of Inria with life science groups working on stochasticity in gene expression [Gandrillon et al., 2012]. External collaborators may be invited to temporally join the working group in order to bring in some specific expertise (in particular mathematicians and physicist).
• The model will be developed by Inria with the same development methodology as the one proposed in task 2.1 above (agile software and test driven development methodology).
• Inria and UU will test the model.
As for all models produced in the course of the EvoEvo project, the final documented stable release will be made Open Source, for potential 3rd party use.

Task 2.3 Modelling environment, population and trophic network (UU, M01 - M10):

In an evolutionary model, the environment model is a crucial element. It imposes external constraints on the evolving organisms, eventually resulting in the selection pressure due to the limited carrying capacity. Moreover, the environment is also modified by the organisms that live in, thus resulting in the creation of new evolutionary niches in which new species can emerge, eventually creating a complex ecosystem and a complex trophic network. Indeed, in all evolutionary experiments, it has been observed that even with an initially clonal population, complex population structures rapidly emerge with a high level of within-population diversity sometimes leading to the emergence of stable polymorphisms [Rainey & Travisano, 1998; Treves et al., 1998] even when it was not expected as in homogeneous environments [Rosenzweig et al., 1994; Laland et al., 1999; Le Gac et al., 2012]. Moreover, in many cases, evolutionary shifts have been detected to organize single cells into cooperating groups of cells [Velicer & Vos, 2009] with selection acting on public good synthesis [Griffin et al., 2004] and social traits [Velicer & Yu, 2003; Fiegna et al., 2006].
Such complex population structure rarely emerges in artificial evolutionary models. In most cases the evolutionary dynamic results in the emergence of a dominant specie that rapidly overcome all other species in the environment thus forbidding the emergence of a complex ecosystem. A notable exception is the Tierra model, developed in the beginning of the 90s [Ray, 1992] in which complex interactions between different species emerged due to the interaction of individuals that allowed inter-individual relationships such as parasitism. The poorness of population structure is due to the environment model that is generally overlooked. In most models, the environment imposes its constraints on the organisms but it is not modified by them. Thus the constraints are uniformly applied on all the organisms, thus resulting in a mostly uniform population structure.
UU has recently developed an extension of the pearls-on-a-string formalism in which organisms are able to release metabolite in the shared environment [Takeuchi et al., 2011]. Similarly, the aevol model developed by Inria has been modified to enable the spreading of a "public good" in the environment [Misevic et al., 2012]. Moreover, both partners are studying the effect of non-stationary environments on the structure of organisms (genomes and networks). In many of these situations, it has been observed that the population acquires the embryo of a structure with the development of polymorphisms, cooperators and resources cycling [Crombach & Hogeweg, 2009].
The objective of task T2.3 is to use this knowhow and others (e.g. the Tierra model [Ray, 1992]) to propose an integrated environmental model able to drive the emergence of a complex ecosystem. This model will in particular includes temporal variations (either cyclic or random), random noise, and metabolites

release/diffusion/consumption. It will thus enable us to study how the population structure contributes to the evolution of evolution phenomenon.

Roles for task 2.3:
In task T2.3, the model conception will be done by a joint work of Inria and UU under the scrutiny of UJF that will validate the choices under the light of bacteria interactions. The development of the environment model and it's integration to the general framework will be done by Inria with the same development methodology as the one proposed in task 2.1 above. Finally the model will be tested by Inria and UU.
As for all models produced in the course of the EvoEvo project, the final documented stable release will be made Open Source, for potential 3rd party use.

Task 2.4 Development of an integrated model (Inria, M06 - M18):

Tasks 2.1, 2.2 and 2.3 will produce independent models that will constitute, together with the already existing models, a panoply of evolutionary models able to be used in a large variety of in silico experimental evolution designs. All these models will be used in WP3 to study specific effects of indirect selection and evolution of evolution evolutionary strategies. However, all these strategies are likely to interact one with the others. For example, robustness can evolve at the genetic level (selection of an appropriate amount of non-coding sequences), at the regulation level (selection of stabilizing motifs), at the molecular level (selection of chaperone proteins), at the phenotypic level (selection of stable phenotypes) or at the population structure (selection of a stabilizing trophic network). Which of these levels will be first selected is an open question that none of the specific models can address. Similar questions arise for the selection of evolvability, open-endedness or variability (e.g. what is the optimal balance of heritable and non-heritable variability in a stochastic environment). In order to tackle such difficult questions, we will need an integrated model that will include all the studied levels of complexity (genome, networks and population) in a single application. This model will also be the ultimate production of WP2 to be transmitted to WP4 as the basis for the development of the computational framework. The development of the integrated model will reuse software elements developed in tasks 2.1, 2.2 and 2.3 as well as the knowhow acquired by Inria and UU in the same tasks. It will also reuse some software elements developed previously for the aevol and the pearls-on-a-string models. Like in the three previous tasks, but here with a strong emphasizing, the risk in this task is to maintain the model complexity low enough for the model to be (1) computable, (2) parameterizable and (3) understandable. However, even if the final model falls in one of these risks, the software development will be exploited in the project as the bridge toward the computational framework (see WP4).

Roles for task 2.4:
The development of the integrated model will be split in three subtasks:
• Model conception. This will be done by Inria and UU. UoY will also participate to the conception of the integrated model since they will lead the next workpackage in which this model will be reused.
• Model development. This subtask will be done by Inria with the same development methodology as the one proposed in task 2.1 above (agile software and test driven development methodology).
• Tests. The model will be tested by Inria and UU.
As for all models produced in the course of the EvoEvo project, the final documented stable release will be made Open Source, for potential 3rd party use.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	INRIA	20.00
2	UJF	5.00
3	UNIVERSITEIT UTRECHT	15.00
4	UNIVERSITY OF YORK	6.00
5	CSIC	6.00
	Total	52.00

WT3:
Work package description

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D2.1	Specifications of the genome-network model	1	2.00	R	PU	3
D2.2	Genome-network model	1	6.00	P	PU	12
D2.3	Specifications of the population model	1	3.00	R	PU	6
D2.4	Population model	1	8.00	P	PU	14
D2.5	Specifications of the realistic network model	1	3.00	R	PU	6
D2.6	Network model	1	8.00	P	PU	14
D2.7	Specification of the integrated evolutionary model	1	3.00	R	PU	12
D2.8	Integrated model	1	8.00	P	PU	18
		Total	41.00			

Description of deliverables

D2.1) Specifications of the genome-network model: Description of the modeling choices for the genome-network integrated model. This model should include a realistic genomic structure as well as a metabolic network translated from the genome. [month 3]

D2.2) Genome-network model: Documented runnable genome-network model suitable for running in silico experiments. This model should follow the choices presented in deliverable D2.1. [month 12]

D2.3) Specifications of the population model: Description of the modeling choices for the population level model. This model should include a realistic population structure enabling niche construction, inter-individual interaction and open-endedness. [month 6]

D2.4) Population model: Documented runnable population model suitable for running in silico experiments. This model should follow the choices presented in deliverable D2.3. [month 14]

D2.5) Specifications of the realistic network model: Description of the modeling choices for the realistic genetic/metabolic network model. This model should include a realistic cellular network model suitable for the study of network variability, robustness and evolvability. [month 6]

D2.6) Network model: Documented runnable network model suitable for running in silico experiments. This model should follow the choices presented in deliverable D2.5. [month 14]

D2.7) Specification of the integrated evolutionary model: Description of the modeling choices for the integrated model. This model should include most of the choices made for deliverables D2.2, D2.4 and D2.6. [month 12]

D2.8) Integrated model: Documented runnable integrated evolutionary model suitable for running in silico experiments. This model should follow the choices presented in deliverable D2.7. [month 18]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead beneficiary number	Delivery date from Annex I [60]	Comments
MS2	Valisation of the genome-network model	1	10	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.
MS4	Validation of network model	1	12	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.
MS5	Validation of the population model	1	12	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.
MS8	Validation of the integrated model	1	16	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.

WT3:
Work package description

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One form per Work Package

Work package number [53]	WP3	Type of activity [54]	RTD
Work package title	In silico experimental study of Evoevo		
Start month	1		
End month	36		
Lead beneficiary number [55]	3		

Objectives

WP3 will use the models developed in WP2 (and the models developed previously by Inria and UU) to study the emergence of variability (plasticity) robustness, evolvability, and population level open-endedness in in silico evolutionary experiments. It will thus produce a generalized knowledge and interpretation of the in vivo experiments. This knowledge will then be used to develop the computational framework and to enable efficient control of evolution in the applications.
The aim of WP3 is to elucidate, by mean of modelling, the extend to which and how the crucial properties of the phenotype-to-genotype (variability, robustness, evolvability, open-endedness) emerge dependent on:
• Patterns and timescales of environment variation,
• Degrees of freedom of the evolutionary process, including
o Defined and evolved genotype-to-phenotype mapping
o Defined and evolved mutational operators
o Defined and evolved coding structure of potential interactions
o Included levels of organization (genomes, networks, metabolism, ecosystems)
o Fitness measure

To enable direct comparison of the results of WP3 and WP1 (thus ensuring validation of the models), we will perform in silico evolutionary experiments that will mimic the in vivo evolutionary experiments of WP1 (Figure 6). However, contrary to what is generally done in in silico evolution, the starting point of in vivo experiments are always wild type organisms (here E. coli or TEV), i.e. an entity shaped by eons of evolution of which the history is of course not known. Their subsequent evolution is studied at the medium and short term to assess (changes in) robustness, evolvability, and open-ended population structuring (niche creation), and the genomic changes associated with these changes.
In contrast to in vivo experiments, for the in silico experiments we do not have such a given wild-type with an unknown evolutionary history, but will have to conduct long-term evolutionary experiments to obtain such "wild-types". This has the advantage that we then know the full evolutionary scenario that shaped it, and can vary this scenario and study the consequences thereof. During the long-term evolution we monitor the evolution of the structure of the entities at different levels of organization (genome, network, population) as well the "mutational profile" (i.e. the effect on fitness of mutations). We will do this using ancestor tracing from different time points such that we can assess the properties of the final successful lineage relative to temporarily fairly successful lineages. Moreover after such prior long-term evolution we can and will conduct medium and short-term evolutionary experiments similar to those used in WP1. Thus we will use "wild-type" strains obtained in various long term evolutionary scenario's as ancestors to continue evolution in a particular fixed environment (but using otherwise the same protocol as during long term evolution) for medium long term, and then compare the resulting strain with its ancestor with respect to robustness to environmental change and mutations, evolvability to other (fixed) environments, and population level niche differentiation and niche creation in heterogeneous environments and through nutrient exchange (Please note, however, that the terms long, medium and short term experiments are not of the same magnitude in the in vitro and the in silico experiments! We will have to find out which time frames are needed under the different evolutionary scenarios).
WP3 is divided into four tasks that directly respond to the tasks of WP1 (except for task 3.1 that corresponds to experiments that have already been conducted by UJF - thus task 3.1 will be compared to published results rather than to results obtained during the EvoEvo project). Each task will have to be divided in two subtasks:

• Develop long-term evolutionary scenarios and evolve "wild-type" species accordingly. The scenario development will be done in close collaboration with the platform development of WP2 and depending on the characteristics of the organisms that will be studied in the short-term experiment.
• Study robustness, evolvability, innovation and population/ecosystem structuring properties of the evolved strains through short/medium term in silico evolution experiment. This will be done by comparing the outputs of different evolutionary scenarios with respect to the properties of the evolved strains.

Description of work and role of partners

Task 3.1 Evolution of variability (Inria, M01 - M36):

In Task 3.1, we will study how digital organisms evolve their own evolutionary pace through the evolution of heritable and non-heritable variability. The former will evolve through the evolution of mutation operators and rates (e.g. by evolution of mutator strains [Taddei et al., 1997]). The later will evolve through the evolution of the stochastic properties of the digital organism's artificial chemistry (developed in WP2, Task 2.3). Indeed, it is now clearly demonstrated that, in microorganisms, the genotype-to-phenotype map is highly stochastic [Elowitz et al., 2002] and that this stochasticity can be used by the organism to acquire specific properties like bet-hedging or differentiation [Veening et al., 2008]. It has been proposed that it can play a crucial role in the emergence of antibiotic resistant strains. In Task 3.1, the contribution of both kind of variability to the evolution will be studied as well as the interactions between them.
Long-term evolution: Task T3.1 will use the platform to be developed in WP2.3 to explore the evolutionary consequences of stochastic gene regulation - under various environmental conditions. We will conduct long term evolutionary experiments, study the evolutionary dynamics (e.g. speed of adaptation) and the (network) structure of the evolved entities and compare with those with strains evolving without stochasticity.
Short-term evolution: The evolved strains will be let evolved in now fixed environments (different from the ones used for the evolution of the "wild-type", and we will compare adaptation to new environments of ancestor and newly evolved strains. Moreover, we will enable the strains to evolve their mutational variability and compare the evolution of mutational variability with the evolution of non-heritable variability. This will enable us to answer questions on the regulation of both kind of variability (e.g. is the variability in the population due to variable gene expression of to mutational events? is phenotypic variability a preadaptation to the new environments? can regulatory adaptation occur without environmental sensors? can non-heritable variability accelerate evolution?). Ultimately, this will enable us to draw the rules of variability evolution, to compare these rules with what is observed in in vivo experimental evolution and to use them to setup evolution of applicative software in WP5.

Note that the results of task 3.1 will not be directly compared with experiments. Indeed, in the LTEE (see WP1), it has been shown that the mutation rates can quickly evolve in microorganisms due to the emergence of mutator strains [Barrick et al., 2009] and subsequent evolution of compensatory mutations [Wielgoss et al., 2013]. Moreover, many experiments have shown that the non-heritable variability can be selected in specific conditions such as a variable environment [Elowitz et al., 2002; Beaumont et al., 2009; Kussel & Liebler, 2005]. Thus, the results of the in silico experiments will be compared with these published results and do not necessitate additional "wet" experiments.

Roles for task 3.1:
• The experiments will be jointly designed by INRIA and UU with the input of UJF and CSIC,
• Experiments will be done by INRIA and results will be jointly analysed by INRIA, UU, UJF and CSIC.

Task 3.2 Evolution of robustness (UU, M01 - M36):

To study the evolution of robustness at different organization levels (robustness of genomes, networks or population), wild-type strains will be produced using the different models produced in WP2. These strains will evolve in changing environment in order to ensure that robustness is a selected property. We will then be able to compare the genome, networks and population structures evolved in environment varying at different speed. Then, we will conduct mid-term evolution of the evolved wild-types in a constant environment (either an environment they have encountered before, or an entirely new environment) for a few thousands of generations. Then, we will be able to
• Characterize mutational robustness by measuring the fitness effect of a large sample of mutations (of the set of allowed mutations in the evolution of the ancestor strain, or knockouts) in the environment from which they were taken.

• Characterize “physiological” (regulatory) robustness of strains with a metabolism and/or environmental sensors by placing them in various environments and determine their fitness at a short timescale (not allowing further mutations).
• Characterize the regulatory robustness of mutants of these strains through systematic gene knock-out
Then, using the integrated model developed in task 2.4, we will study how these different levels of robustness are balanced by the evolutionary process when they are all available in a same organism.

Roles for task 3.2:
• The experiments will be jointly designed by INRIA and UU with the input of WP1 (UJF and CSIC),
• Experiments will be done by INRIA and UU and results will be jointly analyzed by INRIA, UU, UJF and CSIC.

Task 3.3 Evolution of evolvability (UU, M01 - M36):

In task 3.3, we will use the same "wild-type" strains as the one produced in task 3.2. We will then continue evolution of above evolved strains in a constant environment (either an environment they have encountered before, or an entirely new environment) and in highly variable environments for intermediate time duration. We will then characterize the evolvability of the parent strain and the recently evolved strain to high fitness in the new environment with respect to success rate, speed and changes at genome, regulome and metabolic level that this involves. For those strains capable of regulatory adaptation (i.e. with environmental sensors or metabolism) we will measure evolvability in terms of attained fitness relative to the fitness of physiological adapted states.

Roles for task 3.3:
• The experiments will be jointly designed by INRIA and UU with the input of WP1 (UJF and CSIC),
• Experiments will be done by INRIA and UU and results will be jointly analyzed by INRIA, UU, UJF and CSIC.

Task 3.4 Evolution of open-endednes at population level (UU, M01 - M36):

In task 3.4, the model developed in WP2, task 2.2 will be used to study how evolution can shape the environment at the ecological network (evolution of the trophic network) and how this will contribute to the pace and quality of evolution. In particular, we will study how these interactions extend the evolutionary degrees of freedom, by extending the metabolic potential of the evolving entities and by exchanging metabolites among the members of the population.
The models will be used to evolve complex ecosystems with interacting populations. During these long term evolutionary experiments we will study population differentiation and compare the properties of the evolved entities with those evolved in simpler (non-modifiable) environments.
Using the evolved strains, we will change the environmental conditions (i.e. the available metabolites) and study in mid-term evolutionary experiments how the resource utilization change at the individual and at the population level. In other words we will study in which situation can the population collectively solve the problem of consuming all available resources. Then, we will study how resource utilization changes at the individual and at the population/ecosystem level and how the variations of the available resources modify the structure of the population (e.g. emergence of many specialist strains vs. emergence of a few generalist strains).
The results of task 3.4 will be directly compared with the results of task 1.3 (WP1) and with the experimental evolution literature. We will thus be able to propose a theory form niche formation and to use this theory in WPs 4 and 5 in order to allow applicative software to be driven either by a population of evolving entities that collectively "consume" all the available information or by a single entity (although the best one) that would individually provide a global solution. Results of task 3.4 will enable EvoEvo to choose the best of these two strategies (or choose any intermediate strategy) for the applicative framework and applicative software developed in WPs 4 and 5.

Roles for task 3.4:
• The experiments will be jointly designed by INRIA and UU with the input of WP1 (UJF and CSIC),
• Experiments will be done by UU and results will be jointly nalysed by INRIA, UU, UJF and CSIC.

WT3:
Work package description

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	INRIA	15.00
2	UJF	5.00
3	UNIVERSITEIT UTRECHT	25.00
5	CSIC	6.00
	Total	51.00

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D3.1	Evolution of variability; Mechanisms and consequences	3	12.00	R	PU	24
D3.2	Evolution of robustness; Mechanisms and consequences	3	12.00	R	PU	28
D3.3	Evolution of evolvability; Mechanisms and consequences	3	12.00	R	PU	28
D3.4	Evolution of open-endedness; Mechanisms and consequences	3	12.00	R	PU	30
		Total	48.00			

Description of deliverables

D3.1) Evolution of variability; Mechanisms and consequences: A report describing how variability is indirectly selected in the model and its consequences on evolution. [month 24]

D3.2) Evolution of robustness; Mechanisms and consequences: A report describing how robustness is indirectly selected in the model and its consequences on evolution. [month 28]

D3.3) Evolution of evolvability; Mechanisms and consequences: A report describing how evolvability is indirectly selected in the model and its consequences on evolution. [month 28]

D3.4) Evolution of open-endedness; Mechanisms and consequences: A report describing how open-endedness evolves in the model and its consequences on evolutionary dynamics. [month 30]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS9	"in silico" wild-type strain for variability study	3	16	Production of in silico strains. These strains will be available for experiments that will study heritable and non-heritable variability.

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead beneficiary number	Delivery date from Annex I [60]	Comments
MS10	"in silico" wild-type strains for robustness and evolvability study	3	16	Production of in silico strains. These strains will be available to that will study robustness, evolvability and their interactions.
MS11	"in silico" wild-type strains for niche construction study	3	16	Production of in silico strains. These strains will be available for mid-term evolutionary experiments that will study open-endedness and niche construction.

WT3:
Work package description

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One form per Work Package

Work package number [53]	WP4	Type of activity [54]	RTD
Work package title	A computational EvoEvo framework		
Start month	6		
End month	36		
Lead beneficiary number [55]	4		

Objectives

WPs 2 and 3 are developing and producing models of EvoEvo mechanisms and approaches, focused on biological processes. The objectives of this WP4 are to take those biologically-oriented outputs, and develop suitable computational analogues, that form the basis of a novel route to open evolved computational and engineered systems. This framework is used by WP6 for developing computational applications.

The development route is via a computational meta-model. This is essential for developing a coherent bio-inspired computational approach [Stepney et al., 2005; Andrews et al., 2011]. Developing a computational model directly from a biological model runs the risk of confusing biological-contingent detail (for example, the existence of three pathways, say) with the underlying principle (the existence of multiple pathways refined for particular purposes), leading to a rigid, over-constrained, and naïve implementation. The meta-model route instead exposes the underlying principles, abstracts away from the irrelevant details, and results in a more flexible computational analogue.
Specifically, the objectives of this WP are:
1. define a computational meta-model of EvoEvo, by abstracting and interpreting the biological model in a form suitable for in silico implementation,
2. instantiate the meta-model into a computational model suitable for specifying demonstrator applications of EvoEvo,
3. implement the computational model as an executable computational platform, suitable for developing demonstrator applications of EvoEvo.

Description of work and role of partners

Task 4.1 Specification of the EvoEvo framework (UoY, M6 - M18):

Task 4.1 delivers the computational meta-model and computational model encapsulating relevant analogies of the biological EvoEvo processes (WP 2, Task 2.4), in a form suitable for implementing an executable EvoEvo platform (Task 4.2).
We will use the approaches for developing conceptual models of bio-inspired algorithms [Stepney et al., 2005] as refined by the CoSMoS meta-modelling approach [Andrews et al., 2011; Stepney et al., 2013] to perform this task. We will use a domain-driven modelling [Evans, 2004] and meta-modelling approach, to ensure that, despite the abstractions and translations needed along the way, a strong visible link is established between the biological processes and their appropriate computational analogues.
The main subtasks are:
1. Abstract and translate the EvoEvo biological processes into a suitable computational metamodel. This will include the framework that allows evolutionary operators and processes to evolve within some pre-specified paradigm, as developed through the biological model (an EvoEvo “virtual machine” based, for example, on an abstraction of regulatory network). Additionally, abstract mechanisms to allow interaction with the environment will be included.
2. Augment the meta-model with a run time approach to computational evolution, such that an application can encompass the necessary evolving population of individual processes, whilst manifesting as a single “organism”. (Hence an unexpected mutation may kill an individual process, but not the composite organism that is the application.)

3. Augment the meta-model with non-biological components needed for application development and experimentation, including instrumentation, analysis, user interface, and configuration facilities.
4. Instantiate the meta-model into a computational model instance suitable to support the application development workpackage (WP 5). This model forms the specification and design material for the specific computational platform implementation of Task 4.2

Roles for task 4.1:
• INRIA, UU and UoY will jointly specify the EvoEvo framework on the basis of the integrated model produced in task 2.4.
• UoY will be in charge of producing the specification documents.

Task 4.2 Framework development, level 1 (EvoEvo at data level) (UoY, M13 - M30):

Task 4.2 develops and delivers the implementation of the executable EvoEvo platform specified in task 4.1, in a form suitable for application development (WP5).
We will use a best practice software engineering approach suitable fro developing research software. This will include an agile software development methodology [Beck, 2000] and a test driven development approach [Freeman & Pryce, 2010]. We will produce this platform with a major release every five months, and minor releases every one–two months, each of increasing functionality, guided by the specific requirements of the applications (WP4). The intermediate releases will be made available to all the project partners. The final documented stable release will be made Open Source, for potential 3rd party use.
The main subtasks are:
1. Release 1 (month 20) initial skeleton implementation – The technical approaches are tested, interfaces are designed, and the system works “end to end”, but some components remain as “implementation stubs”. The system is usable for initial application design.
2. Release 2 (month 25) prototype implementation – The architecture has stabilized, and all components provide a degree of end-to-end functionality. The system is usable for initial prototype application development by project partners.
3. Release 3 (month 30) full functionality implementation – The system is fully functional and robust, fully documented. It is released for third party application development.

Roles for task 4.2:
• The framework will be developed by UoY and jointly validated by UoY and INRIA.

Task 4.3 Framework development, level 2 (EvoEvo at code level) (UoY, M13 - M36):

The essence of EvoEvo is systems that are “open”, that can modify what they do in unexpected ways. For example, they can evolve new kinds of evolutionary operators. In computational terms, this corresponds to new code (to support such novel operators) being evolved as the code executes. Task 4.2 develops a “semi-open” platform: open within a biologically determined domain, and abstracted into the EvoEvo virtual machine. Here we remove the restrictions imposed by the virtual machine, and explore full computational openness, that can perform direct code self-modification through reflection [Stepney & Hoverd, 2011].
We will use an agile modelling and software development approach, iterating modelling and implementation cycles, to ensure that the desired abstract properties are realizable in code.
This WP will create a computational platform that can support such run-time modifications to the code base in a resilient manner. It will exploit mechanisms of computational reflection in order to create:
1. Mechanisms for adding and modifying existing code
2. A “crash proofing” layer, so that arbitrary and unexpected modifications cannot cause the execution platform to crash. (This is analogous to an unexpected biological mutation killing the cell or organism, but not crashing the “laws of physics”.)
3. Mechanisms for supporting an interacting “population” of code fragments
4. Mechanisms for interacting with the external environment.

The main subtasks are:
1. Modify the meta-model to allow full reflection and self-modification of the computational agents.
2. Instantiate a suitable computational model incorporating such reflection capabilities.
3. Implement a prototype platform incorporating reflection in a “crash proofed” computational environment.

4. Evaluate the performance through a small “toy” application (a stripped down version of the WP5 application). The toy application will be developed using both the basic EvoEvo platform, and the reflective platform, and the different properties analyses to infer the extra, or different, capabilities provided by the full reflective approach.

Roles for task 4.3:
• The second level of the EvoEvo framework will be developed and validated by UoY.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	INRIA	10.00
3	UNIVERSITEIT UTRECHT	5.00
4	UNIVERSITY OF YORK	46.70
	Total	61.70

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D4.1	Computational meta-model definition	4	13.00	R	PU	18
D4.2	Computational model requirements specification	4	9.00	R	PU	18
D4.3	Computational run-time platform	4	14.00	O	PU	30
D4.4	Computational reflective run-time platform	4	18.00	O	PU	36
D4.5	Reflective application	4	6.00	D	PU	36
		Total	60.00			

Description of deliverables

D4.1) Computational meta-model definition: A report containing a description of a CoSMoS approach "Meta-Model": a definition of the meta-model capturing a suitable abstraction of the relevant parts of the biological models (D2.1,2.3,2.5,2.7), suitable for instantiation into a computational model of the evoevo algorithm. [month 18]

D4.2) Computational model requirements specification: A report containing a description of a CoSMoS approach "Platform Model": a specification of the instantiation of the meta-model (D4.1), including instrumentation and interfaces, suitable for implementing as an evoevo algorithm. [month 18]

D4.3) Computational run-time platform: A calibrated, tested and documented implementation of the platform specification (D4.2), suitable for developing the evoevo component of an open-ended application (WP5). A report containing a description of the platform and the CoSMoS approach "Simulation Platform". [month 30]

D4.4) Computational reflective run-time platform: A calibrated, tested and documented reflective implementation of the platform specification (D4.2), suitable for developing a reflective evoevo component of an open-ended application (D4.5). A report containing a description of the platform and the CoSMoS approach "Simulation Platform". [month 36]

D4.5) Reflective application: A report containing a description of a CoSMoS approach "Simulation Experiment": a description and evaluation of a prototype reflective application built using the Computational reflective run-time platform (D4.4) as the evoevo component [month 36]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead beneficiary number	Delivery date from Annex I [60]	Comments
MS7	Draft meta-model	4	14	Demonstrating a suitable computational abstraction of the biological processes
MS14	Draft platform design	4	18	Demonstrating a suitable implementation route from the computational model specification
MS19	Skeleton reflective code design	4	28	“Skeleton” design and implementation, demonstrating a suitable computational reflective approach to EvoEvo is possible within the project development process.

WT3:
Work package description

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One form per Work Package

Work package number [53]	WP5	Type of activity [54]	RTD
Work package title	EvoEvo applications		
Start month	18		
End month	36		
Lead beneficiary number [55]	1		

Objectives

WP5 constitutes the final step of EvoEvo. The computational framework developed in WP4 and the knowledge and know-how generating through the parallel experiments of WPs 1 and 3 will be merged to build proof-of-concept applicative software.

Embryos of living technologies actually exist for more than 15 years (see, e.g. [Sims, 1994; Funes & Pollack, 1999; Lipson & Pollack, 2000]) but they never really demonstrate usability of feasibility. Apart from the technological difficulties of building living technologies, this is due to the lack of real application and to the limit of the toy-problems used to demonstrate the capacities of these technologies. In the EvoEvo project, we decided to directly concentrate on a real application but we carefully designed it such that its difficulty should be manageable in the context of the project. Our objective here will be to design living technologies able to manage the complex, unstable and unstructured flux of information produced by smart houses and smart buildings in order to enable intelligent agents (here personal companions) to adapt to their usage context.
The new wireless sensor technologies have been at the origin of a profound shift in the concept of smart houses, smart buildings and smart cities. It is now relatively easy to spread large population of sensors in the houses/buildings/cities in order to monitor it in real time. However, this shift also profoundly modifies our management of information. Since sensors are no more dependent of costly wiring, they can be added, changed, moved, removed, dynamically while the system is active. In other words, the sensing structure now evolves faster than the structured of the sensed system! On the one hand, this situation creates huge information management difficulties since the sensing system cannot be modelled before it will be used. On the one hand, the generated flux of information will enable to monitor the system with a quality of service that has never been achieved before. The proof-of-concept applications of EvoEvo directly follow from these two points. The objectives will be:
• Application 1: Be able to generate a stable model of the environment despite the evolution of the sensing network. The biological analogy here will be the notion of circadian cycle that is maintained and used in most living beings despite the parallel and different evolution of their biological sensors. This application will be tackled in task 5.1.
• Application 2: Design intelligent agent(s) that “live(s)” in the smart building and are able to use the information flux produced by the building they are integrated in to learn and invent actions in accordance with building usage and building inhabitants. These intelligent agents will thus become personal companions of the building inhabitants and progressively adapt to them so that their presence becomes first acceptable and second useful. These agents will be designed in task 5.2.

Technically, both applications rely on our ability to produce a real flux of information, with a high diversity of sensors, high redundancy and high dynamics. That is why we will ask a smart-sensors company, the HiKoB company (http://www.hikob.com), to equip smart-room(s) with many wireless sensors of different nature (movement, light, temperature, humidity, floor pressure, sound…). However these sensors will not be placed once and for all. On the opposite, they will be added to the system, moved, removed… dynamically during the course of the project. Moreover, the equipped room(s) will not be a showroom specifically dedicated to project. On the opposite, it will be (a) room(s) used by the team (either cafeteria of the team, meeting room or offices). Note that, since the data collected are low-level, they are not personal data such that there is no ethical difficulty here. The generated flux of information will then be used by our EvoEvo software to generate a model of the room(s) and of its “inhabitants” (task 5.1). In a second step (task 5.2) the room(s) will be equipped with smart objects that will have to evolve in order to find their usage/function in accordance with the room(s) users. Ultimately, these objects will become personal companions of the different room users and adapt themselves

to the specific will of their preferred user. In order to lower the equipment cost of these tasks, the smart objects will not be specifically designed for the project. We will use commercial objects that will rely on wireless open technologies in order to be able to connect them to sensor network of the smart room(s). These “agents” will be chosen at the beginning of task 5.2. We will use a collection of agents of increasing complexity in order to test EvoEvo technology on tasks of growing complexity. A proposed set of agents (to be discussed in task 5.2) could be:
• Wireless controllable plugs. These will enable us to duplicate some of the supplies of the room (e.g. light) such that the agent can act on the environment.
• Personal companions. Commercial personal companions will be introduced in the room and acquire behaviour. An example of such companion is the Karotz smart rabbit (http://www.karotz.com, see below).
None of these agents will have a fixed goal. On the opposite, they will have to evolve in order to find it and permanently adapt it to their environment. Note also that the actions of these agents will be sensed by the sensor network deployed in the smart room(s), thus creating the enaction loop [Varela et al., 1991] that must be at the heart of any living intelligent technology.

Description of work and role of partners

Task 5.1 An open stream cluster analysis (Inria, M18 - M36):

In this task, we focus in the application of the evolution of evolution studied in the project to an unsupervised learning task over a data stream to perform cluster analysis, i.e., to find categories in the stream. We propose to tackle this problem in one of its most challenging settings (see B1.2.2 Machine learning) where both the underlying model, that associates objects to categories, and the object descriptors are changing over the stream. The challenge is here to initiate a kind of life long learning (w.r.t. the life of the learner) over a data stream, while the structures of interest in the data and the descriptions of the data themselves are subject to changes. To this aim, we will explore how the quality of the learners can be improved by taking advantage of the various degrees of variability, robustness, evolvability, and open-endedness of evolution that are identified by the EvoEvo project. In the case of a stream of data, the quality of a learner also incorporates an aspect related to time performances, since it includes the ability to keep the model up to date according to the stream rate, and thus it is a situation where the melting of variability/robustness/evolvability/open-endedness is likely to play a central role. The targeted application is to be able to identify categories of states of a room in a building, using a data stream containing measures obtained by probes located in this building and other data sources. In an office, such categories of states could be for instance "meeting in the office", "lunch break in the office", "the office is empty". The challenging difficulties will be to adapt the model to changes related to the kind of data collected, like the addition/removal of sensors, the breakdown of some probes or the modification of their location, and also to adapt the current model to reflect changes in the usages of the room that lead to new categories of states. These various usages could be for example: "office of a single person", "office of a group of workers", "break room", "printing facility room". If possible, the categories identified over time and the category of the current state will be used as complementary input for the personal companion in task 5.2.

Roles for task 5.1:
• The application will be designed by INRIA, UU and UoY with the inputs of WPs 3 and 4,
• The application will be developed and tested by INRIA.

Task 5.2 Evolvable Personal Companion (Inria, M18 - M36):

In task 5.1, the EvoEvo concepts will be used to analyse data without modifying them. In task 5.2, they will be used in a more reflexive way, enabling the evolving system to act in its environment, thus indirectly modifying its own perceptions. This will create the enactive sensori-motor loop that is at the heart of cognitive abilities [Varela et al., 1991].
In this objective, the smart-room(s) will be populated with active objects, ranging from simple controllable plugs to intelligent personal companions like the Karotz companion. Then, all these objects will be controlled by the computational framework developed in task 4.2, receiving as inputs the environment model produced by the application of task 5.1 as well as their own perceptions. Then, they will act on their environment and evolve at the contact of the room users, the organisms being positively selected when the users "accept" their action (i.e. the new environmental conditions) and negatively selected when the users "reject" their action (i.e. actively change the environment in reaction to the agent action). The precise scheduling of task 5.2 as well as the

implementation details will be chosen at the beginning of the task depending on (1) the results of task 4.2, (2) the theoretical knowledge produced in WP3 and (3) the preliminary results of task 5.1.

Roles for task 5.2:
• The application will be designed by INRIA, UU and UoY with the inputs of WPs 3 and 4 and task 5.1,
• The application will be developed and tested by INRIA.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	INRIA	30.00
3	UNIVERSITEIT UTRECHT	5.00
4	UNIVERSITY OF YORK	10.00
	Total	45.00

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D5.1	Impact obtained from EvoEvo mechanisms on data stream cluster analysis	1	20.00	R	PU	36
D5.2	Impact obtained from EvoEvo mechanisms on evolution of a hardware personal companion	1	25.00	R	PU	36
		Total	45.00			

Description of deliverables

D5.1) Impact obtained from EvoEvo mechanisms on data stream cluster analysis: A report on the development and results of the data-stream cluster analysis software. The report should identify the strengths and weaknesses of the evoevo approach for this application. [month 36]

D5.2) Impact obtained from EvoEvo mechanisms on evolution of a hardware personal companion: A report on the development and results of the evolvable personal companion. The report should identify the strengths and weaknesses of the evoevo approach for this application. [month 36]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS18	Application of EvoEvo mechanisms to data stream analysis	1	27	A first version of the evolving data-stream analyser is functional although not efficient.
MS20	Application of EvoEvo mechanisms to personal companion evolution	1	32	A first version of the evolving personal

WT3:
Work package description

Schedule of relevant Milestones				
Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
				companion is functional although not efficient.

WT3:
Work package description

Project Number [1]	610427	Project Acronym [2]	EvoEvo

One form per Work Package

Work package number [53]	WP6	Type of activity [54]	MGT
Work package title	Project management		
Start month	1		
End month	36		
Lead beneficiary number [55]	1		

Objectives

Project management will be carried out in Lyon by the INRIA partner. It will benefit from the experience of the INRIA partner in the coordination of interdisciplinary projects (project leader having leading or co-leading the "Institut Rhône-Alpin des Systèmes Complexe", an interdisciplinary consortium, for more the five years). Management will also benefit from the professional support of the INRIA staff for project monitoring (task 6.1), administrative and financial management (task 6.2) and dissemination (tasks 6.3 and 6.4). This professional-quality management will contribute to the success of EvoEvo by guaranteeing the global quality of the project, timely finalization of the deliverables and reports, tight budget following, and good communication, collaboration and transparency between the partners and towards the European Commission.

Description of work and role of partners

Task 6.1 Consortium Management and project monitoring (Inria, M01 - M36):

INRIA, as coordinator will have a global overview of all tasks through regular contact with task and work package leaders (at least twice a month). Communication will mainly be achieved through e-mail and visio/audio-conferences. Two diffusion lists will be created to facilitate communication between the project members. The "evoevo.coord@inria.fr" list will be used for communication between all local leaders in Lyon, Grenoble, Utrecht, York and Valencia. The "evoevo.all@inria.fr" will be used for communication between all members of the project. Both lists will be managed by the INRIA technical staff. Based on the contract information approved (description of work), Guillaume Beslon (the Project Leader) will support the consortium collaboration in order to keep EvoEvo on track according to deadlines and resources set for each project deliverable. The coordination role of partner 1 (INRIA) will benefit from the interdisciplinary experience of the Beagle team and of the project leader. Having a long experience of projects in interaction with "pure" computer scientists as well as projects in interaction with "pure" life scientists, Guillaume Beslon will be able to discuss/understand all the scientific elements of the project (although he of course cannot be an expert of all the fields covered in the project).
Semester meetings will take place in France, Nederland, UK or Spain to physically assess global and individual progress, discuss planning and future engagements. These meetings will also be means of internal dissemination in order to guarantee that all partners develop a common knowledge of the biological roots of the project as well as of the modelling approaches and of the application specificities. Meeting draft agenda will be proposed 4 weeks ahead and final version sent out 2 weeks before the event. Minutes of the event will be drafted by the coordinator and available for the partners in the following 2 weeks. External experts may be invited to the meetings depending on the needs of the consortium and acceptance of all the participating partners.
Monitoring will assess on continued basis, work progress towards initial objectives, deliverable preparation, writing and finalization after validation by workpackage leader and coordinator. Pre-defined milestones will help to assess regular progress and give green light for go-ahead. In case of delay pre-defined contingency plan will be implemented. Monitoring will be based on regular contact by e-mail, visio-conferences and meetings. The project coordinator will provide electronic tools to enable fast, efficient and secure storage and exchange of the large amount of data that will be generated during the project. In particular, a private website will be implemented in the project to share documents, data and codes. Each member of participating organization will have a private access. This website will enable collaborative work, common editing, progress monitoring and offer the coordinator an immediate overview of the project status. It will also offer distant access to the application

platform developed in WP4 in order to allow all partners to test it. A specific access to the website will be granted to the project officer.

Task 6.2 Administrative & Financial Management (Inria, M01 - M36):

INRIA, being coordinator for several framework projects, will, through their acquired know-how and their European projects/partnership Team, assure regular follow-up and total transparency towards the European Commission. Moreover all partners will have the possibility to have a rapid update on the project status. Official documents will be monitored, updated if necessary and archived at INRIA. Other costs will be monitored through expense reports on semester basis. Hence a global budget table including planned and real expenses will be available every semester.
Funds will be distributed by the coordinator as described in the contract if no major deviation is observed. All official reports for the European Commission will be prepared in advance by the coordinator, using input of all partners and validation of workpackage leaders.
Thanks to regular communication any deviation or need for amendment will be rapidly identified and official requests will be produced by INRIA according to the relevant guidelines available.

Task 6.3 Internal dissemination (Inria, M01 - M36):

EvoEvo is a transdisciplinary project gathering researchers ranging from “wet” biology to “pure” ICT. The success of such a project strongly depends on the ability of the partners to fully understand the questions, difficulties and results of each other’s. The internal dissemination task intends to organize information exchange between the partners in the course of the project.
• Each partner will be responsible for disseminating its knowledge and field specificity in the whole consortium. To achieve this objective, regular internal workshops will be organized, starting from the very beginning of the project. This series of seminar will be organized during the whole project (each partner giving at least one interdisciplinary lecture per year). These workshops will be organized in training session of at least half a daylong and will be co-localized with the global project meetings. If necessary, international experts could be invited as speakers during these workshops. The executive committee will decide workshop agenda at least one month before the workshop. All supports will be made available to the consortium members through the project website.
• All members of the consortium will share experimental and scientific methods in order to ensure compatibility of the results gathered in different fields. In particular, experimental evolution procedures will be shared between partners 2, 5, 1 and 3 (UJJ, CSIC, INRIA and UU). This will ensure the compatibility of the in silico experiments with the in vivo ones. Similarly, software development methods will be shared among partners 4, 1 and 3 (UoY, INRIA and UU) in order to ensure compatibility and exploitability of software and software elements during the whole project. UoY will be in charge of disseminating best practice software engineering approaches suitable for developing research software to all consortium members in charge of software development and testing. This will include an agile software development methodology [Beck, 2000] and a test driven development approach [Freeman & Pryce, 2010].

Inria will be in charge of task 6.3 owing to its long-lasting experience of interdisciplinary exchanges in the Rhône-Alpes Complex Systems Institute.

Task 6.4 Interdisciplinary dissemination (Inria, M01 - M36):

For the successful dissemination and exploitation of the EvoEvo scientific and technical outcomes, dissemination activities runs in parallel with the management activities in coordination with all workpackages and tasks leaders. As for any scientific project, the main dissemination media will be scientific publications and we expect EvoEvo to be able to publish results in the highest impact journals, owing to the research topic, interdisciplinary approaches and quality of the consortium.
Apart from scientific publications, other dissemination media will be used in the project:
• An official project website is planned which will serve as a showroom for disseminating the objectives, public reports and main results of the project. The website will also be used to make the data and public license codes available for the scientific community.
• EvoEvo’s results will be disseminated toward non-initiated public through dedicated pages on the website. Moreover, the partners are invited to popularize the results in order to disseminate the ideas and principle of evolution. In particular, the computational models developed in WP2 could be used as a basis for the

development of serious games to teach evolution and it's consequence, e.g. on the emergence and spreading of new diseases or on the risks of antibiotic treatment misuse.
• The fields of digital genetics, experimental evolution and artificial evolution are all very active but unfortunately only scantly interacting. Besides the classical dissemination tools (publications, conferences, seminars…), we will use our respective contacts to organize two workshops (in years 2 and 3) to bring together scientists from these domains.
• At mid-term a specific report will be drafted for public dissemination (website, newsletters, conferences) with the actual project progress and upcoming highlights.

At the end of the project, a workshop will be organized with CE to present the results. This workshop will serve as a showroom to demonstrate in situ the to two applications developed in WP5.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	INRIA	12.00
2	UJF	2.00
3	UNIVERSITEIT UTRECHT	2.00
4	UNIVERSITY OF YORK	3.30
5	CSIC	2.00
	Total	21.30

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D6.1	Project website	1	2.00	O	PU	1
D6.2	Project communication media	1	1.00	R	PU	3
D6.3	Report of the kickoff meeting	1	2.00	R	CO	3
D6.4	first review report	1	3.00	R	CO	12
D6.5	Mid-term dissemination report	1	2.00	R	PU	18
D6.6	Program of interdisciplinary dissemination workshop	1	0.50	R	PU	32
D6.7	second review report	1	5.00	R	CO	36
D6.8	Final report	1	5.00	R	PU	36
		Total	20.50			

Description of deliverables

D6.1) Project website: Public website of the project. Private website for collaborative work. [month 1]

D6.2) Project communication media: Leaflet and slideshow describing the project objectives. [month 3]

D6.3) Report of the kickoff meeting: Report of the kickoff meeting, including tutorial material. [month 3]

D6.4) first review report: Review report for period M1 to M12 [month 12]

D6.5) Mid-term dissemination report: Mid-term dissemination report. This report should prepare the final dissemination workshop, select conferences where to present the results and journal where to publish position papers to disseminate the results and concepts of the project. [month 18]

D6.6) Program of interdisciplinary dissemination workshop: Program of the interdisciplinary dissemination workshop that will be organized at the end of the project to disseminate ideas and concepts of EvoEvo in the biology and ICT communities. [month 32]

D6.7) second review report: review report for period M13 to M36 [month 36]

D6.8) Final report: Final report of the EvoEvo project. [month 36]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead beneficiary number	Delivery date from Annex I [60]	Comments
MS1	Kickoff meeting	1	1	Production of an efficient action plan for the first twelve months.
MS3	First review report	1	12	Check that the project is still in line with its objectives. Verify that the planned results are still realistically achievable and that the risk assessment does not raise a red flag.

WT4:
List of Milestones

Project Number [1]	610427	Project Acronym [2]	EvoEvo

List and Schedule of Milestones

Milestone number [59]	Milestone name	WP number [53]	Lead beneficiary number	Delivery date from Annex I [60]	Comments
MS1	Kickoff meeting	WP6	1	1	Production of an efficient action plan for the first twelve months.
MS2	Valisation of the genome-network model	WP2	1	10	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.
MS3	First review report	WP6	1	12	Check that the project is still in line with its objectives. Verify that the planned results are still realistically achievable and that the risk assessment does not raise a red flag.
MS4	Validation of network model	WP2	1	12	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.
MS5	Validation of the population model	WP2	1	12	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.
MS6	Production of innovative strains	WP1	2	14	Evolution lineages have been produced in both the viral and bacterial models.
MS7	Draft meta-model	WP4	4	14	Demonstrating a suitable computational abstraction of the biological processes
MS8	Validation of the integrated model	WP2	1	16	Proof-of-concept experiment showing that the model is evolvable and can be used efficiently for in silico experiments.
MS9	"in silico" wild-type strain for variability study	WP3	3	16	Production of in silico strains. These strains will be available for experiments that will study heritable and non-heritable variability.
MS10	"in silico" wild-type strains for robustness and evolvability study	WP3	3	16	Production of in silico strains. These strains will be available to that will study robustness,

Milestone number [59]	Milestone name	WP number [53]	Lead beneficiary number	Delivery date from Annex I [60]	Comments
					evolvability and their interactions.
MS11	"in silico" wild-type strains for niche construction study	WP3	3	16	Production of in silico strains. These strains will be available for mid-term evolutionary experiments that will study open-endedness and niche construction.
MS12	production of strains to study robustness	WP1	2	17	The viral and bacterial strains have been correctly produced and isolated.
MS13	Analysis of innovative strains	WP1	2	17	Innovative phenotype are listed and relevant mutations are identified and validated.
MS14	Draft platform design	WP4	4	18	Demonstrating a suitable implementation route from the computational model specification
MS15	Characterization of the robustness strains	WP1	2	20	The strains have been correctly characterized at the molecular and phenotypic levels.
MS16	Characterization of evolvability	WP1	2	24	Mutation have been identified and validated
MS17	Evolvability: evolution experiments	WP1	2	25	Check that evolution lineages and reconstructed mutants have been produced.
MS18	Application of EvoEvo mechanisms to data stream analysis	WP5	1	27	A first version of the evolving data-stream analyser is functional although not efficient.
MS19	Skeleton reflective code design	WP4	4	28	“Skeleton” design and implementation, demonstrating a suitable computational reflective approach to EvoEvo is possible within the project development process.
MS20	Application of EvoEvo mechanisms to personal companion evolution	WP5	1	32	A first version of the evolving personal companion is functional although not efficient.

WT5:
Tentative schedule of Project Reviews

Project Number [1]	610427	Project Acronym [2]	EvoEvo

Tentative schedule of Project Reviews

Review number [65]	Tentative timing	Planned venue of review	Comments, if any
RV 1	12	Bruxelles	Review planned to verify that the project is well engaged and respects the first objectives.
RV 2	36	Bruxelles	Final review. Can be in Bruxelles or Lyon depending on the interest of demonstrating the applicative part of the project.

WT6:
Project Effort by Beneficiary and Work Package

Project Number [1]	610427	Project Acronym [2]	EvoEvo

Indicative efforts (man-months) per Beneficiary per Work Package

Beneficiary number and short-name	WP 1	WP 2	WP 3	WP 4	WP 5	WP 6	Total per Beneficiary
1 - INRIA	0.00	20.00	15.00	10.00	30.00	12.00	87.00
2 - UJF	45.00	5.00	5.00	0.00	0.00	2.00	57.00
3 - UNIVERSITEIT UTRECHT	0.00	15.00	25.00	5.00	5.00	2.00	52.00
4 - UNIVERSITY OF YORK	0.00	6.00	0.00	46.70	10.00	3.30	66.00
5 - CSIC	40.00	6.00	6.00	0.00	0.00	2.00	54.00
Total	85.00	52.00	51.00	61.70	45.00	21.30	316.00

WT7:
Project Effort by Activity type per Beneficiary

Project Number [1]	610427	Project Acronym [2]	EvoEvo

Indicative efforts per Activity Type per Beneficiary

Activity type	Part. 1 INRIA	Part. 2 UJF	Part. 3 UNIVERS	Part. 4 UNIVERS	Part. 5 CSIC	Total
1. RTD/Innovation activities						
WP 1	0.00	45.00	0.00	0.00	40.00	85.00
WP 2	20.00	5.00	15.00	6.00	6.00	52.00
WP 3	15.00	5.00	25.00	0.00	6.00	51.00
WP 4	10.00	0.00	5.00	46.70	0.00	61.70
WP 5	30.00	0.00	5.00	10.00	0.00	45.00
Total Research	75.00	55.00	50.00	62.70	52.00	294.70
2. Demonstration activities						
Total Demo	0.00	0.00	0.00	0.00	0.00	0.00
3. Consortium Management activities						
WP 6	12.00	2.00	2.00	3.30	2.00	21.30
Total Management	12.00	2.00	2.00	3.30	2.00	21.30
4. Other activities						
Total other	0.00	0.00	0.00	0.00	0.00	0.00
Total	87.00	57.00	52.00	66.00	54.00	316.00

WT8:
Project Effort and costs

Project Number [1]	610427	Project Acronym [2]	EvoEvo

Project efforts and costs

Beneficiary number	Beneficiary short name	Estimated eligible costs (whole duration of the project)						Requested EU contribution (€)
		Effort (PM)	Personnel costs (€)	Subcontracting (€)	Other Direct costs (€)	Indirect costs OR lump sum, flat-rate or scale-of-unit (€)	Total costs	
1	INRIA	87.00	389,964.00	41,000.00	63,000.00	317,994.00	811,958.00	638,263.00
2	UJF	57.00	173,220.00	25,000.00	211,000.00	230,531.00	639,751.00	488,244.00
3	UNIVERSITE	52.00	289,000.00	5,000.00	61,000.00	247,000.00	602,000.00	458,000.00
4	UNIVERSITY	66.00	419,764.00	2,815.00	39,174.00	275,362.00	737,115.00	559,438.00
5	CSIC	54.00	250,648.00	69,650.00	106,250.00	215,557.00	642,105.00	485,055.00
Total		316.00	1,522,596.00	143,465.00	480,424.00	1,286,444.00	3,432,929.00	2,629,000.00

1. Project number

The project number has been assigned by the Commission as the unique identifier for your project. It cannot be changed. The project number **should appear on each page of the grant agreement preparation documents (part A and part B)** to prevent errors during its handling.

2. Project acronym

Use the project acronym as given in the submitted proposal. It cannot be changed unless agreed so during the negotiations. The same acronym **should appear on each page of the grant agreement preparation documents (part A and part B)** to prevent errors during its handling.

53. Work Package number

Work package number: WP1, WP2, WP3, ..., WPn

54. Type of activity

For all FP7 projects each work package must relate to one (and only one) of the following possible types of activity (only if applicable for the chosen funding scheme – must correspond to the GPF Form Ax.v):

• **RTD/INNO =** Research and technological development including scientific coordination - applicable for Collaborative Projects and Networks of Excellence

• **DEM =** Demonstration - applicable for collaborative projects and Research for the Benefit of Specific Groups

• **MGT =** Management of the consortium - applicable for all funding schemes

• **OTHER =** Other specific activities, applicable for all funding schemes

• **COORD =** Coordination activities – applicable only for CAs

• **SUPP =** Support activities – applicable only for SAs

55. Lead beneficiary number

Number of the beneficiary leading the work in this work package.

56. Person-months per work package

The total number of person-months allocated to each work package.

57. Start month

Relative start date for the work in the specific work packages, month 1 marking the start date of the project, and all other start dates being relative to this start date.

58. End month

Relative end date, month 1 marking the start date of the project, and all end dates being relative to this start date.

59. Milestone number

Milestone number:MS1, MS2, …, MSn

60. Delivery date for Milestone

Month in which the milestone will be achieved. Month 1 marking the start date of the project, and all delivery dates being relative to this start date.

61. Deliverable number

Deliverable numbers in order of delivery dates: D1 – Dn

62. Nature

Please indicate the nature of the deliverable using one of the following codes

R = Report, **P** = Prototype, **D** = Demonstrator, **O** = Other

63. Dissemination level

Please indicate the dissemination level using one of the following codes:

• **PU =** Public

• **PP =** Restricted to other programme participants (including the Commission Services)

• **RE =** Restricted to a group specified by the consortium (including the Commission Services)

• **CO =** Confidential, only for members of the consortium (including the Commission Services)

• **Restreint UE =** Classified with the classification level "Restreint UE" according to Commission Decision 2001/844 and amendments

• **Confidentiel UE =** Classified with the mention of the classification level "Confidentiel UE" according to Commission Decision 2001/844 and amendments

• **Secret UE =** Classified with the mention of the classification level "Secret UE" according to Commission Decision 2001/844 and amendments

64. Delivery date for Deliverable

Month in which the deliverables will be available. Month 1 marking the start date of the project, and all delivery dates being relative to this start date

65. Review number

Review number: RV1, RV2, ..., RVn

66. Tentative timing of reviews

Month after which the review will take place. Month 1 marking the start date of the project, and all delivery dates being relative to this start date.

67. Person-months per Deliverable

The total number of person-month allocated to each deliverable.

PART B

COLLABORATIVE PROJECT

Table of Contents

Table of Figures

B1. CONCEPT AND OBJECTIVES, PROGRESS BEYOND STATE-OF-THE-ART, S/T METHODOLOGY AND WORK PLAN

B1.1 Concept and project objective(s)

B1.1.1 Concept: let evolution evolve to build fast-adaptive applications

The ultimate goal of Information and Communications Technologies (ICT) is to improve human life through the extension of human capacities, abilities or communications. Yet, one of the profound movements that traverse modern social and human sciences is that the world cannot be described as a stable system. Humans and societies continuously change due to the many interactions that lead to instability, to the emergence of new social groups, ideas, modes, media. But ICT can only hardly tackle such highly unstable situations: Every encountered situation must have been foreseen long before it occurs, at the time the software is designed. This is due to at least two problems (*i.*) ICT conceptions principles, since the conception is based on a model of the system's environment and (*ii.*) the necessary stability of computer programs when in operation. Moreover, the development cycle of ICT systems (either software and hardware) is so long that its environment is likely to have changed before the first release of the system. A consequence of this necessary stability of ICT systems is that users - *i.e.* humans and society - must adapt to the ICT systems that are supposed to serve them. Inside the ICT world, the same difficulties are at work since software systems cannot efficiently adapt to the emergence of other new programs (or new releases of existing programs) in their environment. Thus, one of the challenges of modern ICT is to develop technologies that are able to adapt dynamically to the evolution of their context, of their user, of the data they receive, and of other systems they interact with - in a single word, of their environment.

The situation in completely different when looking at biology: Evolution, the process that created (and still creates) all the diversity of life, is a process by which organisms permanently adapt to their environment. Moreover, the environment of an organism is never stable as it also depends on the evolution of other organisms. While higher eukaryotes have evolved complex sensori-motor systems to adapt their behaviour to their changing environment, microorganisms are less sophisticated systems that lack complex sensori-motor abilities. However, they efficiently use mutation and selection to dynamically adapt to new conditions. Recent experimental evolution results have shown that they are able to evolve at an amazing speed: in virtually all experimental frameworks that use bacteria or viruses, important phenotypic innovations have emerged in only a few tens of generations (*e.g.* [Rainey & Travisano, 1998; Zhang *et al.,* 2011], reviewed in [Hindré *et al.,* 2012]). These results show that, more than being adapted to a specific condition, **micro-organisms are adapted *to evolve:* evolution has optimized their own ability to evolve, as a primary mean to react to environmental changes**. This "evolution of evolution" [Hindré *et al.,* 2012], also called "second-order evolution" [Tenaillon *et al.,* 2001] or "indirect evolution" [Kirschner & Gerhart, 1998, Reisinger & Miikkulainen, 2006] could offer ICT new paradigms to enable computational systems to dynamically adapt to their environment, *i.e.* to their users, domain of use or condition of use.

The idea to use a bio-inspired evolutionary metaphor to overcome the limits of ICT led to many powerful developments such as genetic algorithms, evolutionary strategies or genetic programming. Most computational evolutionary models and algorithms rely on two important concepts: the genotype-to-phenotype mapping and the fitness landscape. As we will explain below, both concepts must be rethought to fully exploit the "evolution-of-evolution" mechanisms.

The genotype-to-phenotype mapping summarizes in a single conceptual entity (the “mapping”) the complex molecular processes by which information flows from the genetic sequence to the organism’s phenotype. It thus concatenates in a single abstract process different phenomena such as mRNA transcription, gene translation, protein folding, biochemistry and cell dynamics. In typical applications of evolutionary computation, the phenotypic traits are the parameters of the problem to solve and there is a one-to-one mapping from genes to phenotypic traits (Figure 1a). Both the number of phenotypic traits and the number of genes are usually fixed over time.

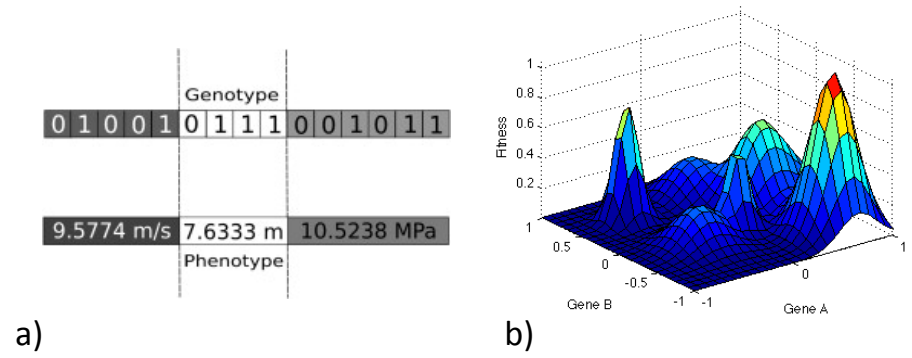


Figure 1 - a) Example of a genotype-to-phenotype mapping for a standard genetic algorithm used in an optimization problem in physics (from [Souza-Lima et al., 2011]). b) Example of a fitness landscape for a simple genome made up of two genes.

The fitness landscape (Figure 1b) is a metaphor that was proposed by Sewall Wright in the 1930’s [Wright, 1932] to provide a visual interpretation of the evolutionary process at the level of a population. Since then, fitness landscapes have become a fructuous concept used in evolutionary biology but also in evolutionary algorithms. According to Elderedge, fitness landscape are « By all odds the most important metaphor in macroevolutionary theory of the past fifty years » [Elderedge, 1989]. The central idea of the fitness landscape is that organisms or populations in evolution can be represented as points on a landscape where the altitude represents the fitness, *i.e.* the reproductive success, of the individuals. Selection can be represented by the local gradient of altitude and the mutation can be represented as a random noise added to the individual positions. In evolutionary computation, initially naive individuals submitted to a variation/selection process progressively climb the peaks of the fitness landscape that represents the problem to resolve. One can make complex reasoning from this metaphor in order to unravel hidden properties of the evolutionary process. At least two classical representations of the fitness landscapes have been proposed in the literature: the Fisher model [Fischer, 1930] in which the fitness landscape has a single peak with a Gaussian shape and the NK-fitness landscapes in which the landscape ruggedness can be

controlled respectively by the N and K parameters [Kauffman & Levin, 1987]. The former has been used to show how evolution is likely to slow down when it approaches the fitness maxima [Orr, 2009]. The latter has been used to show how the complexity of a landscape influences the course and results of an evolutionary process [Correia & Fonseca, 2007].

Although both concepts have proved to be powerful tools for evolutionary biology and evolutionary computation, they have strong limitations that prevent them to be used "as is" for the study of EvoEvo mechanisms. Their main limitation is that in practice, they are not fixed objects that evolution must deal with. On the opposite, both the fitness landscape and the genotype-to-phenotype mapping are very likely to change during the course of evolution.

- Firstly, the fitness landscape depends on the organism's environment, which is not static because it is composed of organisms that evolve themselves. Some authors now use the concept of "fitness seascape" to render the effect of fluctuating landscape on evolution [Mustonen & Lässig, 2009]. A given phenotype may be fit a given time point but become unfit a few generations later.
- Secondly, the genotype-to-phenotype mapping is mediated by complex biochemical components (RNA polymerases, ribosomes, chaperones…). It is thus a stochastic process with many interactions and feedbacks, producing possibly complex phenomena such as bistability. A given genome can give rise to different phenotypes depending on the environment or even on the intrinsic stochasticity of molecular events. Moreover, information pervasively flows from genotype to phenotype, but also from phenotype to genotype (due to, *e.g.* RNA interference, genetic regulation, or environment influence).
- Thirdly, even the number of dimensions in the fitness landscape and in the genotype-to-phenotype mapping cannot be considered constant. Biological evolution proceeds not only by gradual changes due to small mutations inside genes but also – and actually mainly – by reorganizing the genome [Jacob, 1977]. Genes are duplicated, deleted, moved, and exchanged between organisms. In the last years, due to the development of genome sequencing, results are accumulating that show that these events are one of the main forces driving genetic innovations [Blount *et al.,* 2012]. The very structure of the organism, which defines the number of dimensions in the genotype-to-phenotype mapping and in the fitness landscape, is thus changing during the course of evolution.
- Fourthly, the genotype-to-phenotype mapping and the mutation and rearrangement processes are under the control of complex molecular pathways: transcription, translation, regulation, protein folding, DNA repair, DNA recombination, DNA transfer mechanisms, and so on. The proteins involved in those processes are themselves encoded in the genome. Thus, they can also mutate, which will in turn influence the genotype-to-phenotype mapping and the pace of evolution for the other genes. In a given population, several mappings (or several mutation or rearrangement rates) can be in competition. These various lineages will either go to extinction or survive in the long term, depending on how many beneficial mutations the mapping (or the mutation or rearrangement rate) enabled. This phenomenon is called "indirect selection", of a mapping, of a mutation rate, or of a rearrangement rate. It has been shown theoretically and experimentally that the mapping and the mutation rate can be indirectly selected, at least in certain conditions [Wilke *et al.,* 2001; Bedau & Packard,

2003; Knibbe *et al.,* 2007b; Pigliucci, 2008; Crombach & Hogeweg, 2008]. In the most complicated scenario, if we consider the mutation of DNA recombination genes that affect the rate of gene duplications and deletions, even *the rate at which the number of dimensions of the genotype-to-phenotype mapping changes*, is evolvable.

With a fixed-landscape-based description of the evolutionary process, where mappings are defined before the evolution takes place, evolutionary computation cannot tackle changing problems, even less problems that change due to their (incomplete) resolution, or problems whose dimensionality changes over time. So far, in the context of evolutionary algorithms, one path was explored towards more dynamic fitness landscapes: evolving point mutation operators and rates has been extensively tried but with no or little gain of performance. In the EvoEvo project, we will rather explore the other possibility, which is to evolve the structure of the organism itself (*i.e.* the genotype-to-phenotype mapping and its dimensions). **The central concept of EvoEvo is the following: If the genotype-to-phenotype mapping and the fitness landscape are allowed to change over time, if they can be (indirectly) selected, then they can evolve and acquire properties than could favour evolution in changing environments.**

This can be done at the level of genetic structures via rearrangements (number of genes and their redundancy, relative position of the genes along the genome, selection of operonic structures, accumulation of non-coding sequences), and at the level of biochemical networks via molecular interactions (genetic networks, metabolic networks). A same phenotype can be "encoded" by very different networks (with few or many nodes, highly or sparsely connected, scale-free or not…). But all these networks strongly differ in their evolutionary potential. Some will resist most variation events; others will, on the opposite, induce large phenotypic changes even when undergoing a small mutation in one single node. Some will lead to gradual phenotypic changes; others will neutrally accumulate variations and, at some point, abruptly lead to catastrophic changes in the phenotype. Finally, a same phenotype can have different evolutionary possibilities depending on the partnerships it is able to maintain with other species in its environments, *i.e.* on the social structure it is involved in. **EvoEvo will study how organism's genetic, regulatory, metabolic and social structures can be indirectly selected and the resulting effect on the evolutionary pace. It will then exploit the results to create new living technologies that will exploit this knowledge.**

B1.1.2 Objectives

The main goal of EvoEvo is to develop a new family of evolutionary technologies that would exploit the principles of "evolution of evolution" as they are observed in microorganisms. In particular, **EvoEvo aims at creating evolvable software systems that are able to adapt dynamically to environmental changes, these changes being due the user, to new sources of data or to new software systems.**

To reach this main objective, several steps are required to fill the gap between experimental observation of indirect selection and practical applications of the same. Firstly, the biological roots of "evolution of evolution" must be better understood. Indeed, although biologists can see its consequences at the phenotypic level and, by means of genetic sequencing, at the genetic level, the roots of the high adaptation rate of microorganisms are mainly unknown, even in simple, well-known, organisms like the model bacterium *Escherichia coli* or even in

one of the simplest known organisms: RNA viruses. Biologists are only able to observe evolution in highly specific situations. This makes it difficult to decipher the many phenomena at work even in the simplest organism and to propose and test hypotheses for the mechanisms by which evolution evolves. In the absence of very precise hypotheses, it is difficult to get inspiration from micro-organisms to develop new evolutionary computational frameworks that would use evolution of evolution in the context of ICT. Secondly, a bridge is needed between "wet" experimental evolution and artificial evolution in order to enable concepts and methods to easily flow from biology to software engineering, and back. EvoEvo will use computational modelling as this bridge.

Hence, EvoEvo will achieve its main objective through the achievement of four scientific and technological objectives (Figure 2):

1. **Observe, quantify and characterize "evolution of evolution"** in microorganisms at the level of genomes, biological networks and populations. This will be achieved through experimental evolution and bioinformatics. We will gain a better understanding of this phenomenon that is still poorly understood. **EvoEvo will contribute evolutionary theory by allowing understanding of the surprisingly high pace of evolution of microorganisms.**
2. **Simulate "evolution of evolution"** in a computational framework. The simulations will use individual-based models that will help analysing the results of the evolution experiments. They will help us proposing hypotheses on the structural roots of EvoEvo at the levels of genetic sequences, regulation and metabolic networks and cell populations. The models will also constitute the basis of the computational evolutionary platform. **EvoEvo will contribute computational biology by the development of integrated computational evolutionary models that will be made available for the scientific community.**
3. **Design a computational evolution platform** to exploit EvoEvo in applicative software. This platform will be directly inspired from the *in silico* models but both simplifications and generalizations will be made. The former will remove from the models all the biological specificities that are not useful to exploit EvoEvo. The later will enable the framework to be used in different applicative contexts. **EvoEvo will contribute evolutionary computation by the development of a new framework that will use evolution of evolution at its heart.**
4. **Apply EvoEvo to real ICT problems**. Two applications of increasing difficulty will be proposed. The ability to exploit effectively EvoEvo in these applications will constitute the final proofs of concept that evolution of evolution can drive future technologies in an efficient way. **EvoEvo will propose proofs of concept showing the power of the principles developed in the project.**

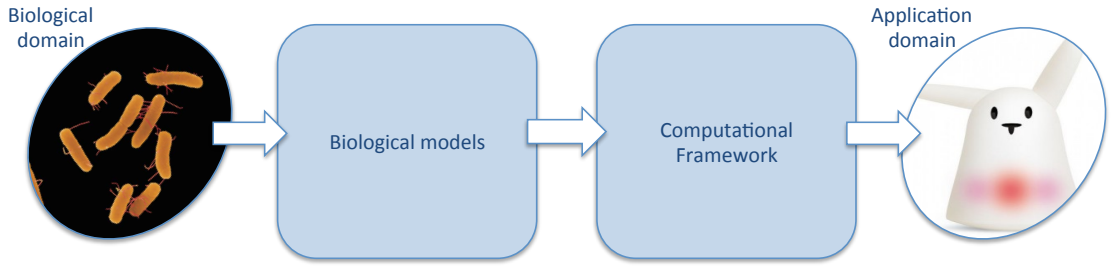


Figure 2 - EvoEvo's route from the biological domain to the living-technologies application domain

B1.1.3 Relevance to topics addressed by the call

EvoEvo addresses target outcome a) of the EVLIT objective, as detailed below.

Empirical, theoretical and synthetic approaches that define the key bio-inspired principles that drive future living technologies and the environment to use them in a controlled way.

In the context of EvoEvo, the key bio-inspired principles addressed will be the ability of evolution to change its own process and conditions of application through "indirect evolution". More precisely, we will focus on the ability of evolution to evolve not only the phenotype of an individual (through mutations at the genome level) but also to evolve the genotype-to-phenotype mapping of this individual. Since this mapping has a direct influence on the way genome mutations will be converted into phenotype variations, the genotype-to-phenotype map can acquire various features that can favours specific evolutionary paths. At least in some conditions, this can lead to second-order evolution that will later-on favours evolution. This key principle gives microorganisms their extraordinary reactivity and adaptability. However, it has never been entirely exploited in evolutionary algorithms, probably because this principle implies a huge complexification of the genotype-to-phenotype mapping in order to give it degrees of freedom that can be used by the indirect-selection process. **EvoEvo will unravel the biological roots of evolution of evolution in microorganisms, and exploit them to create a new generation of evolutionary technologies.**

Creating “living technologies” is a will that is sustained by the extraordinary capacities we observe in the living kingdom (far above the capacities of our artificial systems). However, the roadmap toward this will still contains too many *tierra incognita* to efficiently draw a path toward living technologies. Actually, nobody really knows what these technologies will look like, how they will be implemented (will it be software, hardware, “wetware” technologies?) nor what they will be useful for (will they complement extant technologies or replace them?). Although target a) is our main objective, EvoEvo also aims at a better understanding of what living technologies could be in a near future and how they could be used in a real context.

To this aim, we identified two applicative targets that we think could benefit form adaptable “living” technologies and we will address them to produce proof-of-concepts applications in the form of (1) an evolvable classifier system that will continuously adapt to the data stream it receives (even if the number of sensors, the size or the kind of data change dynamically) and (2) an evolvable personal companion that adapts to its user and condition of use. Both contributions correspond to open ICT problems that emerge today due to the massive deployment of sensors (*e.g.* in smart cities, smart buildings) and to the recent development of personal robot companions (*e.g.* Karotz, Papero…). Although these contributions will probably not constitute living technologies “*per se*” (the route toward real living technologies is clearly longer than a three-years project!), they will enable EvoEvo to also address EVLIT's target b) (*significant steps towards embodying these key principles and showing their usefulness in a technological context*) by providing important insights on what living technologies could be, how they could be integrated to extant technologies and, ultimately, how they could serve humans and societies.

B1.1.4 Approach

The EvoEvo project will specifically address four characteristics of the genotype-to-phenotype mapping in three systems of interest, as shown in Figure 3 and detailed below.

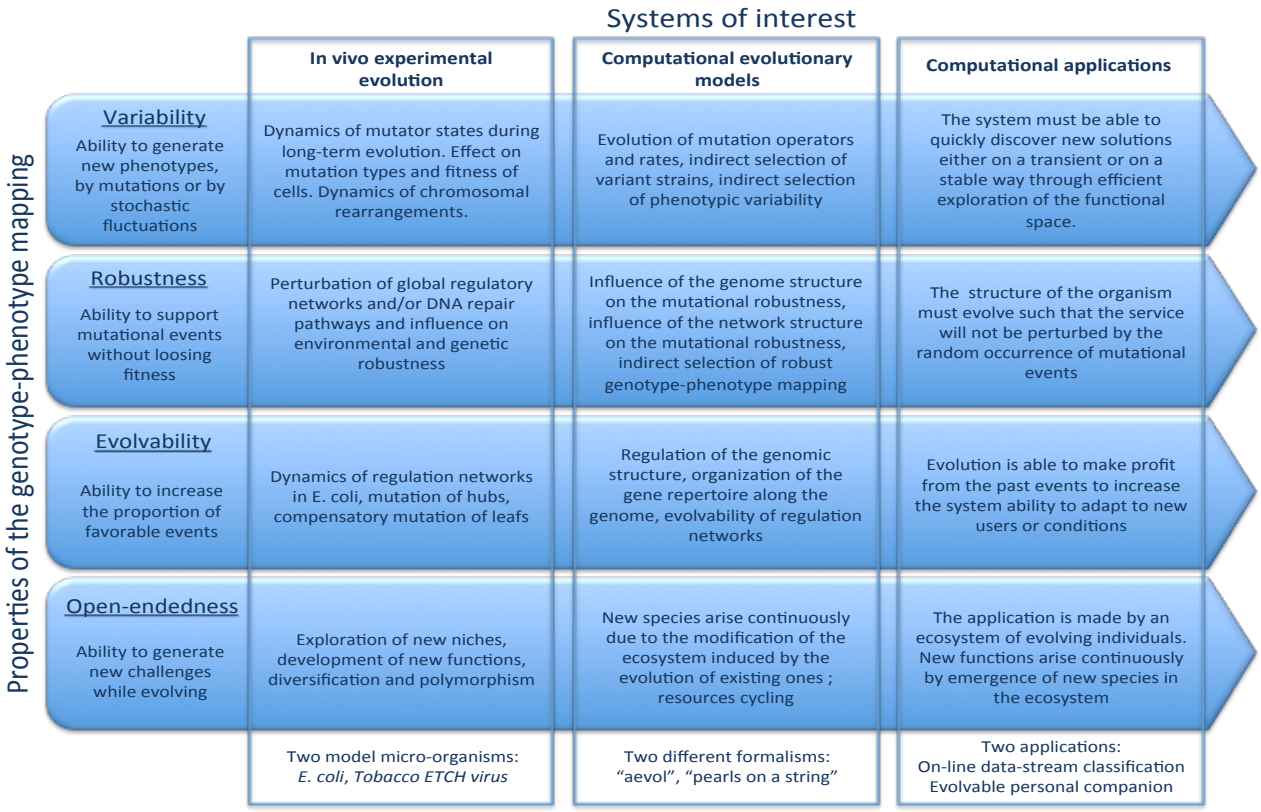


Figure 3 - Approaches towards EvoEvo: four characteristics of the genotype-to-phenotype mapping will be studied in real microorganisms, modelled through computational evolution and implemented in an applicative platform.

The four characteristics of the genotype-to-phenotype mapping that will be addressed in the EvoEvo project are:

- **Variability:** Variability is the ability to generate new phenotypes, by mutations or by stochastic fluctuations. It is a necessary condition for any evolutionary process to take place. However, in biological organisms, the amount of variability is controlled by complex pathways that *e.g.* correct DNA mismatches or breakings. In an ICT context, evolution of variability could help the evolving system to quickly discover new solutions either on a transient or on a stable way through efficient exploration of the functional space. Moreover, in real biological systems, mutational operators are highly diversified, including point mutations, but also large chromosomal rearrangements that can rapidly reshuffle the chromosome organization, extend or reduce the gene repertoire of an organism or even duplicate its entire genome through whole genome duplication.
- **Robustness:** Although mandatory, variability is a very dangerous process since it permanently produces deleterious mutations that lead to mal-adapted individuals. Robustness may evolve to correct these deleterious effects. It enables evolving systems to support mutational events without loosing fitness through *e.g.* canalization or the

selection of structures that creates neutral landscapes. In an ICT context, selection of robustness may favours the emergence of "organism" structured such that the service will not be perturbed by the random occurrence of mutational events.

- **Evolvability:** Depending on the genotype-to-phenotype mapping, the proportion of deleterious/neutral/favourable mutational events may change. Evolvability is the ability of a specific genotype-to-phenotype mapping to increase the proportion of favourable events. This can be done by the selection of specific genome structures or by the selection of specific networks structures. In an ICT context, evolvability will enable evolution to make profit from the past events to increase the system ability to adapt to new users or conditions.
- **Open-endedness:** Biological evolution is not directed towards a specific target. On the opposite, evolution has the ability to generate new challenges while evolving by *e.g.* exploiting new niches created by the evolution of other species. In an ICT context, open-endedness can but exploited when an application is made by an ecosystem of evolving individuals. In such a structure, new functions will arise continuously by emergence of new species in the ecosystem and/or the extinction of maladapted ones.

EvoEvo will specifically exploit chromosomal rearrangements as a primary operator leading to the evolution of organism's structure. It will construct new evolutionary algorithms that will use them as a central way to simultaneously evolve the phenotype of the artificial organisms and genotype-to-phenotype mapping that will drive later-on evolution.

The EvoEvo project is based on up-to-date results and concepts in microbiology, evolutionary biology and experimental evolution. Exploiting these to develop new technologies is not straightforward and needs (1) to conduct additional experiments on micro-organisms to refine these results and concepts in the context of their possible technological exploitation, (2) to be able to abstract these results and concepts from their initial domain of validity (*i.e.* microbial evolution) in order to enable their application in other domains (3) efficiently transfer these results and concept from the life-science domain to the application domain. To achieve these objectives, EvoEvo assembles a highly multidisciplinary consortium that will gather all aspects of the project, from microbiology to software/hardware development. Moreover, to ensure that these three needs are fulfilled, EvoEvo will directly tackle three systems of interest:

1. **Real microorganisms** (the bacteria *Escherichia coli* and the plant RNA virus *Tobacco etch virus*). This will enable us to understand if/when/how the characteristics of the genotype-to-phenotype mapping are selected in real evolving systems that have to adapt to new conditions.
2. **Computational evolutionary models** (specific computational models of evolution at the level of genomes, networks, population and resources as well as an integrated model). These models will be used to generalize the observation made on real microorganisms. They will enable us to propose general laws that could explain biological observation and, in parallel, be exploited in artificial systems.
3. **Computational framework and applications.** The evolution of evolution strategies observed in real microorganism, simulated in the computational evolutionary models,

will be used to develop real ICT applications that can dynamically adapt to their environment. The aim of EvoEvo is to propose a generic computational framework exploiting the principles of evolution of evolution. This framework will serve as a basis for the development of computational applications in various domains. Although the framework should be versatile, we will focus on an application field that will specifically requires ICT applications to dynamically adapt: smart building hosting smart objects. Indeed, in the context of smart-systems (smart-cities, smart-building, smart-objects), the system changes permanently due to the addition of new sensors (*e.g.* in the city or in the building), new pieces of software (*e.g.* a new application in a smart-phone) or new hardware systems (*e.g.* a new personal device). Moreover, the users also change and adapt dynamically to the object (*e.g.* for a personal companion). EvoEvo systems will challenge this problem. A room will be progressively equipped with sensors and smart-objects and the system will have to develop and maintain a stable model of the room and room usage (though the sensors and usage are changing along time). Moreover this model will be used by smart-objects (personal hardware companions) that populate the room, enabling them to be robust to their environment changes, to be able to adapt to the new users, and to discover new functionalities they could offer to their users.

To summarize:

EvoEvo exploits up-to-date results in evolutionary biology and experimental evolution to propose new evolutionary technologies. These new technologies will be based on three main concepts (1) indirect selection of the most efficient fitness landscape structure, (2) indirect selection of highly evolvable genotype-to-phenotype mapping based on network-encoded phenotypes and (3) development of new mutational operators inspired from chromosomal rearrangements in order to evolve rapidly and efficiently the organisms structure. Ultimately, these technologies will be tested in the context of smart-houses and smart objects, two domains that will crucially need software to adapt to the context and users.

B1.2 Progress beyond the state of the art

EvoEvo: A disruptive approach in evolutionary computation

Bio-inspired computation is an old field that contains many approaches ranging from artificial neural networks to artificial immune systems or evolutionary algorithms [Floreano & Mattiussi, 2008]. All these approaches take their inspiration from life science. However, they also stay at a reasonable distance from biology to avoid integrating in the algorithms the whole complexity of living beings. The central idea behind this trend is that the fundamental properties of living systems lies in a small set of rules and that the main goal of bio-inspired computation is to discover and implement these rules in an artificial system. As far as evolutionary computation is concerned, this trend leads to a huge simplification of the genotype-to-phenotype mapping (the extreme situation being Evolutionary Strategies in which the phenotype actually *is* the genotype) and, correlatively, it put the stress on the variation operators that concentrate the complexity and the "intelligence" of the algorithm. This trend has produced highly efficient algorithms (such as CMA-ES [Igel *et al.,* 2006; Igel *et al.,* 2007]) but it is limited to applications in which the problem at stake is stable enough to allow the variation operators to adapt to the structure of the fitness landscape.

EvoEvo is based on a completely opposite idea. It assumes that complexity of the genotype-to-phenotype mapping helps, mainly because it creates a large set of possible evolutionary routes that can be followed to adapt quickly to new situations. In this view, complexity is actually necessary to enable efficient evolutionary dynamics. To some extent, this idea can be compared to Genetic Programming, to Neuro-Evolution or to Developmental Evolutionary Algorithms. However these approaches are theoretical constructions that are only scarcely connected to real evolution and to real organisms. Moreover, none of them succeeded to approach the capacity of real organisms evolution, showing that there are still many decisive points to learn from biology (and to learn *in* biology!) in order to design more efficient evolutionary algorithms. EvoEvo aims at reviving the basic link between evolutionary biology and evolutionary computation. However, contrary to most attempts in this domain, to turn our sight towards the organisms that are the most efficient evolutionary systems: viruses and bacteria.

Many authors have called for a closer to biology approach of evolutionary computation [Banzhaf *et al.*, 2006; O'Neill *et al.*, 2010; Correia, 2010]. EvoEvo will beneficiate from a unique trans-disciplinary consortium ranging from "pure" biology to "pure" ICT with complementary background and domain of expertise. **The wide disciplinary base of EvoEvo's consortium makes it ideally suited to propose disruptive approaches in computational evolution. These approaches will be based on the most up to date knowledge in microbiology and evolutionary biology, including the extraordinary adaptation capabilities of microbes as they have been recently witnessed by many experimental evolution results.**

EvoEvo: A disruptive approach in evolutionary biology

EvoEvo focuses on the extraordinary ability of microorganisms to generate diversity and novelty at a pace that was unthinkable only a few years ago. This ability has been highlighted by many experimental evolution results, with different organisms and in different experimental frameworks (reviewed in [Hindré *et al.,* 2012]). However, its evolutionary roots are only partly understood and the snatches of explanation provided by the analysis of the experimental results contradict our intuition. Indeed, it has been shown that the rapid evolution of microorganisms is based on a profound rewiring of biological networks followed by compensatory mutations [Philippe *et al.,* 2007; Cooper *et al.,* 2008] and that genome reorganization through chromosomal rearrangements plays a crucial role in the whole process [Blount *et al.,* 2012]. Comparison between distant species provides an idea of the plasticity of biological networks but not about how dynamic and constrained they are. With our unique combination of biological and digital materials, we will be able to measure the impact of chance and history in the evolution of genomes and networks, to identify their molecular bases and to decipher between generic properties and specific events occurring in a given lineage. Moreover, **by focusing on the process by which organisms – and especially here microorganisms – are able to generate diversity and novelty (and to *control*, at the population level, the level of generated diversity), EvoEvo will produce many theoretical, practical and technical results in evolutionary biology and microbiology. These results will be very helpful in many domains including bioprocess management, biodiversity and, last but not least, the arms race engaged with microbial pathogens.**

As far as modelling is concerned, evolutionary biology has a long-lasting affair with modelling. However, in evolutionary biology, models can hardly be validated since data are almost inaccessible. Moreover, most of these models cannot account for indirect selection since they assume that the genotype-to-phenotype mapping is a fixed parameter that cannot change during the evolutionary process. EvoEvo will substantially increase the power of evolutionary models by allowing the mapping to evolve and by comparing the model results with the results of *in vivo* evolution experiments. As far as we know, such a direct comparison between *in silico* and *in vivo* setups has never been achieved on any model, thereby providing a considerable step forward in the usability of *in silico* models. As stated by John Maynard-Smith: "*So far, we have been able to study only one evolving system […]. If we want to discover generalizations about evolving systems, we will have to look at artificial ones."* [Maynard-Smith, 1992]; See also [Lenski, 2001; O'Neill, 2003]. However, 20 years later, artificial evolution is still marginal in biology, probably because available models and organisms are not close enough to real organisms to be interpretable in terms of "real biology". This is one of the challenges we will tackle in EvoEvo by joining experimental and computational evolution. One of EvoEvo's returns will be the development of a simulation platform dedicated to the study of indirect selection. This platform will be validated by comparison with *in vivo* evolution experiments and the *in silico* evolution. **By developing new evolutionary models in a close collaboration between biologists and computer scientists, EvoEvo will provide new experimental models that will allow evolutionary biology to study questions it classically overlooks owing to the absence of efficient tools. These questions include the effect of indirect selection of genomes and networks, the contribution of chromosomal rearrangements to microevolution or the emergence of polymorphism in homogeneous environments through niche construction.**

EvoEvo: A disruptive approach in software engineering

A major challenge to build complex adaptive software systems is that current software engineering techniques require software architects to foresee all possible situations that the system may have to face. However, the rapid spreading of pervasive computing, including the development of smart objects, software/hardware agents, smart buildings and cities, creates a complex digital ecosystem in which predicting and modelling all situations is impossible. In this context, biology provides a rich source of inspiration to develop adaptive software systems that will constitute the ground of future living technologies. Indeed biological inspiration has been at the origin of many algorithms and meta-heuristics and, in the recent years, considerable progress has been done in the domain of bio-inspired artificial intelligence (see [Floreano & Mattiussi, 2008] for a general review). However there is still a considerable gap to fill before programs will be able to adapt to their still-evolving environment.

EvoEvo is based on the central idea that today's approaches in bio-inspired evolutionary computation are limited because their degree of freedom is strongly limited by the fixed set of rules that are used to translate the genotype into a phenotype. In other words, current approaches distinguish the software from the data it manipulates. In the so-called "meta-heuristics", the software is fixed and only the data are evolvable. In EvoEvo, we will get inspiration from microorganisms to overcome meta-heuristics approaches and develop a

evolutionary software meta-design in which the evolvable software is an integrated system in which all the organization levels are evolvable, exactly as, in a microorganism, every biochemical component is likely to mutate, change and, ultimately, evolve. The central idea of this evolutionary software meta-design is that evolution must act, in addition to the software function, the rules that trigger the software evolution. This meta-design idea has been proposed to overcome difficulties in complex system design [Doursat, 2008] or in software engineering [Weimer *et al.,* 2009]. **EvoEvo will be the first attempt to use evolutionary software meta-design to build an autonomous piece of software that will be able to control its own evolution in order to continuously adapt to its condition of use. Ultimately, evolutionary software meta-design will enable software to find its own usefulness in the ever-changing digital ecosystem that is the future of ICT.**

We foresee that evolutionary software meta-design will create an important disruption in current software engineering practices. Rather that anticipating on all the aspects of the software usage, software engineering will directly concentrate on the aim of the software and on its ecological niche in the digital ecosystem. Then, the flux of data the system will receive from its environment will trigger the evolution of specific behaviour that can be oriented towards specific purposes through selection rules. While in operation the software will, in turn, produce new data that will modify the local environment, thus enabling its integration with the other pieces of software at work. Once this loop will be stabilized, the software will be adapted to its environment. Now, since the rules of evolution will still be active, it will be able to dynamically change and adapt to any change of its environment. Living beings continuously adapt to their environment and constitute an unavoidable source of inspiration to build a digital ecosystem. However, the modes of adaptation are highly variable in the different living kingdoms. Higher eukaryotes, in particular, dynamically adapt through learning, a process that is, in its essence, independent of molecular evolution. Although evolving learning architectures is a promising research direction, EvoEvo is based on the idea that evolving technologies must start from simplest life forms for which molecular evolution is at the heart of adaptation capacities. **The EvoEvo transdisciplinary consortium constitutes a unique opportunity to develop evolutionary software systems that will get their inspiration from the organisms that are the most evolvable and that have been able to adapt and colonize virtually all ecological niches on earth: bacteria and viruses. By taking inspiration from microorganisms EvoEvo will provide a first step in this long-term vision of evolutionary software engineering.**

B1.2.1 Specific contribution to progress in science and technology

The development of living technologies needs an interdisciplinary approach well balanced between life sciences and technologies and Information and Communication Technologies. EvoEvo is based on state of the art scientific background in both domains. As far as computer science is concerned, EvoEvo is deeply rooted in artificial evolution, machine learning and unconventional computing:

- Artificial evolution is the evident root of EvoEvo, but our approach is disruptive, in the line of position papers by Banzhaf and colleagues [Banzhaf *et al.*, 2006], Correia [Correia, 2010] or O'Neill [O'Neill *et al.*, 2010]. Indeed, EvoEvo incorporates microbiology, evolutionary biology and molecular biology into evolutionary algorithms

that will provide artificial evolution with new approaches and concepts related to indirect selection.

- Machine Learning focuses on the ability of artificial systems to learn from data, and classically uses a stable dataset to train the learning algorithms. More recently, a strong interest for dynamic data (*e.g.* data-streams) has emerged owing to the development of many new sources of data (social and economic data, sensor networks…). EvoEvo will go one step further by focusing on non-stable data-streams in which the data-model changes over time (due *e.g.* to the evolution of sensor technologies).
- Unconventional computing aims at developing new computation schemes that are able to overcome classical limitations of computational systems. EvoEvo is directly linked to this research field since it aims at producing software systems able to self-program and self-adapt to their environment.

As far as life science is concerned, EvoEvo is deeply rooted in evolutionary biology, microbiology and evolutionary computational biology.

- One of the most salient characteristics of EvoEvo is its focus on microorganisms. EvoEvo will benefit from the huge amount of data from microbiology to develop a challenging Systems Biology and dynamic approach that will provide concepts to include in our algorithms, mandatory knowledge to analyse the results of evolutionary experiments, and novel insights into the structure and, more importantly, the forces governing the dynamics of microbial genomes, networks and populations.
- Evolutionary biology provides the core theory of evolutionary computation. However, evolutionary computation overlooks many evolutionary biology concepts. Evolutionary biology is a very active field that highlighted the mechanisms underlying the evolutionary processes. EvoEvo will integrate most up-to-date knowledge and use both experimental evolution and modelling to understand the rules behind the dynamics of these mechanisms themselves.
- Evolutionary computational biology produces models of evolution. On an algorithmic point of view these models are rather close to evolutionary computation [Adami, 2006; Mozhayskiy & Tagkopoulos, 2012; Hindré *et al., 2012*]. However, they are closer from the biological objects they model. EvoEvo will develop and use new evolutionary computational models to study indirect selection at the genome, network and population levels. Ultimately, these models will be applied to the development of the computational framework that will be used to develop EvoEvo's applications.

EvoEvo will strongly benefit from the collaboration of well-recognized teams in all these research fields that together will form an optimally suited consortium for achieving its objectives. Figure 4 summarizes the scientific background of EvoEvo and the domain of expertise of the partners.

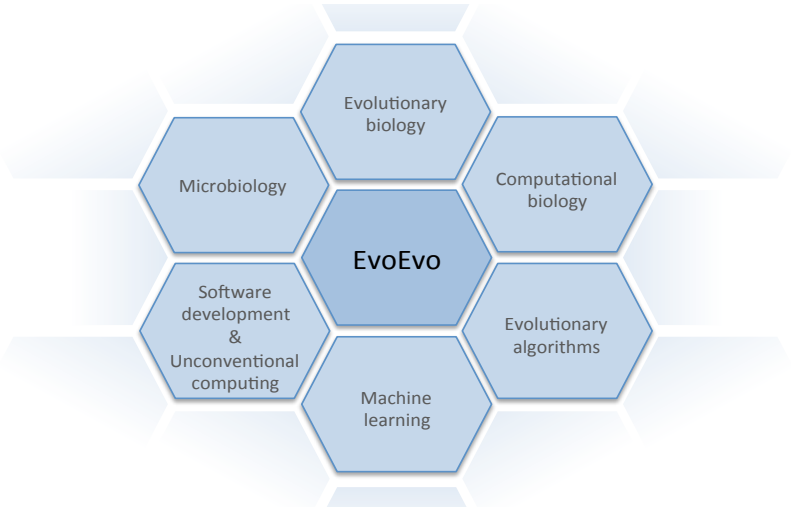


Figure 4 - Scientific background of EvoEvo

B1.2.2 Contribution to information and communication technologies

Evolutionary computation

Classical evolutionary algorithms, with their pre-defined, fixed genome representations and mutation operators, can produce novelties only within that limited representation. Their genome content can evolve, but not the representation because it is constrained by the predefined interpretation rules that constitute the genotype-to-phenotype mapping. Even Genetic Programming, where the representation is itself a computer program, is limited by the fixed terminal functions and simple program structure and cannot escape these constraints because the interpretation of the genotype into a phenotype is too straightforward. Evo-devo algorithms [Hornby & Pollack, 2001; Hornby, 2004] break out of this limitation somewhat, by allowing a genome to develop, through a variety of generative or growth rules (for example, L-Systems), into a richer phenotype. But even here, classical evo-devo algorithms are still confined to a single (albeit much richer) representation, and the developmental rules themselves are fixed. Even schemes that attempt to evolve the growth rules do so in a separate evolutionary domain.

Despite their potential to evolve startlingly novel solutions to optimization problems, classical evolutionary algorithms are forever constrained by their pre-defined structures and operations. Hence, there is a fundamental lack of scalability of these algorithms into more complex spaces. The origin of this limit lies in the closeness of genotype and phenotype representations. Since both are closely linked in evolutionary algorithms, the genotype can only hardly develop its own – phenotypically independent – structure and, when it does so, the consequence are seen as deleterious (like in code-bloat development). Indeed, evolutionary algorithms get their biological inspiration in the Neo-Darwinian theory and, in the case of Evo-Devo algorithms, in the developmental theory. But the mostly ignore the huge knowledge of molecular evolution, thus neglecting the indirect effects of the complex genotype-to-phenotype mapping on the course of evolution. In classical evolutionary algorithms, the data in the genome is used to express the phenotype. In biological organisms the data in the genome also encodes the bio-molecular machines involved in the processes of transcription, translation, metabolism, replication and even mutation. This closely coupled

feedback loop is missing from classical evolutionary algorithms, while it confers to biological organisms their incredible evolutionary capabilities.

More restrictive still, in these algorithms, the evolutionary operators are outside the scope of evolution itself. Some authors have added novel, bio-inspired evolutionary operators, such as gene-duplication operators [Kuo *et al.*, 2006; Leier *et al.*, 2007], but these are still externally imposed, not emergent properties of the evolving system. Others have discovered emergent macro-mutations [Hickinbotham *et al.*, 2010], yet their systems do not have the evolutionary capability to allow these emergent properties to be captured, encapsulated, and exploited by the system: the macro-mutations are *instances* of special one-off events, and cannot be captured as new *types* of event.

EvoEvo moves beyond this current state of the art, by incorporating molecular processes that support the evolution of evolution into the computational algorithms. Moreover, in order to do so, EvoEvo gets its biological inspiration from organisms that are the world-champion of evolution: bacteria and viruses. The complex molecular machinery that enables microorganisms to live and evolved has to be incorporated in the EvoEvo algorithms, using computational mechanisms that are analogues of abstractions of biological processes. This involves several software challenges. Genome representations, evolutionary operators and genotype-to-phenotype mapping all have to be encoded, explicitly or implicitly, in computer code. Making these flexible, mutable, *evolvable* involves making the code itself evolvable.

Mimicking microorganisms to invent new living technologies involves developing a complex artificial chemistry that the “organisms” will use to translate their genetic code into their phenotypic software. This can be done using a virtual machine that provides the “physics” of the evolving system. In this virtual machine, organisms are software segments. Evolution can emerge as a consequence of reproduction and mutation operators embodied in the operational mechanisms of the computational platform [Nellis & Stepney, 2011]. The central interest of this approach is that it enables the emergence of open-ended evolution [Ray, 1992] and, in some cases, of robust genotypic representations [Wilke *et al.,* 2001]. However, this approach still lacks generality: Since the phenotype is directly deduced from the genotype (the latter is the static code while the former is the execution of the code), the translation process cannot evolve and the genotype interpretation is thus fixed.

The EvoEvo will introduce a more general approach of artificial chemistry by taking its inspiration from the molecular structure of microorganisms. Indeed, in these organisms, all the molecular elements are involved in highly complex biochemical networks and the function of a given element cannot be deduced from its genetic sequence. Each molecular element of the organism gets its function from the interactions it is involved in within the network, thus providing evolution with a virtually infinite set of network structures, element functions and, ultimately, genotype-to-phenotype mapping. This approach can be viewed as close to the “neuro-evolution” approach [Stanley & Miikkulainen, 2002; Clune *et al.,* 2011] or to Artificial Genetic Networks evolution [Mattiussi & Floreano, 2007; Marbach *et al.,* 2009]. Yet, a crucial difference is that, in these approaches, the genomes and the phenotypes still exist in a different realm than the laws that translate the former into the latter. In EvoEvo, the artificial chemistry (*i.e.* the biochemical network) will include all these aspects, allowing evolution to act at all links of the chain.

A first approach is to consider biochemical networks as data-structures encoded in the genome, the interpretation of the later being dependent on the former (thus creating the loop that is at the heart of EvoEvo). Although the conception of such software will need a lot of work to translate the knowledge from the biological domain to the ICT domain, its implementation will be relatively straightforward and can directly benefit from the EvoEvo models that will be developed previously in the course of the project.

A more general approach would be to consider that the code that implements the behaviour of the biochemical networks belongs to the evolutionary domain. In this approach, the code itself can mutate and acquire new – unforeseen – properties. This will constitute the ultimate step of EvoEvo, highly risky but also with the opportunity to provide the maximal evolvability/open-endedness. In this second step, EvoEvo will provide a general purpose programming environment that supports fully self-modifying code, either at a low level assembly language level of artificial chemistries [Hickinbotham *et al.,* 2011], or at a higher level of agent based systems [Stepney & Hoverd, 2011] using advanced reflective programming approaches [Maes, 1987].

By providing a flexible evolutionary software platform developed by analogy with state-of the-art biological processes, that incorporates emergent evolutionary processes, EvoEvo would be a first in Computer Science.

Machine learning

Machine learning is seen here as the general research domain aiming at studying and developing ways to learn from data. Note that we will not debate on the sometimes-subtle differences raised between machine learning, data mining and knowledge discovery in databases, and use the term machine learning to encompass them as a whole domain. In the early period of development of the machine learning field, most learning processes studied relied on the hypothesis that all data used in the learning stage where at the learner disposal in some storage system. The corresponding research effort resulted in a deep understanding of the learning in such situations, using targeted representations as for instance the so-called theories (*e.g.* set of formulas in inductive logic programming), models (*e.g.* decision trees in classification) and concepts (e.g. concept learning in description logics). In the last decade of the 20th century, the ever-growing amount of collected data, and at the same time the fact that they were collected over long periods, revealed new issues and opportunities (*e.g.* [Widmer & Kubat, 1996]). Handling this kind of data, commonly termed data streams, gave rise to new challenging settings. In the extreme cases, the data stream is a high rate never-ending stream that cannot even been stored completely and the system can only keep track of summarizations of the data. Learning in this context means that a piece of data can be read and used a single time, and then only summaries of the past data can be used in the process. While the first difficulty was to manage such a dynamic input in the learning task, then a second major concern was the dynamic nature of the underlying model to be learnt. Indeed, in many cases, the models (or theory or concepts) ruling the observed stream are subject to changes over time. So, the problem is no longer only to learn a model and to refine it (*e.g.* enhancing the accuracy of a classifier) as new data arrived in the stream, but also, according to the dynamic modification of the underlying model, to learn from an evolving world. For instance, if classes or clusters appear or are not appropriated any more and are split, merged, or simply disappear, the system has to detect

such changes and to update its own model [Du, 2010; Aggarwal *et al*., 2006]. An even more challenging context is when, for a given data stream, the way of collecting the data is also changing. This is frequently the case, due to the availability of new probes/sensors and/or new processes to collect information, resulting for example in new descriptors of the data appearing in the stream, while some others could turn out to be obsolete and have disappeared. So, to learn from a data stream, a system must be able to adapt its current model to the changes of the underlying model but also to the modification of the kind of data collected in the stream. Facing and studying this learning situation as a whole is still a major open problem in machine learning. The evolution of evolution studied in EvoEvo offers a great potential of application by developing learners that can adapt quickly to changes, while preserving the quality of their models.

Software development and unconventional computing

Conventional classical computation is characterized by the deterministic, discrete, sequential, pre-specified paradigm inherent in the Turing model and von Neumann architecture. Originally designed to model the behaviour of human clerks carrying out well-specified calculations [Copeland, 2008], the Turing model faces difficulties when applied to naturally stochastic, continuous, massively parallel, ill-defined, or open-ended problems. One goal of unconventional computation is the development of non-standard computational models that can overcome some of these limitations. Biology is one domain mined by unconventional computation, since (from a computational perspective) it studies physically-embodied evolved open-ended information processing [Stepney, 2008].

Biology has been a rich source of computational algorithms and architectures, including: artificial neural networks, evolutionary algorithms, artificial immune systems, ant colony optimization, artificial biochemical and signalling networks, L-systems and generative grammars, and many more. However, all these computational schemes have been designed and implemented within the Turing model, with pre-specified closed search spaces and well-defined essentially static fitness landscapes.

At the opposite extreme, unconventional computation studies unconventional substrates: performing directly embodied computation with, for example, chemical reaction-diffusion systems [Adamatsky *et al.,* 2005], biological cells, slime moulds [Adamatsky, 2010], quantum dots, and even (although still only in theory) curved space-time [Hogarth, 1992]. Although promising, many of these approaches are still in their infancy, not yet supporting large-scale computation.

An intermediate route is to exploit the Turing model as the underlying substrate on which is implemented a virtual machine supporting an unconventional computational model (in a manner analogous to the way physical electronics is the underlying substrate supporting the implementation of the Turing model as a virtual computational machine). Although losing the potential efficiency of a fully embodied implementation, it has the advantage of exploiting a mature technology of hardware and software development tools. One such virtual machine is that provided by artificial chemistries [Dietrich *et al.,* 2001]. The computational agents are analogues of chemicals (molecules), with a virtual physics providing spatial mixing, molecular binding, and the semantics of chemical reactions. Implicitly defined rules of binding and reaction potentially allow open-ended chemical construction. “Soft” binding and execution

properties [Hickinbotham *et al.,* 2010] allow a degree of stochasticity, hence mutation, and so can support the evolution of chemical constituents. By encoding as much of the reproduction and error correction mechanisms as possible into the active information-carrying molecular agents, rather than into the fixed physics, such a platform can also support the emergent evolution of evolutionary mechanisms themselves.

By providing an unconventional computational artificial chemistry platform developed from state-of-the-art biological models, specifically to support the evolution of evolutionary processes, EvoEvo would be a first in Computer Science and Artificial Life.

B1.2.3 Contribution to life science

Microbiology

Microbiology is at the heart of our human society whether one talks about species extinction, human impact on complex ecosystems, fighting antibiotic resistance or effects of social networking. The last century witnessed major developments that resulted in a transition in the scale of investigation from the elucidation of individual and separate components of microbial cells (molecules, genes, pathways…) to an integrated view of microbial cell function. EvoEvo has been designed to integrate all this knowledge to go even a step further in the Microbiology research field, in particular to be able to fully understand the dynamic properties of bacterial cells. Such a complete understanding is requested for both a better use of microorganisms for our human needs and a significant improvement in our fighting abilities against microbial pathogens.

The combination of microbiology with genetics and molecular biology allowed the isolation and analyses of mutant strains, resulting in the discovery of the molecular toolbox [Jacob, 1977]. Pathways in microbial physiology, biochemistry, metabolism and gene regulation were dissected resulting in an exquisite understanding of their individual components [Neidhardt, 1996]. The advent of massive genome sequencing technologies applied to thousands of microbial species, including hundreds of *Escherichia coli* isolates, revealed the enormous diversity of the microbial world [Touchon *et al.,* 2009]. Metagenomics further confirmed that the challenge was not anymore to obtain data but to analyse them [Gilbert & Dupont, 2011]. The expression of entire microbial genomes is now qualitatively and quantitatively estimated owing to technologies and tools such as transcriptomics, proteomics, metabolomics, and fluxomics [Han *et al.,* 2011; Güell *et al.,* 2011], uncovering the organization of cellular networks. Moreover, a further transition emerged in the scale of analyses, from the population to the individual-cell level. Indeed, the development of fluorescent reporters together with microscopic and microfluidic devices favoured multi-disciplinary approaches with the fields of mathematics, physics and computer science [Locke & Elowitz, 2009; Ahmed *et al.,* 2010], that further revealed a high diversity at the single-cell level including phenotypic variability [Snijder & Pelkmans, 2011]. All these developments are integrated into Systems Biology, a multi-disciplinary approach devoted to the analysis of the multifaceted complexity of microorganisms.

Both conceptually and empirically, EvoEvo will shed a new light into the development of Systems Biology. Indeed, two main limitations have to be overcome for a complete understanding of the emergence and dynamics of all the genomic patterns including genome

organization, metabolic and regulatory network architecture. First, the intrinsic diversity of the microbial world calls for generic principles that can be applied to any species, including human pathogens and industrially-modified strains devoted to human needs. Second, most Systems Biology studies tend to focus on the static description of genomic patterns within "model organisms" and even on a reference strain of these species. Comparison between distant species provides an idea of the plasticity of cellular networks but not a clear idea of how dynamic and constrained such networks are.

EvoEvo will provide innovative and integrated evolutionary perspectives that are needed to relate the dynamics of adaptive changes to the phenotypic and genotypic landscapes of microorganisms. In particular, it will rely on experimental evolution strategies that provide the missing dimension of systems dynamics. During the last years, experimental evolution provided insights into bacterial adaptation and emphasizes the potential of microbial metabolic and regulatory networks to evolve [Hindré *et al*., 2012]. EvoEvo will include both an evolutionary framework and an integrated approach to relate genomic and regulatory changes to fitness, robustness and evolvability. Ultimately, such a fully integrated evolutionary framework should provide general principles for microbial adaptation, as well as global laws that link evolutionary processes and organismal structure. In particular, we will use this integrated approach to understand the evolution of the evolutionary processes themselves that is the most generic and global approach. Therefore, microbiology can only be considered as being linked to other disciplines like mathematics, computer science, physics, and this integration level will warrant our ability to develop novel innovative ideas combining basic and applied research, and exploiting the wonderful complexity of microorganisms.

Evolutionary biology

The question of "evolution of evolution", often called "indirect selection" or "second-order selection" is an open question in evolutionary biology. In particular, the question of the evolution of evolvability and its relation to the evolution of robustness has received important contributions in the last years. However, the question is still widely opened.

The word evolvability has been used in two different ways [Wagner, 2005b]. First, *a biological system is evolvable if its properties show heritable genetic variation, and if natural selection can thus change these properties*. A second usage links evolvability to evolutionary innovations: *a biological system is evolvable if it can acquire novel functions through genetic change; functions that help the organism survive and reproduce*. Both definitions apply to all levels of biological organization, from macromolecules like RNA and proteins, metabolic pathways, gene regulation networks, to macroscopic traits and whole organisms. In consequence, innovation comprises many different levels, including enzymes with new catalytic activities, novel complex organs, organisms able to colonize new niches, or even new species. The two usages are far from being equivalent. Most importantly, not all systems that are evolvable in the first sense are necessarily evolvable in the second sense. Every system that is evolvable in the sense of being innovative can evolve by means of natural selection. In other words, the ability to innovate is the more profound usage of evolvability as it encompasses the first usage and much more.

Biological systems, from macromolecules to whole organisms, are robust if they continue to function, survive, and reproduce when faced with mutations, environmental change, and internal noise [De Visser *et al*., 2003]. Living organisms are exquisitely complex, yet also highly robust to genetic change on all levels of organization. For example, genome's organization can protect them against deleterious mutations, proteins can tolerate a considerable number of amino acid changes, metabolic networks can continue operating even after removal of important chemical reactions, gene regulation networks continue to function after alteration of key gene interactions, and radical genetic change in embryonic development can lead to an essentially unchanged adult organism [Knibbe *et al.*, 2007a; Rennell *et al*., 1991; Edwards & Palsson, 2000; von Dassow *et al*., 2000]. The more robust a system is, the more mutations in it are neutral, that is, without phenotypic effect.

After reading the above definitions, one may conclude that both concepts are antagonistic: in a highly robust system, mutations will have smaller phenotypic effects than in a less robust (brittle) system. Hence, robustness reduces the amount of heritable genetic variation on which selection may operate and thus, negatively affecting evolvability. In this sense, brittle systems may be more evolvable than robust ones. However, this vision is limited, since neutral mutations may also play a fundamental role in innovation: although many mutations in a robust system do not change its primary function, they can change other system features. In other words, a system fulfilling its primary function in many different configurations is flexible to adopt other features. Similarly, many of these neutral mutations may be revealed upon environmental changes, thus allowing for faster adaptation (*e.g.* [Rutherford & Linquist, 1998; Queitsch *et al.,* 2002]). From this perspective, neutral mutations themselves are the key to evolutionary innovation: robustness implies that many mutations are neutral and such neutrality fosters innovation. Computational and simulation work with Avida [Elena & Sanjuán, 2008] has revealed that robustness may hinder evolvability at the short run but promotes it at the long run by allowing the exploration of large neutral genotypic networks. Similarly, RNA folding simulation studies [Wagner, 2008] has shown that genotypic robustness and evolvability show an antagonistic relationship. However, phenotypic (structure) robustness promotes structure evolvability. Both studies conclude that finite populations of genotypes with a robust phenotype can access large amounts of phenotypic variation while spreading through a neutral network. Population-level processes and phenotypes rather than individual sequences are key to understand the relationship between robustness and evolvability.

All in all, the relationship between robustness and evolvability remain an open question not for the lack of models and theories but for the lack of empirical data that may contribute to confirm or reject some of them. EvoEvo will tackle the tension between evolvability and robustness from a multidisciplinary perspective. At the one side, precise genetic manipulations of two well established biological model systems (an RNA virus and a bacterium) combined with a powerful experimental evolution approach will allow for a direct test of the relationship between genetic (broadly speaking: genomic, network and population) robustness, environmental robustness, phenotypic innovations and evolvability. At the other side, computer models will allow generalization of the data gathered *in vivo* in order to facilitate the connection of these empirical observation with theoretical models proposed in the literature or to produce new theories when necessary. The outcomes of EvoEvo will

represent a major jump on our understanding of the mechanisms by with phenotypic innovations and adaptations occur in biological systems.

Evolutionary computational biology

Since the pioneering work of Haldane [Haldane, 1930-32], Wright [Wright, 1932; Wright, 1984] and Fischer [Fisher, 1930] population genetics has been the leading modelling approach about evolution. This powerful mathematical theory focuses on adaptive (and later also neutral [Kimura, 1977]) processes and their dependence on population size (and structure) and selection regimes. It does, however, largely simplify away from the structure of the evolving entities, and the evolution of the structure, assuming simple linear mapping between genetic information and phenotype.

From a modelling point of view two major lines of research have broken with this tradition, and both use structural information to study evolution.

Firstly, metabolic network theory, and in particular Flux Balance Analysis (FBA) have exploited stoichiometric constraints to study metabolism, and has been remarkably successful in scaling up to the complexity of real metabolic networks and in predicting the outcome of gene knockout experiments (*e.g.* [Papp *et al.*, 2004] and of evolution (*e.g.* [Pal *et al.,* 2006; van Hoek & Hogeweg, 2007]). FBA can do so by considering optimal flux through the network in equilibrium. Recent extensions include gene product expression at a genomic scale [Lerman *et al.*, 2012].

The second comprehensive conceptual and modelling framework to break away from this tradition has been developed taking RNA evolution as a paradigm system [Fontana & Schuster, 1987; Schuster *et al.*, 1994; Huynen & Hogeweg, 1994]. RNA combines genotype (the nucleotide sequence) and phenotype (its folded structure which determines its function) in one molecule. Moreover, the folding, up to secondary structure is computable. By studying the properties of the resulting genotype-to-phenotype mapping and therewith fitness landscape, many novel insights in evolutionary processes have been derived. Most prominently, the structure of this realistic landscape has many remarkable features that enhance evolvability. It contains percolating and intertwining neutral paths, in an otherwise very rugged landscape [Schuster *et al.*, 1994; Huynen *et al.*, 1996]. Because of this structure, only a relatively short evolutionary walk leads to a large variety of phenotypes, and given a phenotype, still allows for exploration of genotype space. It has been shown that on such landscapes evolution automatically leads to robustness, *i.e.* to regions in genotype space with higher than average neutrality [Huynen & Hogeweg, 1994; van Nimwegen *et al.*, 1999]. This enhanced robustness implies (counter-intuitively) also enhanced evolvability, as it leads to larger genetic variability in the population due to drift on the neutral network. Moreover it has been shown that drifting along the neutral network the population encounters many new phenotypes in its close mutational neighbourhood, thus enabling innovation [Huynen, 1996; Wagner, 2012b]. Finally, in a spatial and interacting setting, it has been shown that that niche construction and speciation occur and feed back on the genotype-to-phenotype mapping [Takeuchi & Hogeweg, 2008].

This resulted in the conclusion that "RNA is an ideal evolvable molecule" [Schuster *et al.*, 1994]. However, it is not "just RNA". It turns out that all subsequently studied biological

genotype-to-phenotype mapping share the above-mentioned structural and evolutionary properties to a certain extend. This is studied for protein folding (although less pronounced [Ferrada & Wagner, 2012], metabolic networks [Wagner, 2012b] and gene regulatory networks [Ciliberti *et al.*, 2007b]. Note however that it is not straightforward to design coding structures that define such genotype-to-phenotype mapping, and in the various evolutionary computation approaches they have as yet not been used. Powerful as the RNA model paradigm has been in providing novel insights in evolution, it is still quite restricted in its degrees of freedom, relying heavily on the prior given physical-chemical folding process, which defines the overall landscape, and evolution can only "select" subspaces of the prior defined landscape. In present day organisms, the genotype-to-phenotype mapping involves many layers of transformation and regulation which all are the product of evolution, and subject to evolutionary change even at the short term. Only a few models are able to explore the further consequences [Hindré *et al.*, 2012]. However initial exploratory work indicates that increasing the degrees of freedom has further profound effects. For example, it has been shown that coding structure evolve to structure the mutational neighbourhood beyond a general property as neutrality to specific properties such that random mutation become biased to favourable mutation [Crombach & Hogeweg, 2007; Crombach & Hogeweg, 2008; Draghi & Wagner, 2009], that genome size shows a specific pattern of expansion and streamlining which enhances evolvability and robustness [Knibbe *et al.*, 2007b; Cuypers & Hogeweg, 2012], and that alternative strategies exist with respect to robustness: apart from increased neutrality, increased lethality may evolve under very high mutation rates and or large genomes [Cuypers & Hogeweg, 2012; Krakauer & Plotkin, 2002]. Moreover, interesting phenotypic variation through noise [Kaneko, 2011] and plasticity [Espinosa-Soto *et al.*, 2011] is shown to enhance rapid evolutionary adaptation.

Phylogenetic reconstructing and experimental evolution have suggested similar patterns, *e.g.* large ancestral genomes (*e.g.* [Wolf *et al.*, 2012; Smarda *et al.*, 2008]) and very fast adaptation to changing environments, involving large-scale regulatory (*e.g.* [Ferea *et al.*, 1999]) and/or chromosomal changes (*e.g.* [Dunham *et al.*, 2002]).

The above-sketched developments in evolutionary theory are as yet in their infancy, and the uncovered processes are likely the tip of the iceberg of the phenomena that still have to be discovered and explained. The EvoEvo project aims at further developing this emerging fundamental evolutionary theory, and merge it with the exciting developments in experimental evolutionary biology, which can now be analysed at the genome, and regulome levels and in phylogenetic reconstruction. Moreover EvoEvo will link these insights up with those of evolutionary computation. Along these lines we expect the proposed research will advance fundamental insights in biological evolution, as well as its technical counterpart.

B1.3 S/T Methodology and associated work plan

B1.3.1 Overall strategy and general description

EvoEvo aims at achieving the objectives stated in section B1.1 through a workplan containing 6 workpackages: WP6 management; WP1 *in vivo* experiments; WP2 model design; WP3 *in silico* experiments; WP4 computational framework design; WP5 applications. Figure 5 illustrates the essential characteristics of the project and the way their declination in the workplan ensures that the global objectives are achieved.

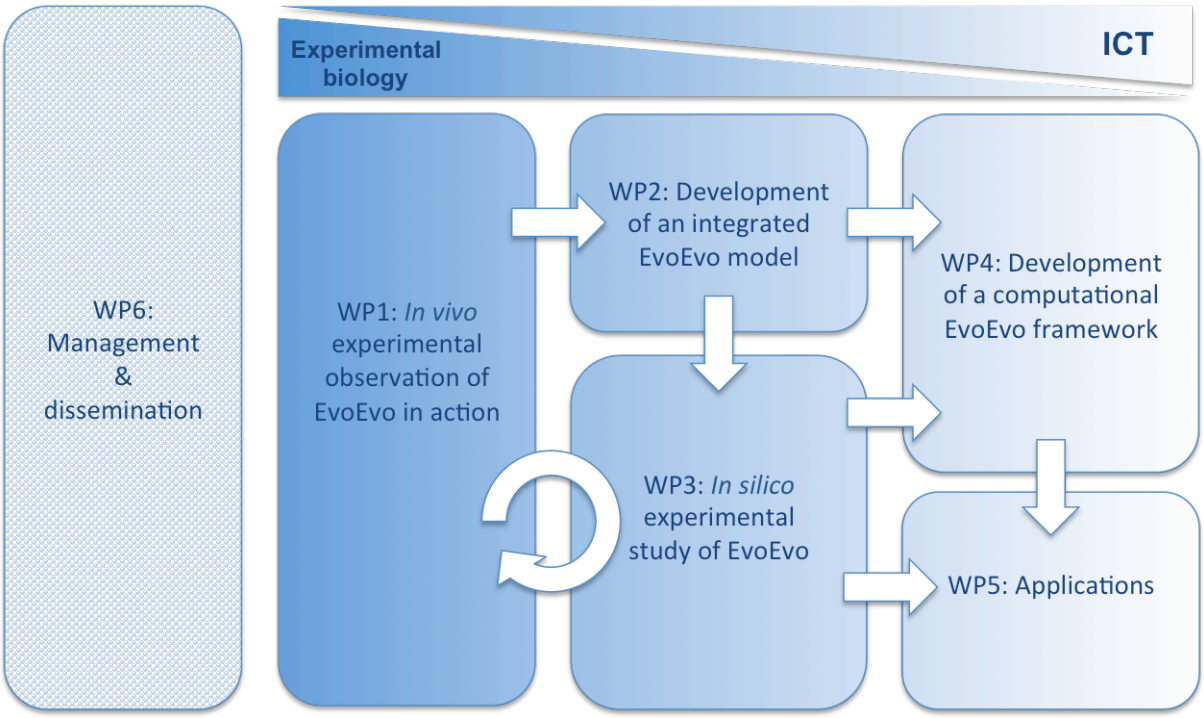


Figure 5 - EvoEvo workplan overview

EvoEvo workplan is based on three principles that guaranty that the biological foundations of EvoEvo are effectively and efficiently transmitted to the computation application through modelling and framework development steps. The three principles are the following:

- **Principle 1: A route from evolutionary biology (WP1) to artificial evolution (WP4) through modelling (WP2).** The workplan is organized to guarantee a continuity research and development from "wet" experimental biology to ICT applications. It benefits from the pivotal role of modelling: models are computational artefacts that mimics the phenomenon observed *in vivo*. They thus constitute an intermediate step between life science and ICT. Now, models must not be mistaken for applicative code. Their objectives are - and must stay - clearly different. That is why the transition from life science (WP1) to applicative code is done in two steps. The models developed in WP2 will be reinterpreted to develop a computational EvoEvo framework (WP4) that will benefit from them but that will also introduce simplification and/or generalization of the model's bio-like structures.
- **Principle 2: Parallelism between *in vivo* experimental evolution (WP1) and *in silico* experimental evolution (WP3).** One of the most salient features of EvoEvo's workplan is the complementarity between WP1 (*in vivo* experimental observation of EvoEvo in action) and WP3 (*in silico* experimental observation of EvoEvo in action). Models of evolution are very difficult to validate through direct comparison with *in vivo* experiments. The organization of WP1 and WP3 is such that the same experiments can be done in both frameworks (Figure 6). Then we will be able to compare the results of the two experiments and thus (1) to validate the computational models before they are used as a basis for the conception of the computational framework (WP4) and (2) to propose generalizations of the

results of "wet" experiments. These generalizations will help proposing evolutionary hypotheses that will later on be used to enhance the applicative performances of EvoEvo in WP5.

- **Principle 3: Applicative targets (WP5) that will make profit from both the computational framework designed in WP4 and from EvoEvo knowledge produced in WP3.** To ensure that EvoEvo will produce results that fulfil target a) of the EVLIT call (*Empirical, theoretical and synthetic approaches that define the key bio-inspired principles that can drive future living technologies and the environment to use them in a controlled way*), we will develop proof-of-concepts applications that will demonstrate the power of EvoEvo. This will be done in WP5 through the development of two target applications in the context of smart buildings and Internet of things. The workplan is organized such that WP5 benefits from both WP3 (that will provide knowledge about the most interesting phenomenon susceptible to increase the evolution capacities of the applications) and WP4 (that will provide the computational framework used for the target applications).

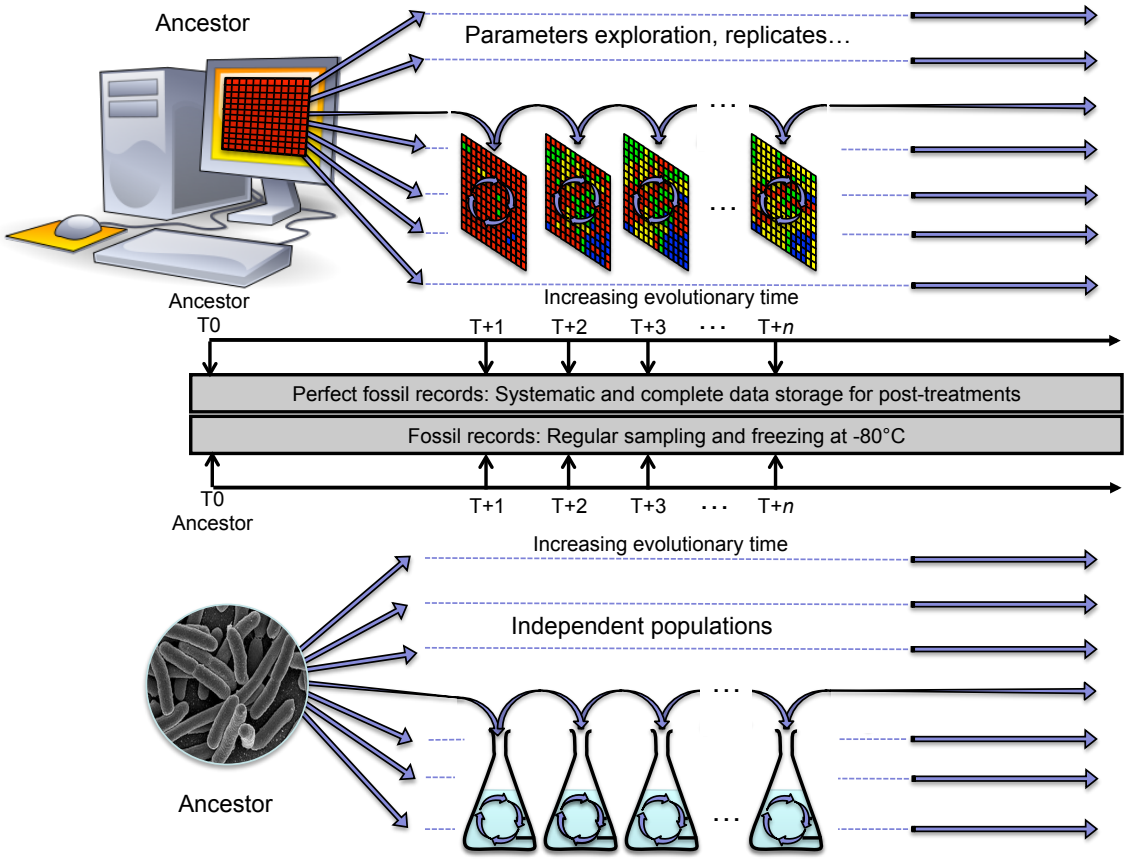
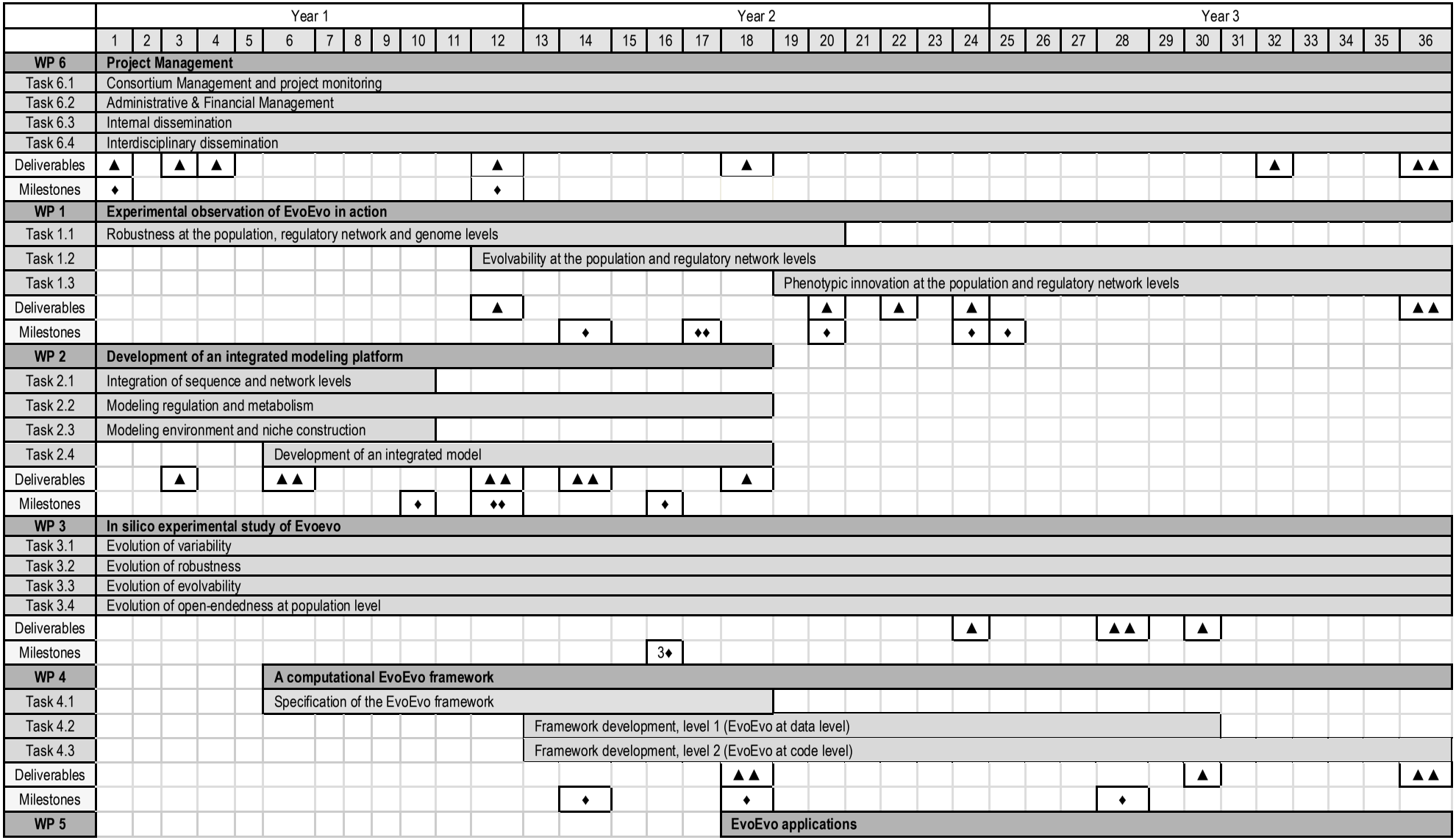


Figure 6 – Parallel *in silico* (top) and *in vivo* (bottom) experimental evolution. The experiments conducted *in vivo* will be precisely reproduced in the computational modelling framework.

B1.3.2 Timing of work packages and their components

The following table shows the global timing of the project for each workpackage (see also table WT1) and for each task (month 1 being the month that starts at the start date of the grant agreement). The date of delivery of the project deliverables and milestones are reported on the table (please see tables WT2 and WT4 for a detailed description of deliverables and milestones).


Year 1
Year 2
Year 3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
WP 6
Project Management
Task 6.1
Consortium Management and project monitoring
Task 6.2
Administrative & Financial Management
Task 6.3
Internal dissemination
Task 6.4
Interdisciplinary dissemination
Deliverables
Milestones
WP 1
Experimental observation of EvoEvo in action
Task 1.1
Robustness at the population, regulatory network and genome levels
Task 1.2
Evolvability at the population and regulatory network levels
Task 1.3
Phenotypic innovation at the population and regulatory network levels
Deliverables
Milestones
WP 2
Development of an integrated modeling platform
Task 2.1
Integration of sequence and network levels
Task 2.2
Modeling regulation and metabolism
Task 2.3
Modeling environment and niche construction
Task 2.4
Development of an integrated model
Deliverables
Milestones
WP 3
In silico experimental study of Evoevo
Task 3.1
Evolution of variability
Task 3.2
Evolution of robustness
Task 3.3
Evolution of evolvability
Task 3.4
Evolution of open-endedness at population level
Deliverables
Milestones
3♦
WP 4
A computational EvoEvo framework
Task 4.1
Specification of the EvoEvo framework
Task 4.2
Framework development, level 1 (EvoEvo at data level)
Task 4.3
Framework development, level 2 (EvoEvo at code level)
Deliverables
Milestones
WP 5
EvoEvo applications

	Year 1												Year 2												Year 3											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Task 5.1																		An open stream cluster analysis																		
Task 5.2																		Evolvable Personal Companion																		
Deliverables																																				▲▲
Milestones																											♦					♦				

B1.3.3 Graphical presentation of the components showing their interdependencies

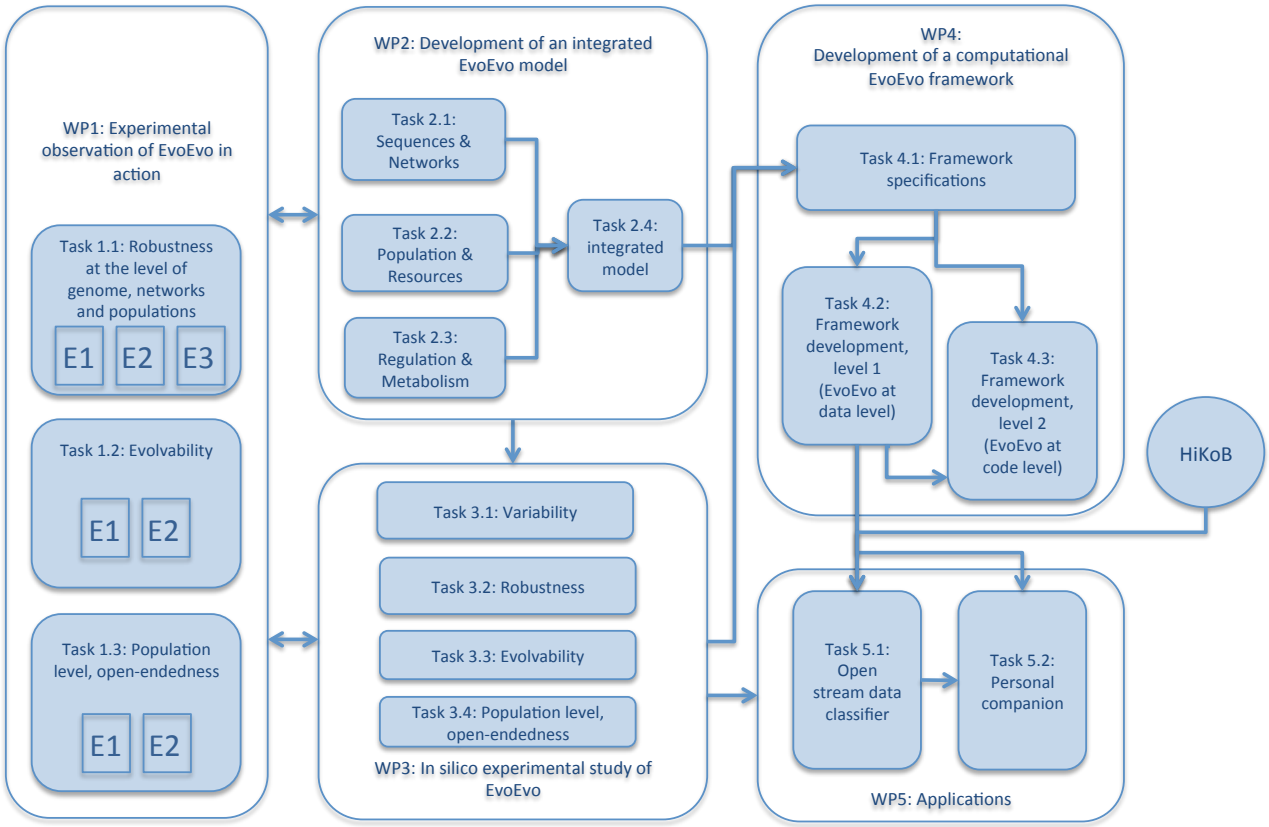


Figure 7 - Workplan components and interdependencies

Figure 7 describes the relationships between the different workpackages and tasks in the workplan. We also display a box called “HiKoB” to show the input of the subcontracted part (sensor network). Arrows represent dependencies between the workpackages and the tasks. Arrows linking workpackages indicate that the whole result of the workpackage will be used (or is used). The figure illustrates the main characteristics of the workplan:

- **Close integration of WPs 1, 2 and 3**.
  - WP1 will use two model organisms (*E. coli* and TEV). Both organisms have been extensively studied by UJF and CSIC and others. This knowledge from WP1 will be used in WP2 to design the models in tasks 2.1, 2.2, 2.3 and 2.4.
  - Models designed in WP2 as well as their condition of usage (*e.g.* range of acceptable parameters) will be used in WP3 to launch long-term evolution that will produce *in silico* “wild-type” strains.
  - Experiments performed in WP1 on *E. coli* and TEV will be modelled in WP3 by implementing mid-term *in silico* evolutionary experiments using the wild-type strains.
  - The comparison between the results of WP1 and WP3 will be used to develop a theoretical conceptual framework of EvoEvo.

- **Organization of WP2**. Building an integrated model of EvoEvo promises to be one of the most difficult and most risky parts of the project. To reduce the risks, the integrated model will be designed in two steps.
    - Starting from the initial knowledge of Inria and UU, tasks 2.1, 2.2 and 2.3 will produce models integrating parts of the functionality (respectively genome and network, metabolism and resources, realistic network structure). All these models will be tested individually and used for some experiments in WP3.
    - The complete model will be constructed from the know-how produced in tasks 2.1, 2.2 and 2.3. This will reduce the risk inherent to task 2.4 as well as giving as alternative plan to be sure to have enough material to start WP4 in time.
- **The development route of WP4**. Developing a coherent bio-inspired computational approach is difficult and can only be achieved through a precise workplan in order to avoid the risk of confusing biological-contingent details with the underlying principles [Stepney et al., 2005; Andrews et al., 2011]. WP4 will follow the CoSMoS methodology (figure 8) to avoid this risk by:
    - Using as its input an evolvable, tested, computational model rather than the whole biological knowledge.
    - Sharing information between designers of the computational model and designers of the computational framework in order to build, first, a computational meta-model (through abstraction and translation of the biological and modelling knowledge and software) and second, a computational model that will instantiate the choices in a precise, programmable system. These two steps will be performed in task 4.1.
    - Developing two frameworks of increasing complexity (tasks 4.2 and 4.3). The first one will implement evolution of evolution at the data level (*i.e.*, the “organisms” are be represented as data-structures that can (1) evolve and (2) reorganize themselves in order to “evolve to evolve”). This will be a manageable step forward the computational model. Thus the risks are here minimal. That is why this framework will be used in the applications of WP5. This first framework will be developed in task 4.2. The second framework will implement evolution of evolution at the code level (*i.e.,* the “organisms” are represented by data-structures and by the code that manipulates them. Thus both the data-structures and the code can evolve and “evolve to evolve”). This direction is highly risky but it will prepare the future steps of evolving technologies. Task 4.3 will concentrate on this aspect of EvoEvo. Its output will be one of the final outputs of the whole project (the reason why there is no output arrow from this task).
- **The development of proof-of-concepts applications in WP5**. WP5 intends to develop applicative software that will be effectively usable in real conditions. We will tackle smart-houses and smart-objects problem in two steps: Task 5.1 will tackle the problem of model adaptation to analyse a data stream when the data placed in the stream come from a network of sensors that changes over time (sometimes faster than the sensed system itself). Task 5.2 will tackle the problem of behavioural adaptation for personal companion that will use the data produced by the sensor network. The input of WP5 will come from three different sources:

- We will use a real sensor network that will be deployed in “smart-room(s)”. This activity will be provided through subcontracting with the HiKoB company.
- The evolvable software will be the one produce by task 4.2 (evolvable framework at data-level).
- EvoEvo knowledge and knowhow will come from WPs 1 and 3. It will enable efficient analysis and parameterization of the applications in order to ensure that EvoEvo is at work in the applications and to analyse its consequences on the behaviour of the system.
-

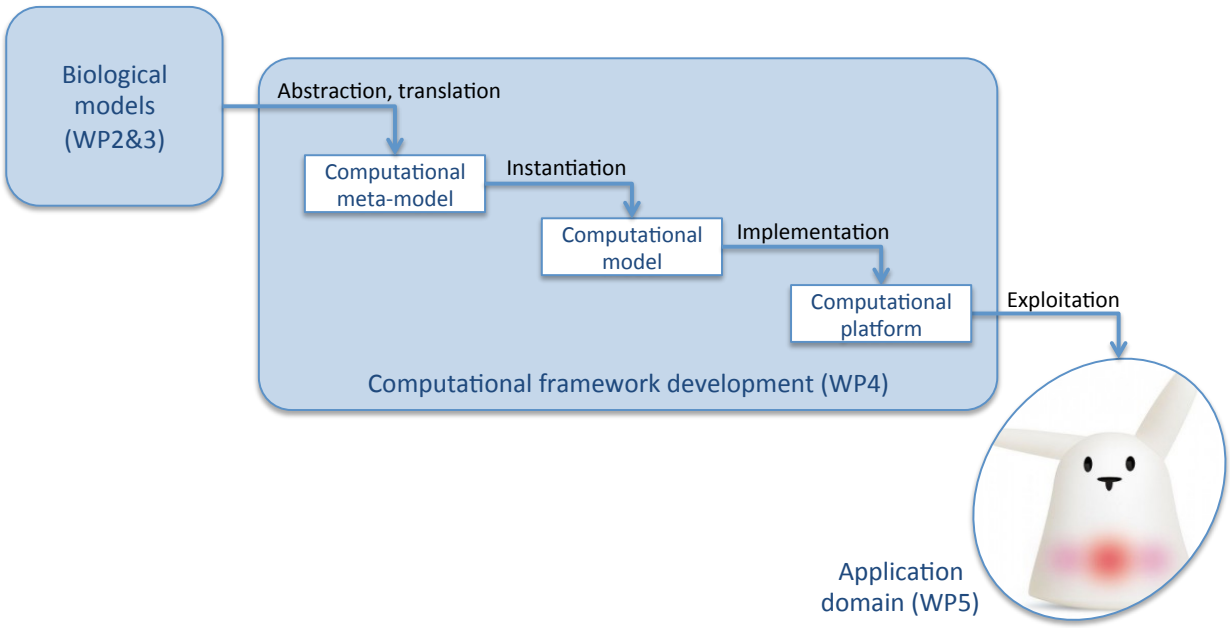


Figure 8 – CoSMoS methodology used for the development of the computational framework (inspired from [Stepney et al., 2005; Andrews et al., 2011])

B1.3.4 Risks and associated contingency plans

The consortium has, prior to the start of EvoEvo, identified the main risks of the project. These are presented in the table below and have been integrated in the workplan design (see previous section). Risks management will be performed throughout the whole project duration in order to control the identified risks and to update the risk analysis. Risk management will be included as part of the regular periodic project management reports. This periodic review will ensure early warnings of potential risks and that appropriate correction activities are performed.

WPs involved	Explanation of risk	Probability	Impact	Mitigation plan
WP1 and WP3	All tasks of WP1 are wet evolutionary experiments that are inherently risky. There	Low. All methods used in the proposed experiments are	The concerned experiment will not produce usable results	(1) Analysis of the failure will be carefully done by UJF and CSIC in order to understand it in the

	is always a risk that the analysis of the evolved clones shows no exploitable changes	well mastered by UJF and CSIC[1]. However, experimental biology always has an inherent risk associated.		context of EvoEvo. (2) WP3's results will be compared with existing literature with the help of UJF and CSIC
WP2	Cannot evolve the genome-network model (Fitness does not improve for all tested parameter sets)	Low considering the knowhow of partners INRIA and UU	Very high. The genome-network model is the core of the integrated model	Organize a global project meeting to simplify the model. This could be done by focusing on one of the existing models (aevol or pearls-on-a-string).
WP2	Cannot evolve the environment model (Fitness does not improve for all tested parameter sets)	Low considering the knowhow of UU	Very high. The environment structure will enable niche construction that is the core of open-endedness.	The environment model will be replaced by an existing model [Takeuchi *et al.,* 2011] that has been shown to enable open-endedness.
WP2	Cannot evolve the realistic network model (Fitness does not improve for all tested parameter sets)	Medium: The realistic network model will rely on complex parameters and need a high level of computation resources	Medium: The realistic network model will be used to study the evolution of variability	Many formalisms have been proposed in the literature to model networks with different degrees of realisms. INRIA and UU will collectively choose a lower-level model that could be used in task 3.1.
WP2	Cannot evolve the integrated model (Fitness does not improve for all tested parameter sets)	Medium considering the risk management performed in tasks 2.1, 2.2 and 2.3.	High: the integrated model cannot be used as the input of WP4	Organize a global project meeting to simplify the model. This could be done by focusing on one or two of the models produced in tasks 2.1, 2.2 or 2.3.
WP3	All tasks of WP3 are *in silico* evolutionary experiments that are inherently risky. There is always a risk that the analysis of the evolved clones shows no exploitable changes	Low to Medium depending on the experiments	The concerned experiment will not produce usable results	(1) Analysis of the failure will be carefully done by INRIA and UU with the help of UJF and CSIC in order to understand it in the context of EvoEvo. (2) The corresponding experiment of WP1 will be analysed on its own.

[1] Tasks 1.1, 1.2 and 1.3 involve various technologies, most of which are available and well mastered by partners UJF (bacterial experimental evolution, gene inactivation, growth cultures in microtiter plates using a Tecan microplate reader, transcription profiling, competition experiments, allelic exchange experiments, biochemical and molecular analyses, analyses of polymorphic populations including growth curves and competition experiments under negative frequency-dependent selection, genome re-sequencing and analysis) and CSIC (experimental evolution of TEV populations, fitness determinations, analyses of TEV genome sequences, characterization of symptoms induced upon TEV infections, transcription profiling of infected and non-infected plants).

WP4	Cannot transpose the evolutionary models into a computational meta-model	Low considering that a first transposition step has already been done in WP2	Communication issues between life-science partners and ICT partners. Technical issues to reproduce EvoEvo in applicative software	The consortium will focus on the lower complexity models in order to select the minimal properties to be implemented in the meta-model
WP4	Cannot transpose the computational meta-model into the first level computational framework	Low considering that the first level framework will be close to the models	High: The development of the proof-of-concept application is not possible	Simplify further the computational framework by reconsidering the computational meta-model
WP4	Cannot transpose the computational meta-model into the second level computational framework	High owing to the ambitious objectives of task 4.3	One of the final results of the project cannot be achieved	Reconsider the objective of task 4.3 by isolating parts of the software that can evolve and parts that cannot
WP5	The first experimental validation (online data-stream classifier) is not convincing	Medium	One of the final results of the project cannot be achieved	Validation of the project on a simplified dataset
WP5	The second experimental validation (evolving personal companion) is not convincing	High owing to the highly ambitious objectives of task 5.2	One of the final results of the project cannot be achieved	Use the computational framework to evolve agent behaviour on a classical benchmark

B2. IMPLEMENTATION

B2.1 Management structure and procedures

EvoEvo is a research-focused project that gathers 5 partners with complementary expertise. The size of the consortium enables an **efficient structure** that **monitors progress continuously** and **acts rapidly** on any matter.

B2.1.1 Project Organizational Structure

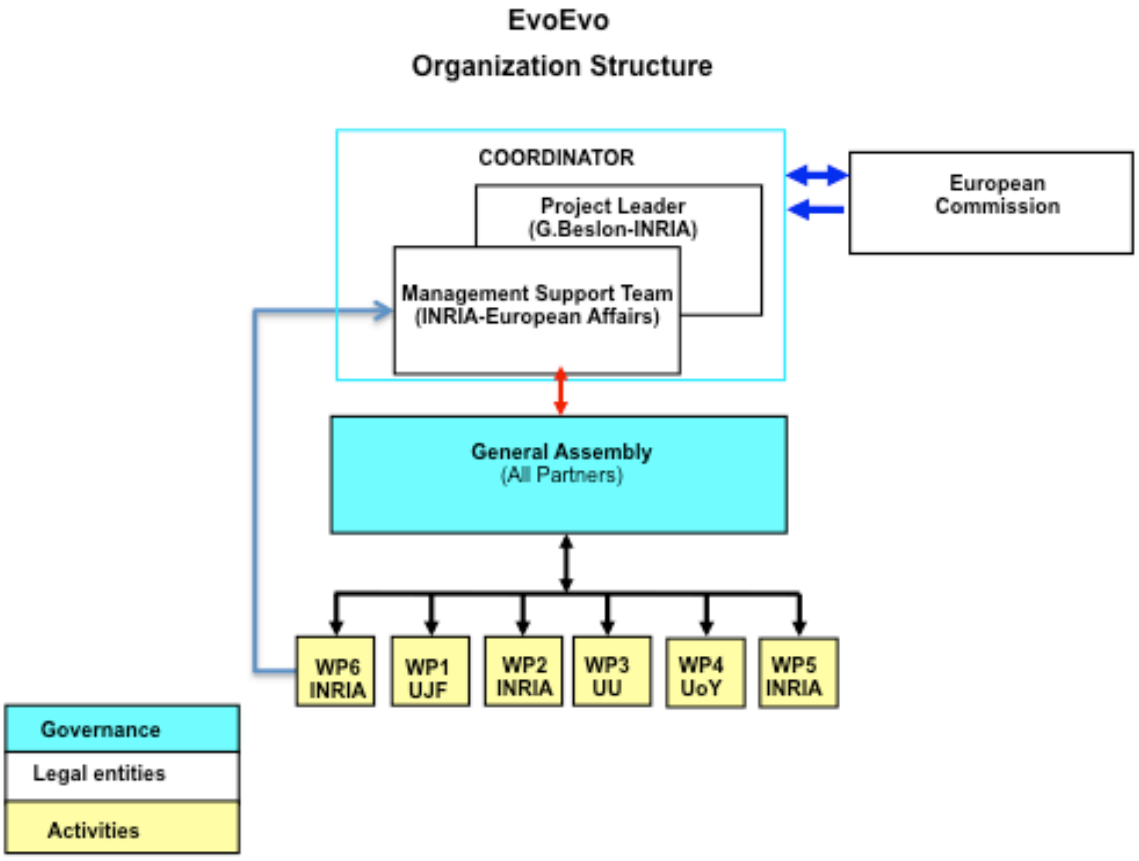


Figure 9 - EvoEvo's governance structure

INRIA, and more specifically Guillaume Beslon, as Project leader will be in charge of Administrative and technical Management of the project.

For the Administrative management, he will be assisted by the INRIA European project team and the workpackage leaders will support the technical side.

B2.1.2 Roles

INRIA is responsible for the day-to-day execution of the project and will ensure the timely delivery of project objectives and deliverables by continuously monitoring how closely project progress is following the plan. Guillaume Beslon from INRIA, as Project Leader will work with the Work Package Leaders to identify issues and propose suitable corrective actions (*e.g.*

resource reallocation, taskforce creation, etc.) that might require approval by the Executive Board, which he chairs. He is responsible for calling Executive Board meetings and reviews as well as compiling and distributing minutes and actions. He defines the procedures for change control (proposed changes to the plan), risk management and quality. Finally he is also responsible for the administrative management of the project that includes the provisioning of periodic reports and financial statements, as well as interacting with the Management Support TEAM of INRIA to ensure efficient distribution of EU funding. Moreover this team will be able to offer support in the regular administrative and legal tasks with regard to the project and the European Commission.

The main roles are to:

- Define high-level technical strategy and drive the project team to implement according to that strategy.
- Ensure that the project maintains its scientific and technological objectives, its relevance to the ICT program and its strategic objectives
- Lead the technical discussions at the project level and try reaching resolution of conflicts by consensus.
- Keep regular contact with the partners ensuring that the project stays on track.
- Interface with the European Commission side of the project administration: reviews, reports, etc.
- Hold regular meetings every 6 months with the Executive Committee for financial and administrative management.
- Maintain accurate records of costs, resources and time scales for the project.
- Pull together information for the management reports and arrange for their delivery to the European Commission.
- Determine project communication strategy and employ the tools required to implement that strategy.
- Coordinate the preparation of annual reviews and reports: specifically, administrative and financial effort, cost tracking, and official reporting to the Project Officer.

Executive Committee:

The Executive Committee is the main decision making team. It is formed by the leading participant from each partner, who can be represented by some member of the same partner if explicitly appointed. Since several Work Packages are lead by a different partner, all the Work Packages are represented in the Executive Board. The Executive Board will be **responsible for the overall direction of the project and any strategic decisions**. Each partner will have one vote. In case of a tie, the vote of the **chairperson (INRIA**, as Coordinator) decides, and the EC is notified. No major modification will be made to the project without EC approval. The Executive Committee will **meet at least twice a year in addition to a monthly conference call.** The detailed description of roles and responsibilities of the Executive Board will be defined accurately in the Consortium Agreement, which will be based on DESCA – The simplified FP7 model Consortium Agreement supported by the European Commission.

Members of the Executive Committee will be:

- Mr Guillaume Beslon - INRIA
- Mr Dominique Schneider - UJF
- Mrs Paulien Hogeweg - UU
- Mrs Susan Stepney - UoY
- Mr Santiago F. Elena - CISC

Workpackage Leaders

- WP1: Leader UJF (D. Schneider)
- WP2: Leader INRIA (G. Beslon)
- WP3: Leader UU (P. Hogeweg)
- WP4: Leader UoY (S. Stepney)
- WP5: Leader INRIA (G. Beslon)
- WP6 (Management): Leader INRIA (G. Beslon)

The roles of the Work Package Leaders are to:

- Coordinate, monitor and manage the activities and tasks under their responsibility, and to ensure the timely achievement of the objectives and milestones of the work packages.
- Ensure the timely production of the deliverables and their quality.
- Meet with the coordinator and arrange regular technical meetings every 6 months (meetings will be combined with the Executive committee meetings to save travel times and costs).
- Ensure the accurate recording of times, costs and resources, and report any discrepancies immediately to the project leader/coordinator.
- Prepare the internal and external reports expected for the workpackage, and assist in the production of the overall management reports of the project
- Organize technical presentations of the work package activities, and to ensure proper involvement and visibility of the active technical members.
- Inform the Executive Committee about progress of activities and possible critical issues.

In addition to the scheduled meetings, the project leader will promote asynchronous conference calls and face-to-face meetings between partners if required.

B2.1.3 Risks

During the project preparation a risk assessment activity has been performed and risks for each work package have been identified and potential risk mitigation actions have been proposed and integrated to the workplan.

As part of the normal operation of the project an updated risk analysis will be performed every 6 months and will be included as part of the regular periodic project management reports. This periodic review will ensure early warnings of potential risks and that appropriate correction activities are performed.

Addressing impact risks is part of the normal operation of the project. Such risks will be addressed in periodic management meetings. The Project leader will report on risks and issues to the Executive Committee and will keep a Risk / Issue Log for the project as well as assign actions or contingency plans to be executed as required, so as not to impact the overall outcome or objectives of the project. In any case, the EC Project Officer will be adequately informed either directly by the Project Leader or through the regular management reports.

B2.1.4 Information Flow

The exact structure of information flow in the consortium will be set up at the beginning of the project, following the kick-off meeting.

Some of the information flow procedures that will be used are:

- A project web server will be set up by the coordinator to act as a repository for all the working documents, minutes, reports and deliverables. The server will have a public part with a generic description of the project, copies of the public deliverables, and any documents that are declared as public by the consortium. It will also contain a password protected project private part.
- Internal communications will be made via e-mail correspondence, project meetings, and by web/telephone conferences. Contact information will be recorded in the project web server, so as to be available to all project partners and EU.
- Management reports will be prepared in a structured fashion and be kept in the project web server. The main content will be the progress reports and management reports from the work package leaders. The project coordinator will integrate this information and produce the necessary periodic reporting for the EC, and make the information available to them.

Minutes will be kept of all the formal project meetings, and kept on the project webserver.

B2.1.5 Conflict Resolution

The project work includes many different kinds of activities in which an effective and efficient, lightweight decision process is required. The methods for reaching an agreement over an issue to be decided will mainly be by technical discussions, in person during the project meetings, and by email and telephone conferences in between. They will be followed by written confirmations in the form of informal messages, document sections, or more formal meeting minutes, which everybody will be able to review for final approval. Agreement will be reached whenever possible by consensus. If this fails the coordinator can appoint a smaller group of participants to discuss the technical details in depth and deliver a more detailed technical document for review by the rest of the partners, possibly with proposed resolutions.

When a particular issue raises more contention, a position statement will be requested from the proposer, and partners will be invited to contribute with pros, cons, and alternatives.

Once everybody has had an opportunity to read and contribute the issue will be discussed. In the end, if consensus is impossible, simple majority voting will be used. In case of a tie, the coordinator vote decides.

Usually, for technical or scientific issues the above rules are enough. However, management or possible commercial aspects may require a different approach. Any Work Package Leader will report potential conflicts of this kind immediately to the coordinator. Should this happens the Project Leader will attempt to solve the issue by discussion, or if needed by calling a technical or administrative meeting of the partners. If the called meeting cannot resolve the conflict, resolution is tried at the Executive Board level. Any conflicts that cannot be resolved through the principles above will be handled according to the dispute resolution provision set forth in the Consortium Agreement.

B2.1.6 Summery of Project Organization

	Administrative Management	**Strategic Management**	**Executive Management**	**Operational Activities**
WHAT?	Project and financial reports Communication tools and dashboard	New orientations, conflict solving, corrective actions, budget allocation, projects results impact	Implementations of project through WP, inputs to administrative management.	Research and innovation, dissemination, training.
HOW?	Following instructions from the General Assembly, interacts with WP leaders for monitoring and reporting	Top-down decisions	Top-down decisions and bottom-up reporting to the General Assembly	Joint activities, meetings, …
WHO?	Management Support Team, Coordinator	General Assembly	WP Leaders	WP Leaders

B2.2 Beneficiaries

Participants No.	1	Organisation name	Institut National de Recherche en Informatique et en Automatique

Working at the crossroads of computer sciences and mathematics, over the last 40 years INRIA's researchers have been developing the scientific foundations for a new field of learning: computational sciences. When associated with other scientific disciplines, computational sciences can be used to offer new concepts, languages, methods and teaching aids which open up new avenues for exploration and understanding of complex phenomena. Working in project-teams, Inria researchers mix fundamental and applied research in an innovative blend to produce their results.
The BEAGLE INRIA research team gathers researchers from different domains (mainly computer science, bioinformatics, mathematics and computational biology). It's goals is to develop computational models of the cell in order to unravel the laws that govern its structure and behaviour. More precisely, the team focuses on the evolutionary origin of cellular structures (genome structures, networks structures) and on the biophysics law of diffusion in the cytoplasm and in the nucleus of the cells. All Beagle members are strongly involved in interdisciplinarity through collaborations with experimental biology labs both at the national and at the international levels.

No.	Collaborator Profile
C 1	**Prof. Guillaume Beslon** Guillaume Beslon is a full professor at the Computer Science department of the Institut National des Sciences Appliquées de Lyon (INSA-Lyon) and the leader of the INRIA Beagle team (that he created in 2011). He is a computer scientist whose fields of research are artificial life, modelling of biological processes and bio-inspired computation. Guillaume Beslon is strongly involved in interdisciplinarity collaborations with experimental biology labs for more than ten years. He has collaborations with teams working in molecular and cellular biology as well as teams working in evolutionary biology. He is also involved in the organization of interdisciplinarity proactive actions at the regional and national levels. He co-leads the Rhône-Alpes Institute of Complex Systems (IXXI, with the physicist P. Jensen).
C 2	**Dr. Carole Knibbe** Carole Knibbe is an associate professor at the Computer Science department of the Université Lyon 1, and a member of the INRIA Beagle team. Her contributions include the study of indirect selective pressures on gene number and on the amount of non coding DNA, and the indirect selective pressures on the organization of transcription and on the amount of chromosomal rearrangements. In 2011 and 2012, she coordinated a two-years CNRS interdisciplinary research project called "Analyse, simulate and experiment the evolution of bacterial genomes", involving five partners from microbiology, bioinformatics, mathematics and computer science.
C 3	**Dr. Christophe Rigotti** Christophe Rigotti is an Assistant Professor at INSA-Lyon (University of Lyon). He works at the LIRIS laboratory (UMR5205 CNRS) in the team DM2L (Data Mining and Machine Learning) and in the INRIA Beagle team. His main research interests include sequential pattern mining, analysis of sequences of events, and high performance analysis by condensed representations and constraint propagations. He has been involved as senior researcher or work package leader in several IST projects in the field of machine learning/data mining: CINQ European project (No IST-2000-26469), AEGIS European project (No IST-2000-26450) and IQ European project (No IST FP6-516169).
C 4	**Dr. Yann Gripay** Yann Gripay is an associate professor at INSA-Lyon. His work focuses on query processing in distributed dynamic environments (*e.g.* pervasive computing, ambient intelligence), for which he proposed the SoCQ framework, a unified data model combining traditional data with data streams and distributed web services. He is contributing to the "KISS" INS ANR Project (2012-2015) dedicated to innovative ways of securely managing personal data and documents through intelligent devices.

Participants No.	2	Organisation name	Université Joseph Fourier Grenoble 1

University Joseph Fourier Grenoble (UJF) is located in the Rhône-Alpes region, 2nd French region in terms of research. UJF is a research intensive university in an international high tech environment:

- 15,200 full-time students, excluding PhD students, 1500 lecturers/researchers, 1500 administrative/technical staff
- 50 laboratories in 4 core areas : Mathematics-Information sciences-Technologies-Communication; Chemistry-Life sciences-Health-Biotechnologies; Material sciences-Nanosciences-Engineering-Earth-Universe-Environment; Humanities
- major international and national research centres (ESRF, ILL, EMBL, CEA)
- many multinational firms in nano- and micro-electronics and biotechnologies (ST microelectronics, Hewlett-Packard, Becton Dickinson, Schneider Electric, Roche Diagnostics, BioMerieux)
- 3 major science parks (MINATEC, BIOPOLIS, MINALOGIC).

UJF has great experience in European framework programmes (FP). The first projects were submitted in 1990, during the 3rd FP. It has managed the participation in 69 projects under FP6 (2002-2006). Since the beginning of FP7, UJF laboratories have succeeded in 95 projects, 25 being coordinated by UJF, and 7 in the ERC program (three Starting Grants and four Advanced Grants).
EvoEvo will benefit from the expertise of the GEM team which develops the interface between basic, environmental and medical research by interdisciplinary collaborations, including with modelling and computer scientists, population geneticists and medical practitioners.

No.	Collaborator Profile
C 1	**Prof. Dominique Schneider** Dominique Schneider is a full professor at UJF in Grenoble and the head of the GEM team. He is a molecular microbiologist who has specialized in the evolution theory and its experimental investigation. The ecological and molecular mechanisms underlying bacterial adaptation are investigated using evolution experiments, including the longest-running experimental evolution using *Escherichia coli* that has been developed by Richard Lenski (Michigan State University, USA). The relationships between genome sequence and structure (mutation rates, large chromosomal rearrangements), genome expression (plasticity of regulatory networks, global gene expression) and the evolutionary dynamics of bacteria are investigated. It allows a better understanding of how natural selection is able to re-shape and improve entire genomes, and which functions are more plastic over evolutionary time. This evolutionary perspective is fully complementary to most Systems Biology approaches and addresses how evolvable are genomic and network features and what are the molecular bases of such evolvability. EvoEvo will enable to go a step forward by integrating a modelling framework in experimental evolution.
C 2	**Dr. Joël Gaffé** Joël Gaffé is an assistant professor in Biology at UJF in Grenoble and a member of the GEM team. His fields of research are oriented toward microbial molecular biology, genetics and evolution using experimental evolution. His main interests are related to the understanding of the dynamics of genome sequence and structure, including mutation rates, chromosomal rearrangements (inversions, deletions, duplications) and transposition of IS elements in the context of bacterial adaptation. The impact of genome dynamics on bacterial phenotypes (fitness, resistance to stress) is investigated.
C 3	**Dr. Thomas Hindré** Thomas Hindré is an associate professor at UJF in Grenoble and a member of the GEM team. He works on the dynamics of global regulatory networks during a long-term evolution experiment with *Escherichia coli.* The molecular mechanisms involved in the rewiring of regulatory networks during evolution are investigated with special interests on the link between bacterial chromosome topology and expression. Hence, nucleoid-associated proteins (NAPs)-encoding genes were repeatedly affected by mutations during bacterial adaptation. The impact of these mutations on the whole chromosome topology and on regulatory network architecture is investigated respectively by measuring local DNA superhelicity at different genomic loci and by quantifying disturbance of central regulons owing to these NAP mutations. This work will contribute to a better understanding of molecular mechanisms ensuring bacterial adaptation and eventually lead to identification of new targets for anti-bacterial compounds.

Participants No.	3	Organisation name	Utrecht University

Utrecht University (UU) is amongst Europe's leading research universities, renowned for both its innovative cross-disciplinary approach to research, and its emphasis on excellent quality in education. It is ranked first in the Netherlands and 12th in Europe according to the Shanghai ARWU 2012 ranking. Around 30,000 students study for their Bachelor or Master degree at Utrecht University. It hosts around 3,000 fte scientific personnel including PhD students and postdoctoral researchers. Each year over 7,000 scientific publications are published and over 450 PhD students obtain their degree. The university has much experience in European projects. So far in FP7, it participates in 77 projects in the cooperation programme, in 19 ERC projects and in 60 People projects. The Theoretical Biology & Bioinformatics group participates amongst others the International Training Network in quantitative immunology QuanTI.

The Theoretical Biology and Bioinformatics group at the Department of Biology, Faculty of Science at Utrecht University is one of the world-leading groups in the field of theoretical biology. The Theoretical Biology & Bioinformatics programme develops formal approaches of modelling and bioinformatics to better understand functioning and evolution of biological systems. The group combines large-scale computer simulations, mathematical modelling and bioinformatics with a strong biological background. The group has coined the term bioinformatics for the study of informatic processes in biotic systems and has pioneered many new approaches to handle and interpret the enormous complexity of biological systems and the massive amount of data that are available. These approaches include *e.g.* tree based multiple alignment, individual (agent) based modelling, spatial evolutionary models, and cell based modelling. Important contribution have been made *e.g.* towards the theory of multilevel evolution, functioning of the immune system and the dynamics of development.

No.	Collaborator Profile
C 1	**Prof. Paulien Hogeweg** Paulien Hogeweg has founded the Theoretical Biology and Bioinformatics group at Utrecht University, and is now Honorary Professor and PI in that group. She is a biologist, and a pioneer in computational biology complex systems research. She, together with Ben Hesper, coined the term Bioinformatics for the study of informatic processes in biotic systems. Important contributions include methodological developments (multiple sequence alignment, individual based modelling, cell based developmental models (CPM. aka. GGH (Glazier, Graner, Hogeweg) model but her main contributions are iin biological theory formation. She elucidated the role of self-organization in spatial eco-evolutionary models, as well in models of behaviour, and development. By using RNA folding as paradigm for non-linear genotype-phenotype mapping the evolutionary consequences of such mapping, as well as the consequences of evolution on this mapping have been revealed. Extending this paradigm, to include gene regulation, and metabolism, her current research focus is on unravelling the mutual interrelation between levels of organization, how these are shaped by evolution, and how they shape further evolution. This research has led to insights in long term information integration in evolutionary processes, which shapes the mutational profile, and therewith the propensity to variability, robustness, evolvability and innovation and therewith the evolution of evolution: the topic of this proposal. The EvoEvo project will enable us to go forward by integrating in vitro experimental evolution results in the formulation and the analysis of the in silico evolutionary systems, and by collaborating on model development with the Lyon group.
C 2	**Dr. Kirsten Ten Tusscher** Kirsten ten Tusscher is a PI in the Theoretical Biology and Bioinformatics Group at Utrecht University. She obtained her PhD in that group in 2004 on modelling electrical excitation in the heart. Her break-through model of human cardiac tissue (the TNNP) model is widely used. After a postdoc period in Utrecht and a junior group leader position in Oslo, she returned as a staff member to Utrecht in 2009, switching her research focus to her old love of evolution and development. She is currently studying models of the evolution of speciation processes and of developmental pathways, the common theme being the linking of micro and macro evolutionary patterns. Recent results include the demonstration that genome and network organization can evolve so as to facilitate population polymorphism despite obligatory sexual reproduction, and the demonstration that functional but NOT architectural modularity is associated with increased robustness and evolvability of developmental patterning networks.

Participants No.	4	Organisation name	University of York

The University of York, founded in 1963, has nearly 16 000 students and over 30 academic departments and research centres. York concentrates on strong viable departments, and teaching and research of the highest quality. York is one of the top ten universities in the UK for teaching and research – and is first in the UK and eighth in the world in the Times Higher Education world rankings of universities less than 50 years old.

The Non-Standard Computation Group (NSC), Department of Computer Science (led by Prof. Stepney) researches reality-based computing approaches that seek their inspiration in the natural world (biology, chemistry, and physics). This includes computational modelling of biological systems; the abstraction of biological principles into novel computational methods; modelling computational dynamical systems and their emergent properties; investigating novel physical substrates and their computational properties.

The **York Centre for Complex Systems Analysis** (YCCSA) provides an inspiring interdisciplinary research environment, bringing together a critical mass of 80 resident researchers and visitors from the Departments of Biology, Chemistry, Computer Science, Electronics, Management, Mathematics, and Physics. Since 2010 YCCSA has been housed in purpose-built facilities, demonstrating York's commitment to interdisciplinary research.

No.	Collaborator Profile
C 1	**Prof. Susan Stepney** Susan Stepney is Professor of Computer Science at the University of York, and leads the Non-Standard Computation Group. She is one of the founder members of the York Centre for Complex Systems Analysis (YCCSA), and since 1 Jan 2012 has been its Director. Originally a theoretical astrophysicist, she spent 18 years in commercial R&D, participating in and managing commercial research projects, including software development. On moving to academia in 2002, she began researching aspects of novel computation, and founded the Non-Standard Computation research group. She is chair of UKCRC Grand Challenge 7 in Non-Classical Computation, and has research interests in a range of non-classical computing approaches. She is a member of the EU Coordination Activity TRUCE "Training and Research in Unconventional Computation Europe". Her research interests include novel computational paradigms inspired by physical, chemical and biological processes. This involves input from domains as diverse as NMR, bacterial plasmids, and biochemical networks, used to inspire and validate broad conceptual models of *in materio* computation. This has more recently led to work in computational Artificial Chemistries, including the invention of "sub-symbolic AChems", and work on self-modifying and reflective software systems, both designed to support open-ended evolution in software. Additionally, her research interests include developing software systems related to biological processes. This covers (i) the development of flexible approaches to extracting and defining bio-inspired algorithms from observed biological process via a principled conceptual abstraction framework; (ii) the modelling, design, implementation and validation of rigorous and trustworthy scientific simulations, through the development of the CoSMoS (Complex Systems Modelling and Simulation) approach. EvoEvo brings together her interests in open-ended evolution in software, and bio-inspired software development.
C2	**To Be Nomitated Research Assistant** A research assistant will be appointed for the whole duration of the project. He/She will contribute to WPs 2,4,5 for a total of 36 person-months. We seek for a candidate with a PhD in unconventional computing, evolutionary computation of bio-inspired computation. Experience of bio-ICT interdisciplinary projects is absolutely essential. Some knowledge in evolutionary and molecular biology would be also positively considered. The candidate will take responsibility for the specification and development tasks of WP4 as well as for interactions with modelling experiments of WP2 and 3.
C3	**To Be Nominated Research Assistant**

	A research assistant will be appointed for the specific developments linked to WP4. He/She will contribute to WP4 for a total of 24 person-months. We seek for a candidate with a PhD in unconventional computing, evolutionary computation of bio-inspired computation. Experience of bio-ICT interdisciplinary projects is absolutely essential. Some knowledge in evolutionary and molecular biology would be also positively considered. The candidate will take responsibility for the development tasks of WP4 as well as for interactions with applicative developments of WP5.

Participants No.	**5**	**Organisation name**	Agencia Estatal Consejo Superior de Investigaciones Científicas

CSIC is the largest public institution dedicated to research in Spain and the third largest in Europe. Belonging to the Spanish Ministry of Economy and Competitiveness through the Secretary of State for Research, Development and Innovation, its main objective is to develop and promote research that will help bring about scientific and technological progress, and it is prepared to collaborate with Spanish and foreign entities in order to achieve this aim. According to its Statute (article 4), its mission is *to foster, coordinate, develop and promote scientific and technological research, of a multidisciplinary nature, in order to contribute to advancing knowledge and economic, social and cultural development, as well as to train staff and advise public and private entities on this matter.*
CSIC plays an important role in scientific and technological policy, since it encompasses an area that takes in everything from basic research to the transfer of knowledge to the productive sector. Its research is driven by its centres and institutes, which are spread across all the autonomous regions, and its more than 15,000 staff, of whom ore than 3,000 are staff researchers and the same number again are doctors and scientists who are still training. CSIC has 6% of all the staff dedicated to Research and Development in Spain, and they generate approximately 20% of all scientific production in the country. It also manages a range of important facilities; the most complete and extensive network of specialist libraries, and also has joint research units. Its multidisciplinary and multisectorial nature means CSIC covers all fields of knowledge. Its activity, which covers everything from basic research to technological development, is organised around eight scientific-technical areas. CSIC main functions are: Multidisciplinary scientific and technical research, Scientific and technical advice, Transfer of results to the business sector, Contribution to creation of technologically-based companies, Training of specialist staff, Management of infrastructures and large facilities, Promotion of the culture of Science, and Scientific representation of Spain at international level.

No.	Collaborator Profile
C 1	**Prof. Santiago Elena** PhD in Molecular and Evolutionary Genetics (1995 Univ. de València) working on experimental evolution of RNA viruses under the supervision of Pr. A. Moya and Pr. E. Domingo. He did a postdoc on bacterial experimental evolution (1996-1997, Michigan State University) under the supervision of Pr. R. E. Lenski. In 1998 he become Assistant Professor of Genetics (Univ. de València) and in 2001 Associate Professor of Population Genetics. In 2002 he moved to the Spanish National Research Council (CSIC). In 2005 he was promoted to CSIC Professor. Since then, he has been working on experimental evolution of plant viruses and viroids, artificial life, theoretical population genetics, and molecular evolution. In 2008 he was elected as External Professor of The Santa Fe Institute (NM, USA). Since 2009/11, he has been the director of IBMCP Department of Virology.
C2	**To Be Nominated postdoctoral researcher**[2]

[2] According to Spanish legislation, it is not possible at this stage to list names in the proposal since the position have to be publicly offered via CSIC job offers web portal and all potential candidates must be evaluated in equal conditions.

	A post-doctoral researcher will be appointed for the whole duration of the project. He/She will contribute to WP1. We seek for a candidate with a PhD in molecular virology or molecular evolutionary genetics and a strong motivation for combining experimental and computational work. Experience in basic molecular biology techniques (e.g., cloning, RT-PCR, qPCR, site-directed mutagenesis, sequencing) is absolutely essential. Some knowledge in programing (C++, Perl or Python) would be also positively considered. The candidate must be familiar with advanced topics in evolutionary genetics. The candidate will take responsibility for the experimental evolution work with TEV necessary to accomplish deliverables 1.1, 1.2, 1.3, and 1.5 (robustness and evolvability).
C3	**To Be Nominated Predoctoral researcher**[3] A PhD student will be appointed for years 2 and 3 of the project. He/She will contribute to WP1. We seek for a candidate with a master degree in molecular genetics and experience in plant virology. The candidate must be also familiar with advanced topics in evolutionary genetics and sequence analyses, including NGS data. The candidate must show a strong motivation for both experimental and computational work. Some knowledge in programing (C++, Perl or Python) as well as in the use of computer packages for statistical analyses (R or SPSS) would be desired. The candidate will be responsible under the supervision of Prof. Elena for the experimental work with TEV necessary to accomplish deliverables 1.4 and 1.6 (genome architecture). The candidate will also be responsible for analyzing the RNA-seq and Illumina sequencing data."

B2.3 Consortium as a whole

EvoEvo will beneficiate from a unique trans-disciplinary consortium that ranges from "pure" biology to "pure" ICT. The five partners have complementary background and domain of expertise. Moreover, all partners have past experience of inter-disciplinary work and individually gather competencies that include, though at different levels, both biology and computer science. Every partner has its own specific competencies and experience that is necessary to achieve the objectives and work of the EvoEvo project but with and overlap of knowledge that will enable an effective collaboration (Figure 10). INRIA, as coordinator has the overall overview.

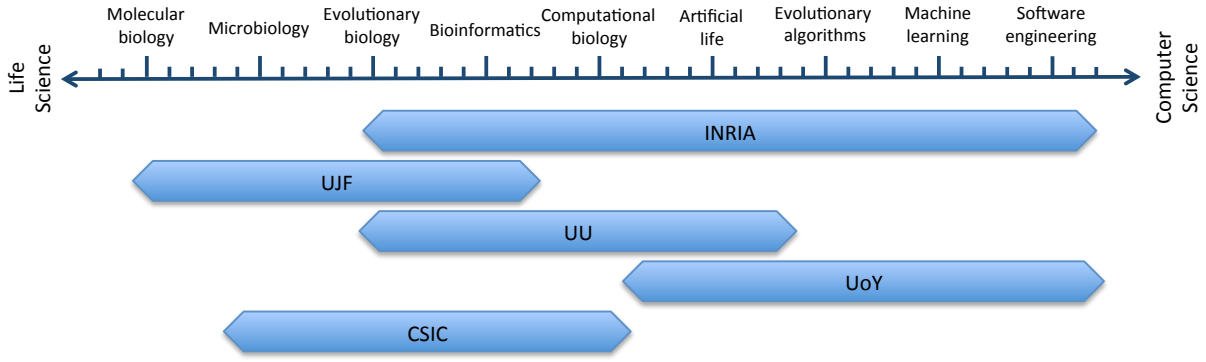


Figure 10 - Overview of the domains of expertise of the partners

[3] According to Spanish legislation, it is not possible at this stage to list names in the proposal since the position have to be publicly offered via CSIC job offers web portal and all potential candidates must be evaluated in equal conditions.

UJF working system consists in the longest-running evolution experiment with bacteria. During that experiment, an ancestor of *Escherichia coli* is propagated in a defined environment since more than 55,000 generations, which represents more than 2 million years at the human level. Their input in the project will consist in the understanding of the relationships between on one hand genome structure and expression, and on the other hand evolutionary processes in Bacteria.

CSIC is working on the mechanisms governing the evolution of RNA viruses and retroviruses. Among other topics, there are two that are more directly related to this proposal: the evolution of RNA genomes architecture and the adaptation of virus to new hosts (emerging viruses) and the evolution of specialist and generalist viruses (evolvability). Then newly acquired knowledge will serve as input for UU and INRIA

UU's interest is in modelling informatic processes in biotic systems, currently focussing on multilevel models of evolution. Their input will be developing particular models and study features like robustness, phenotypic variation and evolvability, the evolution of evolvability, and the open-endedness of evolution, interpret the results in biological terms so as to contribute to biological theory, explore the evolution of regulation of mutational process and simulate these processes, and analysing the results in multiple ways

INRIA is developing computational models of genome and transcription networks architecture evolution and focus on variability, robustness and evolvability determinants of these architectures as well as their consequences on the evolutionary process. Being a computer science institute, INRIA also conducts research in "pure" ICT: data-mining, bio-inspired artificial intelligence, data-management and databases. Thus, this partner has a global vision of all the project aspects, from biological questions to modelling to the final applications.

These novel approaches to "open" systems, based on experimental and theoretical work with biological processes, will be suitable for translating into a computational system, where it can be theoretically studies, and applied to "open" system applications (there the system needs to adapt online to ever-changing and unpredictable environment and inputs). UoY will deliver computational frameworks and applications, demonstrating the computational use of the discovered biological approaches, and feeding back into biological aspects that have the most computational power.

B2.3.1 Subcontracting

INRIA, UJF and CSIC have planned subcontracting, that represents 4% of the total project budget.

WP1 tasks need biological analyses involving specific competences and machinery not available in the consortium. In particular genome sequencing and global transcription profiling are necessary to complete all the research objectives of WP1. UJF and CSIC have planned subcontracting costs for these analyses:

- UJF subcontracting costs are for:
    - Genome sequencing: the experiments with the SOLiD technology will be performed on a French Ibiza platform dedicated to such projects (although all

reagents and materials will be bought directly by Partner 2). A budget of 18,000 euros will be required for UJF.
    - Global transcription profiling: the experiments with Affymetrix arrays will be performed on the ProfileXpert platform (Lyon) dedicated to such projects (although all reagents and arrays will be bought directly by Partner 2). A budget of 7000 euros will be required for UJF.
- CSIC subcontracting costs are for determining the genomic sequence of either individual clones or mixed virus populations at different stages of the evolution experiments and as a control for the success of introducing mutations in the genome. Sanger sequencing will be used for confirmation of mutagenesis events, for the rightness of cloning and to discard the incorporation of unwanted mutations during cloning. The amount and fate of genetic diversity in TEV populations during the different evolution experiments will be determined by Illumina HiSeq 2500 technology. To evaluate the extent of phenotypic innovation at the regulatory level, the RNA expression profiles of plants infected with different genotypes of TEV will be evaluated by RNA-seq using the same NGS technology. Given the large amount of sequencing involved in the project (both classic and NGS), based in our long experience, we have estimated sequencing costs of 67 650 €. CSIC will open a public concourse for selecting the company to be subcontracted to perform sequencing. The company chosen will be the one offering the best quality/prize ratio.

INRIA will subcontract the installation of the smart room(s) to a French company: HIKOB. The installation of the smart room(s) is necessary for testing the software developed in WPs 4 & 5 but it is not a direct product of the researches conducted in the project. That is why the consortium has chosen to subcontract this part rather than including a specific task (note that none of the partners have the necessary skill to equip smart rooms with sensor networks).

In the EvoEvo project, HiKoB proposition is to equip a one or many rooms with nearly ~100 HiKoB WOLF sensor nodes with the goal to provide redundant measures associated to :

- Temperature / humidity,
- Light detection,
- Sound detection
- The opening / closing of doors using accelerometers and inertial measurement units
- Proximity detection using ultrasound sensors
- Constraints and presence detections using strain gauges that can be mounted on chairs, tables, …

Depending on the advancement of the HiKoB R&D developments and of the EvoEvo project, it will be discussed the opportunity to equip a single room with high-level sensors (e.g., covering the whole floor of the room using a sensitive ground as already developed and deployed by HiKoB in a similar research project) or many rooms with simpler ones.

Approximative budget: 41K€

- 1 HiKoB GATEWAY: ~1200€
- 100 HiKoB WOLF: ~300€ per unit (to be defined more precisely depending on the exact measure requirements)

- HiKoB LivePulse (supervision software): ~1450€
- Software integration fee (API definition & implementation): ~1850
- Deployment support fee (2 days): ~1100€.
- Yearly support fee (9 days): ~4950€
-

HiKoB is a leading provider of wireless and scalable instrumentation systems that generate strategic data and information on physical resources and assets to feed your information system.

HiKoB has been awarded twice by the Ministry of Higher Education and Research as part of the national competition for the creation of innovative technology companies in 2011 and in 2012. HiKoB has been finalist of the IBM SmartCamp 2012 in Paris, of the Cleantech Republic Awards 2012 and awarded at the Trophée BREF Rhône-Alpes 2012.

With a +15 years of cumulated experience in the field of wireless sensors networks, HiKoB brings a strong track records of more than 2000 wireless sensor nodes deployed at an operational level.

Finally, within EvoEvo we will have subcontracting related to Certificates on Financial Statements (9 815 Euros):

Some partners have assigned subcontracting budget in the Management category to subcontract the certification of their financial statements. The budget has been calculated according to the FP7 rule about the need of a certificate only when there has been an accumulated funding claim of 375 000 Euros.

- Costs of these certificate Certificates on Financial Statements depending on the partners:
  - 5 000 Euros for Universiteit Utrecht (partner 3),
  - 2 815 Euros for University of York (partner 4)
  - 2 000 Euros for CSIC (partner 5; CSIC will subcontract with AudiHispania Grant Thornton SL. This company will be responsible for auditing the project).
- No certificates costs needed for Inria (partner 1) and UJF (partner 2) as they will be produced by their Competent Public Officers (Agent Comptable).

B2.3.2 Third parties

In EvoEvo Project, Inria represents the BEAGLE Team. BEAGLE Team is a Joint Research Unit for which Inria will represent the following Partner Institutions, Third Parties linked to Inria via the **special clause 10** in the Grant Agreement and Consortium Agreement:

- Université Claude Bernard Lyon 1 (UCBL)
- Institut National des Sciences Appliquées de Lyon (INSA)

All administrative documents related to the general agreements between INRIA and INSA-Lyon, between INRIA and UCBL as well as the documents attesting the creation and composition of the BEAGLE team are attached to this document (appendix B2).

Inria has an analytical accounting system allowing it to declare its actual costs (both direct and indirect). It fills in Form C with its own costs only**: 254 474 Euros as direct costs and 174 301 Euros as indirect costs.**

UBCL and INSA as Third Parties linked to Inria, carry out part of the work attributed by the Grant Agreement to Inria. However, as they are unable to identify with certainty their actual indirect costs, they use the flat rate of 60% for indirect costs. They fill in Form C with their own costs only:

- UCBL may charge costs related to the expenses of Carole KNIBBE: **37 254 Euros as direct costs and 22 352 Euros as a flat rate**. Carole KNIBBE will contribute to WP2 and 3.
- INSA may charge costs related to the expenses of Guillaume BESLON and Christophe RIGOTTI: **202 236 Euros as direct costs and 83 023 Euros as a flat rate.** Note that Mr. Yann GRIPAY, who is enroled in the project, belongs INSA-Lyon but not to the Beagle Team. Therefore, also the contribution of Mr. GRIPAY will be mentioned in the project reports, INRIA will not ask for funding for his contribution. Guillaume BESLON will contribute to WP2, 3, 4, 5 and 6. Christophe RIGOTTI and Yann GRIPAY will contribute to WP4 and 5.

This will imply during the project that:

- Inria remains the beneficiary of the present project and will thus fill in the financial forms needed for the management report.
- In Inria's Form C, UCBL and INSA Lyon will be indicated as Third Parties of Inria as well as the costs handled by it.
- Inria, UCBL and INSA Lyon will prepare and submit Certificate on Financial Statement as requested by the EC Financial guidelines

B2.4 Resources to be committed

The overall budget of the Evoevo project for the period of **36 months** is **3 432 930 Euro** (see table WT8 for details), which is appropriate for the scope of the planned activities. The requested EC contribution is **2 629 000 Euro**. The budget dedicated to research is distributed as follows.

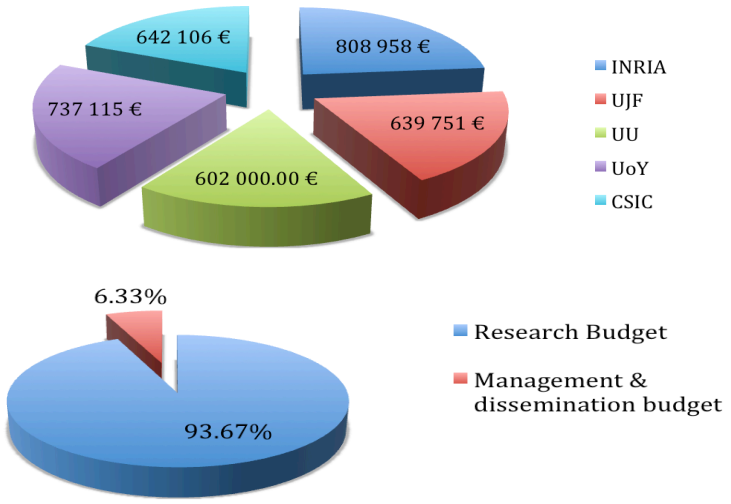


Figure 11 - Research Budget Distribution

B2.4.1 Human resources

The total effort dedicated to the project is equal to **316 PMs** (table WT6). The detailed budgets necessary to carry out the tasks foreseen within the EvoEvo project are indicated in the A3 forms and the summary table in the subsection below. Here we will tackle and explain the main budget items.

The major part (**93,7%**) of the PMs committed to the project will be spent on research activities. The remaining **6,3%** will be allocated to the management activities led by the project coordinator, INRIA, as well as all dissemination activities for all partners. Those will include technical, administrative and financial management as well as the dissemination activities such as participation to conferences, FET events, and the preparation and organisation of the workshop. In the figure below one can observe the similar percentages when budgets are compared.

The **personnel and overhead costs represent 76,8% of the total direct costs**. Other direct costs consist mainly of travel and subsistence costs since the consortium partners will attend the workshop planned in the activities of WP6 and extended semester meetings in order to benefit from interdisciplinary exchange. The complexity of the project, the particularity of its topic and the trans-disciplinary approach will require participation of all partners in common work sessions and meetings. What's more, the travel and subsistence expenses related to the participation of external experts will be taken care of by INRIA.

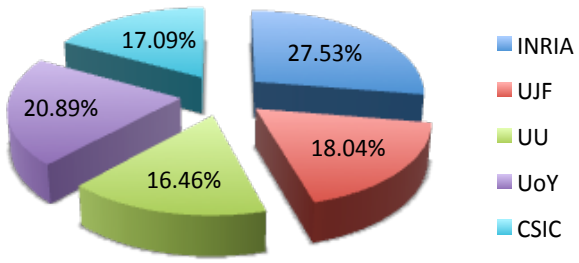


Figure 12 - Distribution of Resources.

The above figure clearly shows that the input of manpower is well-balanced within the consortium. The main increase is related to the coordination and management activities led by INRIA and their transversal involvement in the project. The requested resources are necessary to obtain the critical mass of skilled researchers able to achieve the deliverables and milestones of EvoEvo. Thereby, the consortium can achieve the objectives of the project. As it can be seen here below WP1 requires more involvement than the other WP because it is the only one dealing with living material and experiments have longer delays.

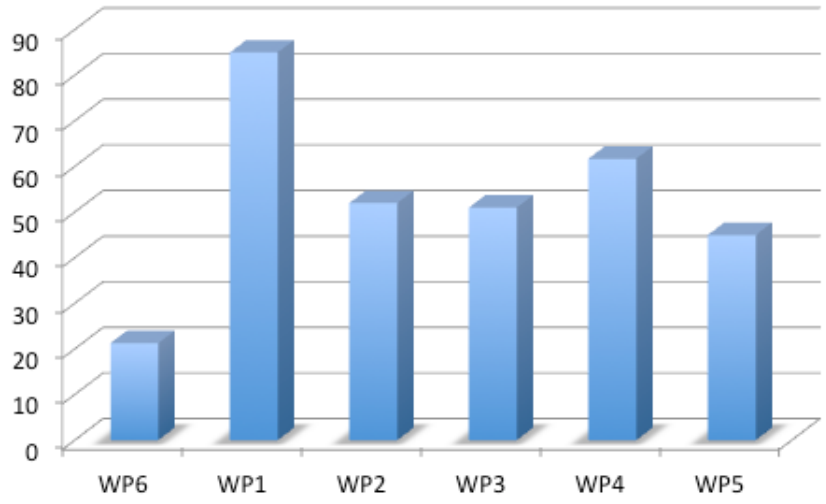


Figure 13 - Person.months by WP

The required resources have been estimated analytically and reflect the activities that each partner will carry out within the project. The distribution of workload between partners is well-adjusted and in line with the specific expertise they will bring to the consortium at different stages of the project. Therefore, we can clearly see that the partners use respectively the most PMs resources in the Work Packages they are leaders of.

B2.4.2 Direct costs

B2.4.2.1 Consumables

Consumables budget represents 9.4% of the total budget. Their main usage is for WP1 and the life science/biology research. Dealing with living material is more expensive but essential to the results needed for the work of the other WP.

Consumables for partners UJF (173,500 €) can be detailed as follows:

- Microbiology media (routine cultures, competition experiments, allelic exchanges, phenotypic tests) (30 000 euros)
- Biolog plates (12 000 euros)
- Molecular biology kits, PCR experiments, cloning experiments, electrophoresis (43 500 euros)
- Routine sequencing of DNA fragments and synthesis of oligonucleotides (20 000 euros)
- Reagents for genome sequencing (48 000 euros). A total of 188 genomes will be sequenced. We will perform these experiments directly and therefore ask for all the reagents needed.
- Reagents for global transcription profiling (20 000 euros). A total of 54 transcription profiles will be performed. We will perform these experiments directly and therefore ask for all the reagents needed and the chips.

A budget of 92,000 € was requested by CSIC to cover the cost of experimental work with TEV infecting different plant species. This amount is based in our extensive previous experience working with this experimental pathosystem:

- The proposed work requires a large amount of chemicals (e.g., salts, buffers, organic solvents, sugars, hormones, dNTPs, agarose…), molecular biology kits (e.g., site-directed mutagenesis, in vitro transcription, DNA and RNA purification and preparation, restriction enzymes, polymerases, synthesis of primers for PCR and sequencing, RT-PCR and PCR reactives, fluorochromes for qPCR and DNA labeling…), growth media (e.g., agar, LB, MS), plastic and glassware (e.g., 96-wells PCR plaques, a variety of test tubes of diverse volumes, petri dishes of different sizes, boxes, sample bags, flasks, beakers…), and greenhouse supplies (e.g., plastic pots, artificial substrates, nutritive media, light tubes, plastic for localized irrigation, labels, trays, phytosanitary treatments against insects and fungi…).
- In addition, space usage in IBMCP greenhouses is subject to fees (40 €/table/month × 8 tables = 320 €/month). Small pieces of equipment such as automatic pipettes, liquid nitrogen containers, freezer racks and boxes, and electrophoresis trays will be acquired as needed. Finally, office materials (e.g., paper, pens, permanent markers, staples, printers ink, envelopes, notebooks, CDs and DVDs, …) and express courier expenses are also included in this budget.

UU and UoY have planned consumables (respectively 36,000 € and 7,038 €) for maintenance and upgrade of computer hardware.

INRIA planned 20,000 € for consumables in order to complete the smart room and feed the EvoEvo applications with redundant low-level data streams from the environment, other sensors will be added, using different technologies with complementary features. For example, two technologies seem pertinent: EnOcean devices are wireless and energy autonomous sensors for basic ambient measurement (temperature, humidity, luminosity), with low rates and rough precision; Kinect-like devices, that are initially gaming input devices, provide movement detection, gesture recognition, and object tracking capabilities, as well as

low-level depth measurement. Those sensors (and associated hardware and software gateways) will contribute to the sensitivity of the smart room required for EvoEvo applications in tasks T5.1. For task T5.2, INRIA also plans to buy actuators, like EnOcean smart power outlets, and innovative user-interaction devices, like the Karotz “rabbit”. Those actuators and robot-like devices are required to make EvoEvo applications interact with the smart room (e.g., controlling room lights) and actually “live” with people in the smart room (e.g., displaying messages or even talking to people, moving around them).

B2.4.2.2 travels and subsistence

All partners have planned travels to attend EvoEvo meetings, Conferences and Ad hoc meetings project meetings needs for peer to peer collaborations.

- INRIA planned a 28,000 € budget for travels to attend project meetings (14x1,000 €), conferences (4x1,500 €) and ad-hoc meetings for joint work with other partners (for software specification and development – 8x1,000€).
- UJF planned a 22,500 € budget for travels to attend project meetings (14x1,000 €), conferences (3x1,500 €) and ad-hoc meetings for joint work with other partners (for model specification and tests – 4x1,000€).
- UU planned a 20,000 € budget for travels to attend project meetings (14x1,000 €), conferences (2x1,500 €) and ad-hoc meetings for joint work with other partners (for model specification and tests – 3x1,000€).
- UoY planned a 23,692 € budget for travels to attend project meetings (14x1,000 €), conferences (4x1,500 €) and ad-hoc meetings for joint work with other partners (for software specification and development – 4x923 €).
- CSIC planned a direct cost of 11,250 € to cover travel and per diem expenses. This will include all necessary travels to the different project meetings (as described in WP6) as well as to participate at least in one international conference (e.g., ESEB 2015) to present the results derived form the project.

B2.4.2.3 Management and dissemination

INRIA planned a 15,000 € budget for management and dissemination activities (organization of project meetings in Lyon, organization of the final dissemination workshop, publication fees…). UJF planned a 15,000 € budget for management and dissemination activities (organization of project meetings in Grenoble, publication fees – as the leader of WP1, UJF will have to support publication fees in open access life science journals). UU planned a 5000 € budget for organization of project meetings in Utrecht and publication fees. UoY planned a 8,444 € budget for publication and conferences fees. CSIC planned, a managing direct cost of 3 000 € to pay for the expenses associated to the organization of a meeting of all teams in Valencia (e.g., catering and rental of a conference room for two days).

B2.4.3 subcontracting

Subcontracting costs are 143 465 € (4% of the total budget) and correspond to provisions for INRIA, CSIC and UJF. INRIA will subcontract the installation of a smart room for the experiments details in WP5, to a French company called HIKOB. CSIC and UJF must perform genome sequencing (DNA and RNA) and global transcription profiling (subcontracting costs are detailed in section B2.3.1 above in this document).

B2.4.4 financial tables

The following table synthetize the global budget for each partner (see main text for details):

Eligible cost		**INRIA**	**UJF**	**UU**	**UoY**	**CSIC**
Personnel costs (€)		389,964.00	173,220.00	289,000.00	419,764.00	250,648.00
Subcontracting (€)		41,000.00	25,000.00	5,000.00	2,815.00	69,650.00
Other direct cost (€)	Travels and subsistence	28,000.00	22,500.00	20,000.00	23,692.00	11,250.00
	Consumables	20,000.00	173,500.00	36,000.00	7,038.00	92,000.00
	Management and dissemination activities	15,000.00	15,000.00	5,000.00	8,444.00	3,000.00
Indirect costs (€)		317,994.00	230,531.00	247,000.00	275,362.00	215,557.00
Total costs (€)		811,958.00	639,751.00	602,000.00	737,115,00	642,105.00
Requested EU contribution (€)		638,263.00	488,244.00	458,000.00	559,438.00	485,055.00

B3. IMPACT

B3.1 Strategic impact

B3.1.1 Impact on computer science: a realistic route towards living technologies

There are many analogies between computer and life sciences. For instance, the ideas of digital ecosystem, intrusion-detection immune system or brain-machine analogy try to integrate in computer systems some characteristics of living systems. In most cases however, the only characteristics that are looked out are functional (*what* the system is doing) but not structural (*how* it does it). Even in computational systems that are really inspired from biology (*e.g.* evolutionary computation or artificial neural networks), the structural characteristics - that give them their functional properties - of the living system are overlooked. As a direct consequence, at the functional level our computational artefacts are far from the properties of real living systems, even the "simplest" of them: microorganisms. This is actually not surprising since the rules that link the structural properties to the functional ones in living beings are still mostly unknown and the few things we know are still highly uncertain. To say it simply, we just don't know how biological systems create and process information. For example, it is far from intuitive to predict how environmental challenges impact on the structure of regulatory or metabolic networks and why a new organization is more appropriate in given conditions.

EvoEvo is based on two ideas: first, the initial step towards living technologies requires a strong collaboration between biologists and computer scientists to initiate a virtuous cycle between discovering rules of biological information processing and exploiting them in artificial systems. Second, current artificial evolutionary systems are oversimplistic when compared to biological evolving systems because they neglect the major rule that evolution evolves. This evolution of evolution has led to contemporary evolving systems that evolve much more efficiently than our computational artifacts and, probably, than the RNA systems at the origin of life. Bacteria and viruses in particular have pushed the power of evolution at its paroxysm since they rely on it even for immediate adaptation to environmental fluctuations. For example, they can cope with various stresses (temperature, pH, osmotic, oxidative) by modifying their global gene expression profiles in a time scale of few minutes. Following this idea, one could argue that artificial evolution would benefit from copying more closely these biological systems in order to evolve as efficiently. In recent years, different position papers have claimed that introducing more biology in artificial evolution would help [Banzhaf et al., 2006; O'Neill *et al.,* 2010]. In EvoEvo, we propose not only to integrate biology in artificial evolution but in a differently important way: first, by identifying the mechanisms resulting in Evolution of Evolution and second, by implementing them in artificial systems to create living technologies. Thus, **EvoEvo will produce evolving systems able to accelerate the course of their own evolution**.

- **At short-term these evolving systems will be able to adapt dynamically to their environment**. This will be tested through two application benchmarks.
- **At mid-term, such systems will be able to grow in complexity and to naturally become living technologies**. We argue that life complexity cannot be directly copied to build living technologies because: (1) it is not understood fully by biologists, and (2)

it results from evolution that, unlike engineers, creates highly intricate systems. In contrast, we argue that Evolution of Evolution might be copied once understood since it is the actual process, and not its result. Thus EvoEvo is the most realistic - if not the only - way towards living technologies. The goal here will be to understand *how* a specific system structure is able to perform *what* functions it is doing.

- **At long-term, these living technologies will cause a profound shift in the direction of future ICT design, engineering and applications**. Living technologies will not anymore require to be programmed since they will spontaneously find the optimal behaviour. These systems will be reactive to their environments, fully autonomous, self-healing and innovative to find behaviours that will step out of the human imaginative power that will undoubtedly become the main limitation of traditional programmable technologies in a very near future.

B3.1.2 Impact on life science

Life science aims at investigating the composition and functioning of biological systems, as well as their interactions together and with their environment. Life science relies on five major actions: observe, quantify, understand, manipulate, and predict. The cellular, biochemical and molecular components of biological systems have been described at multi-scale levels: from molecules, cells, organs, and organisms, to populations and ecosystems. All these traits are combined into the single most important, integrative and complex phenotype of all, which is the Darwinian fitness of an organism in a particular environment. By Darwinian fitness, biologists mean the ability of a given organism to leave viable offspring in a given environment. In addition, computational models of large-scale physiological systems have been proposed in the recent years. The merge of computational and mathematical modelling techniques with the development of experimental platforms for massive and high-throughput molecular data collection gave rise to a new discipline called Systems Biology, whose aim is to develop detailed quantitative models for the functioning of cells, tissues, organs, and individuals and to predict how these systems will respond to perturbations.

As stated by evolutionary biologist Theodosius Dobzhansky in 1973: “nothing in Biology makes sense except in the light of Evolution”. Darwin’s evolutionary theory (and its more recent developments) is the only hard theory in all Life science. It provides the unifying mechanism and conceptual framework to understand how organisms evolve to optimize their fitness, resulting in the beautiful and extraordinary diversity of phenotypes of contemporary living organisms. Random mutations on the genome, together with flexible regulatory networks provide the raw material for this process.

The interactions between biological systems and their environment have major impacts on virtually all societal issues (public policy, ethics, public health, communication, transport, science, education, behaviour of complex systems, translation to therapeutic and engineering applications, mutual impact of human populations and environment…). Indeed, our planet is a heterogeneous and constantly changing environment that has major impacts on the distribution and behaviour of all living organisms from microbes to mammals. In turn, their complex adaptive responses to environmental challenges impose modifications of the environment. Understanding the complex interactions between species and with their environment is a prerequisite to improve life conditions in the context of our human societies, communities and public policies. Favourable outcomes such as ecosystems management,

control of infectious diseases, development of new cancer and neurodegenerative therapies, or increased production of food will depend from our ability to predict and manipulate these evolutionary systems.

EvoEvo provides an integrative and synergistic framework of life science, evolutionary biology, population genetics, computer science, mathematics and engineering to understand the dynamics of these interactions, by investigating the heart of life sciences: evolution of the evolutionary processes themselves. It will address the properties of molecular and cellular components and how they can interact and can be integrated in larger networks and systems for the dynamic behaviour of biological systems, resulting in adaptation of living cells. Finally, EvoEvo will extend life sciences to the challenging objective of manipulating and producing new biological and hybrid systems. This challenge is the first step of EvoEvo to be integrated in artificial evolution with the aim of building living technologies.

B3.1.3 Social impact

Impact on public health

Because of its strong interdisciplinary nature, our project has implications in the fields of computer science and microbiology but also in public health.

The impressive adaptive abilities of microorganisms have broad harmful impacts in our human society, even just considering the race engaged with our attempts to fight the various diseases they cause. Examples are numerous, from the emergence of new diseases to the development of antibiotic-resistant bacteria or the difficulties to fight nosocomial diseases. Hence, it is well known that the development of human technologies like air-conditioning systems resulted in the conditions favouring the emergence of diseases as legionellosis. Bacterial abilities to form biofilms resulted in difficulties during surgeries and increased care to avoid contaminations leading to nosocomial diseases. Efficient gene exchanges by horizontal transfer, combined with our human transport systems, resulted in the fast dissemination of multi-drug resistant bacterial pathogens. Considering only the two model organisms that will be scrutinized in the EvoEvo project (*E. coli* and TEV), one can already identify a huge impact on health and food, not to talk about the generalization of the results to other bacterial and virus species:

- Usually commensal, *E. coli* can also emerge as a severe pathogen that causes many human and animal diseases. The *E. coli* infections are of increasing concern worldwide, with ~2 million human deaths per year. Unarguably, *E. coli* is one of the most important human pathogens in Europe and one of the principal causes of morbidity and mortality from community- and hospital-acquired extraintestinal infections, responsible for more than 80% of community-acquired urinary tract infections [Denamur *et al.*, 2002], newborn-meningitis, and sepsis. Recently, *E. coli* strains have also been associated with cancers and chronic inflammatory disorders. As a human disease, *E. coli* infections have a high public health importance as well as substantial economic burdens. As an animal and zoonotic disease, *E. coli* also plays a crucial role in human food safety, animal welfare and economy of production. Furthermore, the *E. coli* species is problematic in term of resistance with the recent emergence of strains producing extended-spectrum beta-lactamases and

carbapenemases. The 2011 German *E. coli* O104:H4 outbreak of haemorrhagic diarrhoea is an example of the cocktail of high virulence and resistance that can emerge in this species. The outbreak rapidly expanded in the entire European continent causing more than 52 deaths and a total cost of more than one billion euros (including compensation packages paid to farmers by EU and affected countries).

- Plant viruses are one of the main causes of plant diseases in the world, resulting in billions of dollars lost per year by reduced plant production, quality and quantity [Thresh, 2006; van der Vlugt, 2006]. More specifically, potyviruses represent a large group of plant viruses that are able to infect different crop species. *Tobacco etch virus* in particular has a wide host range and can affect various plants like pepper, tomato or tobacco. It is an economically important plant pathogen, causing millions of dollars of crop loss worldwide.

Pathogenic microorganisms often emerge or re-emerge owing to dynamic and fast evolutionary changes of mutational processes and regulatory networks that allow them to colonize new host species, transmit between individual hosts by new means, resist therapeutic antibiotics, and so forth. The evolutionary rate is a prominent factor of the ability of microorganisms to adapt to new environmental conditions and possibly to develop virulence traits resulting in new emergent diseases. Understanding variability processes, evolution and dissemination of mutant isolates (either more pathogenic or more resistant against drugs) is necessary to fight epidemics. For example, a high proportion of mutants with increased mutation rates (so-called hypermutators with 10- to 100-fold increases in mutation rates due to deficient DNA repair systems) have repeatedly been identified in clinical isolates of bacterial pathogens. Hypermutation results in increased progression of diseases and drug resistance. Moreover, genome dynamics has been shown to underlie changes in bacterial lifestyle, including for pathogens [Gomez-Valero *et al.*, 2007].

Pathogen evolution is often supposed to be governed solely by a direct selective pressure (fitness increase). Yet it is far from being the only evolutionary pressure since it permanently interacts with genetic drift, mutational biases and indirect selective pressures for variability, robustness, evolvability or host-selectivity. The precise causes, consequences and mechanisms of these indirect pressures in microorganisms, which are *terra incognita* for modern biology, will be investigated in our project, in real organisms but also through wider results obtained though *in silico* modelling.

The development of predictive models of microbial variability, robustness, evolvability and open-endedness will allow for a much more efficient battle against pathogens by integrating all levels of biological checkpoints, thereby also either reducing the resistance abilities or discovering new therapeutic targets. Analyses of evolution experiments in the laboratory, but also of clinical isolates of bacterial pathogens, revealed that changes in the structure and expression of genomes are the major contributors of bacterial evolutionary processes. For example, fluctuations of mutation rates have been shown to result from the fit to the environmental conditions but also from the effect of the mutations that are produced by elevated mutation rates [Wielgoss *et al.*, 2013]. Moreover, bacterial adaptive abilities rely on the reorganization of chromosomal structures through large rearrangements (inversions, deletions, duplications, amplifications) and on the rewiring of metabolic and regulatory networks [Hindré *et al.*, 2012]. This is also the case for experiments where resistance to increasing doses of antibiotics was selected for. While mutations in the expected genes

encoding the antibiotic targets were detected, changes also occurred in global regulator-encoding genes, thereby leading to modifications of regulatory networks. In all these experiments however, the relationship between these important structural and functional changes and evolutionary processes like robustness, evolvability and phenotypic innovation is mostly unknown, and must be investigated to fully understand how specific structures of genomes and networks are able to result in specific answers to given environmental challenges. It will give us new insights on drug design (*e.g*. by identifying global regulators that are evolutionarily conserved) and dose delivery (*e.g.* by taking into account its effect on evolutionary dynamics). Moreover, successful establishment as a full pathogen is often associated with the loss of some repair genes and a genome reduction [Moran, 2002]. In that respect, EvoEvo will contribute in the future to simulate the emergence of pathogens, a problem of major importance with the recent world changes that bring in closer contact pathogens from wild animals and humans. Additionally, taking into account the evolutionary mechanisms is likely to give new insights on emergent disease dynamics and on environmental conditions promoting their spread.

To conclude EvoEvo addresses crucial questions for public health: what are the mechanisms fine-tuning the balance between robustness and evolvability in bacteria and viruses? What are the mechanisms underlying the fluctuation of mutation rates? What are the evolutionary conditions favouring adaptation to new host species? EvoEvo is likely to have a direct impact on public health by enabling a deeper understanding of the mechanisms that enable microorganisms to rapidly escape human efforts to fight them and colonize new niches.

Impact on social usage of information technologies

Electronic devices and computers are nowadays almost everywhere, from desktop computers in offices to smart phones in pockets and sensors on the walls. With current networking capabilities, such devices are technically able to exchange all kind of data and information, in order to build a word-wide “pervasive” or “ubiquitous” environment that provides new services to people. This vision is however constrained by the capacity for application and service developers to handle the heterogeneity and dynamicity of such an environment. Current developments require a mix of imperative code for application logic, declarative code for data management, and network protocols for communications. Some recent works are targeted towards a fully declarative service-based approach to simplify application development [ActiveXML, SoCQ], with run-time dynamic discovery of the environment and dynamic interactions to benefit from all computing resources, event those that are intermittent (especially in mobile environments like smartphones).

However, these approaches still aim at fulfilling a pre-defined goal, *i.e.* their application logic is defined at design time. Introducing evolution of this application logic enables applications and services to better integrate changes in the computing environment and to adapt to new settings and user requirements automatically. Introducing evolution of evolution would further free applications and services from their design-time constraints, enabling them to discover and benefit from unknown computing resources, to adapt their behaviour to the ever-changing real world setting (*e.g.* learning different users requirements), and to propose new kinds of behaviours and interactions, that would be assessed and selected according to automatically defined and evolving rules.

Designing applications and services that are ready to discover the computing *terra incognita* of their current and future surroundings is an expected breakthrough for modern computer science. This innovative approach would reduce the complexity to develop intelligent software, and hence intelligent objects and services. In particular, it would make practicable a mass-customization of such intelligent entities that would be able to interact in a personalized way with users and to create pertinent innovative services, even within "niche" markets with particular settings.

B3.1.4 Expected impacts listed in the work programme

The Evlit workprogram lists two major expected impacts. EvoEvo directly targets these two expected impacts as detailed in the following.

Foundations, approaches and proofs of concept for a radically new type of living technology.

Evolution of evolution is likely to enable artificial evolving systems not only to increase their quality relative to an extrinsic fitness criterion but also to grow in complexity and to dramatically increase their evolutionary capacity in the context of their usage. EvoEvo will provide important knowledge about the mechanisms of evolution of evolution (*foundations*). By developing a computational framework able to use evolution of evolution, it will propose a way of exploiting these principles in real computation systems (*approaches*). Ultimately, by applying the framework to two real-life applications, EvoEvo will prove the usability of this approach (*proof of concept*). Moreover, the EvoEvo approach, by enabling evolution acceleration, will enable to create a shift from artificial evolution to living technologies, exactly as biological evolution experienced successive transitions from the initial self-replicating molecules to today's complex organisms [Szathmáry & Maynard-Smith, 1994; Szathmáry & Maynard-Smith, 1997].

Possible contributions beyond the area of ICT (manufacturing, chemistry, biology, agriculture).

EvoEvo is an interdisciplinary project that includes computer scientists and biologists. The foundations of evolution of evolution that will be scrutinized in the projects are likely to have an important impact on biology, health and agriculture. Indeed, evolution is at the heart of many contemporary problems such as new disease emergence, antibiotic resistance, invasive species or biodiversity. In all these situations, understanding the determinants of variability, robustness, evolvability or niche construction of the species at stake (*e.g.* viruses, bacteria, parasites…) will enable a better management and control of these phenomena. Indeed, many attempts to fight microbial pathogens do not directly integrate their abilities to evolve, thereby resulting in escape to new treatments. This level is actually the heart of EvoEvo that will integrate all dimensions of bacterial behaviour since it is aimed at understanding the head of the pyramid, *i.e.* the mechanisms allowing evolution to evolve. EvoEvo will tackle these questions with a large range of tools, from experimental biology on viruses and bacteria to modelling, thus producing theoretical knowledge on evolution that could be directly translated towards public health or resource management.

B3.1.5 Measurement of the EvoEvo impact

EvoEvo is a challenging concept. Although we aim, in the project, at a better understanding and at proposing applications of the concept (from life science to ICT applications), we also expect a strong impact through global dissemination of the concept and tools developed in the project. Moreover, we expect them to disseminate in the life science community as well as in the ICT community.

In order to measure this impact, we propose to use two indicators:

- Bibliometric indicators. At the beginning of the project[4], the "EvoEvo" request gives no hit in Web of Science (WoS), DBLP or PubMed and only three correct hits in google scholar (other hits being noise due to translation errors and/or constant names in mathematical equations). We will use these bibliometric indicators to measure the spreading of the concept in the life science communities (via WoS or PubMed) and ICT communities (DBLP). In particular, we will track whether other researchers than the ones directly involved in the project use the concept.

- Use of the software. We will propose the models developed in WP2 to the computational biology community and maintain a list of external users of the software developed in the project.

As claimed in section B1.1.3, it is impossible to predict what will be living technologies and how they will interact with human beings. In the EvoEvo project, we will implement our models to "let them live" in smart rooms and smart agents. We will use this implementation to test the receptivity of such technologies. To this aim, we will interact with research teams specialized in social aspects of ICT in order to propose a small questionnaire to the room user. This survey will enable us to measure the social and personal impacts of such technologies.

B3.2 Plan for the use and dissemination of foreground

To have a disruptive impact on science, technology and society, EvoEvo will have to efficiently disseminate its results and know-how. A detailed dissemination plan will be proposed at mid-term of the project (deliverable D6.5). Preliminary EvoEvo's dissemination plan includes:

- **Academic Dissemination:** To achieve the best possible scientific dissemination of the project results, it is important to publish them in the highest ranked journals, conferences and workshops. To this aim, the interdisciplinary nature of the project must be taken into account since communication media vary in the different scientific communities. In particular, we expect life-science results to be published in the highest impact journals. The models developed in WP2 will be distributed freely to the academic community under the open source General Public Licence (GPL –

[4] Accessed May 19th 2013.

http://www.gnu.org/licenses/gpl-3.0.en.html). In particular, we will propose them to other EVLIT projects that expressed a need for evolutionary models (*eg.*, EVOBLISS, EVOPROG or PlasWires). Computer science results will be sent to high-impact software-engineering and machine learning conferences as well as more specialized ones within fields such as artificial evolution, self-adaptation, complex systems, artificial life, autonomic and intelligent computing and unconventional computing. To this aim, EvoEvo will join the FET coordination action TRUCE (http://www.truce-project.eu). The idea is to introduce the major project results to the software engineering and machine learning community at large, while also stimulating further advances in research in related areas.

- **Online Dissemination:** A website (www.evoevo.eu - the domain name is the property of the project leader for the duration of the project) will provide a first access point for scientific and non-scientific parties interested into the EvoEvo project. Key results will be published on that website as well as developed software (all software developed in the project will be available freely under the General Public Licence). In particular, all computational models will be made freely available through the website. The long-term objective of the website is to create a community of interested parties around computational evolution, digital evolution and EvoEvo, to accelerate their involvement and to create awareness of the research results.
- **Popular science:** Evolution is an important concept that has many consequences on everybody's life. However, these consequences are often misunderstood. EvoEvo's website will contain dedicated pages for non-initiated public in order to popularize the concept of evolution as well as the project results. In particular, the *in silico* models will be made available as serious games in order to teach evolution and to offer training tools to familiarize the public (including physician) to questions like nosocomial disease emergence, antibiotic resistance emergence and good practices in antibiotic usage, biodiversity or evolutionary consequences of global warming.
- **Interdisciplinary dissemination**: Owing to its large interdisciplinarity, the EvoEvo's consortium is able to trigger interdisciplinarity between computer scientists and life scientists, at least in Europe. In particular, the fields of experimental evolution, digital evolution and artificial evolution share many concepts and methods, but they only rarely communicate. All partners of the consortium will share their contacts in these fields to organize an interdisciplinary workshop on the topic of experimental practices of evolution at the end of the project (M36). We expect this workshop to enable expertise sharing between the different fields concerned by experimental evolution.

As for IPR, the consortium will sign at the start of the project the consortium agreement, based on the DESKA model. Nevertheless, platforms and models will be open source and information will flow easily between all partners. Since the developments will be open source, there will be little to protect at the end of the project.

In the project, INRIA will produce scientific results on evolutionary theory as well as applications of evolution of evolution to demonstrate the potential of the approach. Therefore, INRIA will report its results in scientific conferences (in biological evolution as well as in artificial evolution) and in renowned journals (with a preference for open access journals).

The two applications will be a source of publications and will serve as dissemination media in an electronic format (*e.g.* videos of the hardware personal companion).

For UJF and CSIC, the objective is to understand the constraints that apply on bacterial cells and viruses. These results will be published in renowned journals. Moreover, evolutionary data can be used to feed evolutionary models, with the goal of being able to predict the adaptive abilities of bacteria and viruses. In any case where it will be possible, joint publications with computational models will be submitted. Moreover, since evolutionary theory is at the heart of all biological processes, the results of EvoEvo will be inserted into the courses of the Master program at UJF, with the aim of favouring contacts between students registered in Masters from different fields.

UU, as most partners, will report the results in scientific conferences (ranging from conferences primarily on biological evolution, to meetings on artificial evolution and complexity theory – and most importantly they will publish the results in renowned (and mainly open access) journals. As well in ongoing teaching programs.

UoY intend to develop the Computational Platform that will be Open Source code with related documentation. In addition, a User Guide / documentation – detailing how the proof-of-concept applications were developed using the platform, as exemplars for others to follow.

Finally, figure 14 presents a tentative logotype for the EvoEvo project.

Figure 14 - Logo of the EvoEvo project (tentative)

B4. ETHICAL ISSUES (IF APPLICABLE)

	YES	PAGE	COMMENT
Informed Consent	NO		
Does the proposal involve children?	NO		
Does the proposal involve patients or persons not able to give consent?	NO		
Does the proposal involve adult healthy volunteers?	NO		
Does the proposal involve Human Genetic Material?	NO		
Does the proposal involve Human biological samples?	NO		
Does the proposal involve Human data collection?	NO		
Research on Human embryo/foetus	NO		
Does the proposal involve Human Embryos?	NO		
Does the proposal involve Human Foetal Tissue / Cells?	NO		
Does the proposal involve Human Embryonic Stem Cells?	NO		
Privacy	NO		See comments in section B4.1
Does the proposal involve processing of genetic information or personal data (e.g. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)	NO		
Does the proposal involve tracking the location or observation of people?	NO		
Research on Animals	NO		See comments in section B4.2
Does the proposal involve research on animals?	NO		
Are those animals transgenic small laboratory animals?	NO		
Are those animals transgenic farm animals?	NO		
Are those animals cloning farm animals?	NO		
Are those animals non-human primates?	NO		
Research Involving Developing Countries	NO		
Use of local resources (genetic, animal, plant etc)	NO		
Benefit to local community (capacity building ie access to healthcare, education etc)	NO		
Dual Use	NO		See comments in section B4.3
Research having potential military / terrorist application	NO		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	YES		

B4.1 Use of data generated by sensor networks

The proof-of-concept applications of the EvoEvo project will be the classification of data streams generated by smart houses and the use of the classes for training personal companion. All the sensors used in the smart-room that will be designed for these experiments are low-level. They will generate environmental data not directly related to the specific individuals present in the room (e.g. pressure, temperature, sound…). In particular, we will not use video. Moreover, the data streams will be used on-line and not stored to produce public databases or benchmark. Therefore, there are no ethical concerns on the use of these data in the applications presented here.

B4.2 Use of biological material

All EvoEvo biological experiments will be performed on microorganisms (E. coli and TEV). We are therefore not concerned by ethical questions related to animal experimentation. Nevertheless, we emphasize that all experiments will be performed using the most rigorous safety conditions, including the warranty that no microbial organism (genetically-modified or not) will be released outside the laboratory. The UJF and CSIC partners are fully equipped with the most high-standard tools to ensure isolation of cultures. Moreover, all bacterial strains used in EvoEvo derive from the laboratory strain B which is not pathogenic and does not normally colonize the human intestine. These traits did not change during 55,000 generations of evolution in the laboratory, making this bacterium a safe model in EvoEvo. The specific TEV strain used in the experiments has been attenuated to avoid its transmission by aphid vectors. All plant work will be confined in a BLS-2 biosafety level greenhouse.

B4.3 Dual use

It is always difficult to ensure that a research of any kind and field will never be used in the military domain. This is particularly true for biotechnological applications (consider e.g. the DARPA programs in synthetic biology) and for Information and Communication Technologies. EvoEvo will shed light on properties of the evolution of microorganisms that will enable a better understanding of their ability to quickly colonize new niches. In EvoEvo, we will use the knowledge on these abilities, not the abilities themselves. Moreover, we would like to emphasize that a better knowledge of the evolutionary properties of microorganisms will directly help to efficiently address ethical questions for example about the use of engineered strains in the environments (e.g. as depolluters). Indeed, it will give better insights on the eventual risks of such strains to escape from their initial objective and to start colonizing unexpected niches and/or presenting unexpected properties.B5. Gender aspects (optional)

B5. GENDER ISSUES (IF APPLICABLE)

Two of the five partners are lead by women – Susan Stepney and Paulien Hogeweg; four of the nine investigators involved in the project are women. The gender balance of the project is thus almost in equilibrium and we are all aware of gender aspects in science and technology.

APPENDIX B1: REFERENCES

[Adamatzky, 2010] A. Adamatzky, 2010, *Physarum Machines: Computers From Slime Mould*. World Scientific, 280 p.

[Adamatzky *et al.,* 2005] A. Adamatzky, B. de Lacy Costello, T. Asai, 2005, *Reaction-Diffusion Computers*. Elsevier, 334 p.

[Adami, 2006] C. Adami, 2006, Digital genetics: unravelling the genetic basis of evolution, *Nat. Rev. Genet.,* 7(2):109-118.

[Aggarwal *et al.,* 2006] C. C. Aggarwal, J. Han, J. Wang, P. S. Yu, 2006, A Framework for On-Demand Classification of Evolving Data Streams, *IEEE Transactions on Knowledge and Data Engineering*, 18(5):577-589.

[Agudelo-Romero *et al.,* 2008] P. Agudelo-Romero, P. Carbonell, M. A. Pérez-Amador, S. F. Elena, 2008, Virus adaptation by manipulation of host's gene expression, *PLoS ONE* 3:e2397.

[Ahmed *et al.,* 2010] T. Ahmed, T. S. Shimizu, R. Stocker, 2010, Microfluidics for bacterial chemotaxis. *Integr Biol (Camb)* 2(11-12):604-629.

[Aldana *et al.,* 2007] M. Aldana, E. Balleza, S. Kauffman, O. Resendiz, 2007, Robustness and evolvability in genetic regulatory networks, *Journal of Theoretical Biology*, 245(3):433-448.

[Alon, 2007] U. Alon, 2007, Network motifs: theory and experimental approaches, *Nat Rev Genet.* 8(6):450-461.

[Anantharaman *et al.,* 2005] T. S. Anantharaman, V. Mysore, B. Mishra, 2005, Fast and cheap genome wide haplotype construction via optical mapping. *Pac. Symp. Biocomput.* 2005:385-396.

[Ancel & Fontana, 2000] L. W. Ancel, W. Fontana, 2000, Plasticity, evolvability, and modularity in RNA. *J. Exp. Zool.* 288 (3):242-283.

[Andrews *et al.,* 2011] P. S. Andrews, S. Stepney, T. Hoverd, F. A. C. Polack, A. T. Sampson, J. Timmis, 2011, CoSMoS process, models, and metamodels. *CoSMoS workshop, Paris, France, August 2011*, Luniver Press, pp.1-13.

[Bantinaki *et al.,* 2007] E. Bantinaki, R. Kassen, C. G. Knight, Z. Robinson, A. J. Spiers, P. B. Rainey, 2007, Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity, *Genetics* 176(1):441-453.

[Banzhaf *et al.,* 2006] W. Banzhaf, G. Beslon, S. Christensen, J. A. Foster, F. Képès, V. Lefort, J. F. Miller, M. Radman, J. J. Ramsden, 2006 From artificial evolution to computational evolution: a research agenda, *Nature Genetic Review*, 7(9):729-735.

[Barrick *et al.,* 2009] J. E. Barrick,D. S. Yu, S. H. Yoon, H. Jeong,T. K. Oh, D. Schneider, R. E. Lenski, J. F. Kim, 2009, Genome evolution and adaptation in a long-term experiment with *Escherichia coli*, *Nature*, 461(7268):1243-1247.

[Batada & Hurst, 2007] N. N. Batada, L. D. Hurst, 2007, Evolution of chromosome organization driven by selection for reduced gene expression noise. *Nat. Genet.* 39(8):945-949.

[Beaumont *et al.,* 2009] H. J. Beaumont, J. Gallie, C. Kost, G. C. Ferguson, P. B. Rainey, 2009, Experimental evolution of bet hedging, *Nature*, 462(7269):90-93.

[Bedau & Packard, 2003] M. A. Bedau, M. H. Packard, 2003, Evolution of evolvability via adaptation of mutation rates, *BioSystems* 69(2-3):143-162.

[Beck, 2000] K. Beck, 2000, *Extreme Programming Explained*, Addison Wesley, 224 p.

[Beslon *et al.,* 2010a] G. Beslon, D. P. Parsons, Y. Sanchez-Dehesa, J.-M. Peña, C. Knibbe, 2010,. Scaling laws in bacterial genomes: A side-effect of selection of mutational robustness? *BioSystems,* 102:32-40.

[Beslon *et al.,* 2010b] G. Beslon, D. P. Parsons, J.-M. Peña, C. Rigotti, Y. Sanchez-Dehesa, 2010, From digital genetics to knowledge discovery: Perspectives in genetic network understanding. *Intelligent Data Analysis journal (IDAj)* 14:173-191.

[Bloom *et al.,* 2006] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, 2006, Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. USA,* 103(15):5869-5874.

[Blount *et al.,* 2012] Z. D. Blount, J. E. Barrick, C. J. Davidson, R. E. Lenski, 2012, Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 489(7417):513-518.

[Ciliberti *et al.,* 2007a] S. Ciliberti, O. C. Martin, A. Wagner, 2007, Robustness can evolve gradually in complex regulatory gene networks with varying Topology. *PLoS Comput. Biol.* 3:e15.

[Ciliberti *et al.,* 2007b]] S. Ciliberti, O. C. Martin, A. Wagner, 2007, Innovation and robustness in complex regulatory gene networks, *Proc. Natl. Acad. Sci. USA*, 104(34):13591-13596.

[Clune *et al.,* 2011] J. Clune, K. O. Stanley, R. T. Pennock, C. Ofria, 2011, On the Performance of Indirect Encoding Across the Continuum of Regularity, IEEE Transactions on Evolutionary Computation, 15(3):346-367.

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

[Copeland, 2008] B. J. Copeland, 2008, The Modern History of Computing. In: *The Stanford Encyclopedia of Philosophy*, 2008 ed.; Zalta, E.N., Ed.; Archived online: http://plato.stanford.edu/archives/fall2008/entries/computing-history/ (accessed on 17 December 2012)

[Correia & Fonseca, 2007] M. B. Correia, C. M. Fonseca, 2007, How redundancy and neutrality may affect evolution on NK fitness landscapes, *CEC2007*, pp. 2842-2849.

[Correia, 2010] L. Correia, 2010, Computational evolution: taking liberties, *Theory Biosci*., 129(2-3):183-**1** 91**.**

[Crombach & Hogeweg, 2007] A. Crombach, P. Hogeweg, 2007, Chromosome rearrangements and the evolution of genome structuring and adaptability, *Mol. Biol. Evol.,* 24(5):1130-1139.

[Crombach & Hogeweg, 2008] A. Crombach, P. Hogeweg, 2008, Evolution of evolvability in gene regulatory networks, *PLoS Comput. Biol.,* 4(7): e1000112.

[Crombach & Hogeweg, 2009] A. Crombach, P. Hogeweg, 2009, Evolution of resource cycling in ecosystems and individuals, *BMC Evol. Biol.,* 9:122.

[Cuypers & Hogeweg, 2012] T. D. Cuypers, P. Hogeweg, 2012, Virtual genomes in flux: an interplay of neutrality and adaptability explains genome expansion and streamlining, *Genome Biol. Evol.,* 4(3):212-229.

[Denamur *et al.,* 2002] E. Denamur, S. Bonacorsi, A. Giraud, P. Duriez, F. Hilali, C. Amorin, E. Bingen, A. Andremont, B. Picard, F. Taddei, I. Matic, 2002, High frequency of mutator strains among human uropathogenic *Escherichia coli* isolates, *J. Bacteriol.* 184(2):605-609.

[De Visser *et al.,* 2003] J. A. De Visser, J. Hermisson, G. P. Wagner, M. L. Ancel, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao, J. M. Cheverud, S. F. Elena, W. Fontana, G. Gibson, T. K. Hansen, D. Krakauer, R. C. Lewontin, C. Ofria, S. H. Rice, G. von Dassow, A. Wagner, M. C. Whitlock, 2003, Evolution and detection of genetic robustness. *Evolution* 57(9):1959-1972.

[Dittrich *et al.,* 2001] P. Dittrich, J. Ziegler, W. Banzhaf, 2001, Artificial Chemistries - A Review, *Artificial Life,* 7(3):225-275.

[Doursat, 2008] R. Doursat, 2008, Organically grown architectures: Creating decentralized, autonomous systems by embryomorphic engineering. In: Organic Computing, R. P. Würtz, Springer-Verlag, pp. 167-200.

[Draghi & Wagner, 2009] J. Draghi, G. P. Wagner, 2009, The evolutionary dynamics of evolvability in a gene network model, *J. Evol. Biol.,* 22(3):599-611.

[Du, 2010] K. L. Du, 2010, Clustering: A neural network approach, *Neural Networks*, 23(1):89-107.

[Dunham *et al.,* 2002] M. J. Dunham, H. Badrane, T. Ferea, J. Adams, P. O. Brown, F. Rosenzweig, D. Botstein, 2002, Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae, *Proc. Natl. Acad. Sci. USA.,* 99(25):16144-16149.

[Edwards & Palsson, 2000] J. S. Edwards, B. O. Palsson, 2000, Robustness analysis of the *Escherichia coli* metabolic network. *Biotechnol. Prog*. 16(6): 927-939.

[Eldredge, 1989] N. Eldredge, 1989, *Macroevolutionary Dynamics : Species, Niches and Adaptive Peaks*, McGraw_Hill.

[Elena, 2012] S. F. Elena, 2012, RNA virus genetic robustness: possible causes and some consequences, *Curr. Opin. Virol*. 2(5):525-530.

[Elena *et al.,* 2006] S. F. Elena, P. Carrasco, J. A. Daròs, R. Sanjuán, 2006, Mechanisms of genetic robustness in RNA viruses. *EMBO Rep* 7(2):168-173.

[Elena & Rodrigo, 2012] S. F. Elena, G. Rodrigo, 2012, Towards an integrated molecular model of plant virus interaction. *Curr. Opin. Virol*. 2(6):713-718.

[Elena & Sanjuán, 2008] S. F. Elena, R. Sanjuán, 2008, The effect of genetic robustness on evolvability in digital organisms. *BMC Evol.Biol*. 8:284.

[Elowitz *et al.,* 2002] M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, 2002, Stochastic gene expression in a single cell. *Science* 297(5584):1183-1186.

[Espinosa-Soto *et al., 2011*] C. Espinosa-Soto, O. C. Martin, A. Wagner, 2011, Phenotypic plasticity can facilitate adaptive evolution in gene regulatory circuits, *BMC Evol. Biol.,* 11:5. doi: 10.1186/1471-2148-11-5.

[Evans, 2004] E. Evans, 2004, *Domain-Driven Design: tackling complexity in the heart of software*, Addison Wesley, 560 p.

[Ferea *et al.,* 1999] T. L. Ferea, D. Botstein, P. O. Brown, R. F. Rosenzweig, 1999, Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc. Natl. Acad. Sci. USA.,* 96(17):9721-9726.

[Ferrada & Wagner, 2012] E. Ferrada, A. Wagner, 2012, A comparison of genotype-phenotype maps for RNA and proteins, *Biophys J.,* 102(8):1916-1925.

[Fiegna *et al.,* 2006] F. Fiegna, Y-t. N. Yu, S. V. Kadam, G. J. Velicer, 2006, Evolution of an obligate social cheater to a superior cooperator. *Nature,* 441(7091):310-314.

[Fisher, 1930] R. A. Fisher, 1930, *The Genetical Theory of Natural Selection*, Oxford: Clarendon Press.

[Floreano & Mattiussi, 2008] D. Floreano, C. Mattiussi, 2008, *Bio-Inspired Artificial Intelligence Theories, Methods, and Technologies,* MIT Press, Boston, 659 p.

[Fontana & Schuster, 1987] W. Fontana, P. Schuster, 1987, A computer model of evolutionary optimization. *Biophys Chem.,* 26(2-3):123-147.

[Frank, 2007] S. A. Frank, 2007, Maladaptation and the paradox of robustness in evolution, *PLoS ONE* 2:e1021.

[Franke *et al.,* 2011] J. Franke, A. Klözer, J. A. G. M. de Visser, J. Krug, 2011, Evolutionary accessibility of mutational pathways, *PLoS Comput. Biol*., 7:e1002134.

[Freeman & Pryce, 2010] S. Freeman, N. Pryce, 2010, *Growing Object-Oriented Software, Guided by Tests*, Addison Wesley, 284 p.

[Funes & Pollack, 1999] P. Funes, J. Pollack, 1999, Computer Evolution of Buildable Objects. *In Evolutionary Design by Computers. P. Bentley (editor)*. Morgan Kaufmann, San Francisco. pp. 387-403.

[Gandrillon *et al.,* 2012] O. Gandrillon, D. Kolesnik-Antoine, J. J. Kupiec, G. Beslon, 2012, Chance at the heart of the cell, *Prog Biophys Mol Biol.,* 110(1):1-4.

[Gilbert & Dupont, 2011] J. A. Gilbert, C. L. Dupont, 2011, Microbial metagenomics: beyond the genome. *Ann Rev Mar Sci* 3:347-371.

[Gillespie, 1977] D. T. Gillespie, 1977, Exact Stochastic Simulation of Coupled Chemical Reactions. *The Journal of Physical Chemistry,* 81: 2340-2361.

[Gomez-Valero *et al.,* 2007]. L. Gómez-Valero, E. P. C. Rocha, A. Latorre, F. J. Silva, 2007, Reconstructing the ancestor of *Mycobacterium leprae* : The dynamics of gene loss and genome reduction, *Genome Research*, 17(8):1178-1185.

[Gosset *et al.,* 2004] G. Gosset, Z. Zhang, S. Nayyar, W. A. Cuevas, M. H. Saier Jr., 2004, Transcriptome analysis of ,-dependent catabolite control of gene expression in *Escherichia coli*, *J. Bacteriol.* 186(11):3516-3524.

[Griffin *et al.,* 2004] A. S. Griffin, S. A. West, A. Buckling, 2004, Cooperation and competition in pathogenic bacteria. *Nature* 430(7003):1024-1027.

[Güell *et al.,* 2011] M. Güell, E. Yus, M. Lluch-Senar, L. Serrano, 2011, Bacterial transcriptomics: what is beyond the RNA horiz-ome? *Nat. Rev. Microbiol*. 9(9):658-669.

[Haldane, 1930-32] J. B. S. Haldane, 1930-32, A Mathematical Theory of Natural and Artificial Selection, Part I-IX. *Mathematical Proceedings of the Cambridge Philosophical Society* 23: 838–844.

[Han *et al.,* 2011] M. J. Han, J. W. Lee, S. Y. Lee, 2011, Understanding and engineering of microbial cells based on proteomics and its conjunction with other omics studies. *Proteomics* 11(4):721-743.

[Hickinbotham *et al.,* 2010] S. Hickinbotham, E. Clark, S. Stepney, T. Clarke, A. Nellis, M. Pay, P. Young, 2010, Diversity from a monoculture: effects of mutation-on-copy in a string-based artificial chemistry. *ALife XII,* MIT Press, pp. 24-31.

[Hickinbotham *et al.,* 2011] S. Hickinbotham, S. Stepney, A. Nellis, T. Clarke, E. Clark, M. Pay, P. Young, 2011, Embodied genomes and metaprogramming, *ECAL 2011,* MIT Press, pp. 334-341.

[Hindré *et al.,* 2012] T. Hindré, C. Knibbe, G. Beslon, D. Schneider, 2012, New insights into bacterial adaptation through in vivo and in silico experimental evolution, *Nature Reviews Microbiology,* 10:352-365.

[Hogarth, 1992] M. Hogarth, 1992, Does General Relativity allow an observer to view an eternity in a finite time? *Foundations of Physics Letters*, 5(TODO):73–81.

[Hornby, 2004] G. S. Hornby, 2004, Functional scalability through generative representations: the evolution of table designs, *Environment and Planning B*, 31(4):569-587.

[Hornby & Pollack, 2001] G. S. Hornby, J. B. Pollack, 2001, The Advantages of Generative Grammatical Encodings for Physical Design, *CEC 2001*, IEEE Press, pp. 600-607.

[Huynen & Hogeweg, 1994] M. A. Huynen, P. Hogeweg, 1994, Pattern generation in molecular evolution: exploitation of the variation in RNA landscapes, *J. Mol. Evol.,* 39(1):71-79.

[Huynen *et al.,* 1996] M. A. Huynen, P. F. Stadler, W. Fontana, 1996, Smoothness within ruggedness: the role of neutrality in adaptation, *Proc. Natl. Acad. Sci. USA.,* 93(1):397-401.

[Huynen, 1996] M. A. Huynen, 1996, Exploring phenotype space through neutral evolution, *J. Mol. Evol.,* 43(3):165-169.

[Igel *et al.,* 2006] C. Igel, T. Suttorp, N. Hansen, 2006, A Computational Efficient Covariance Matrix Update and a (1+1)-CMA for Evolution Strategies, *Proc. of GECCO'2006, ACM Press*, pp. 453-460.

[Igel *et al.,* 2007] C. Igel, N. Hansen, S. Roth, 2007, Covariance Matrix Adaptation for Multi-objective Optimization, *Evolutionary Computation*, 15(1):1-28.

[Jacob, 1977] F. Jacob, 1977, Evolution and tinkering. *Science* 196:1161-1166.

[Kaneko, 2011] K. Kaneko, 2011, Proportionality between variances in gene expression induced by noise and mutation: consequence of evolutionary robustness, *BMC Evol Biol.,* 11:27, doi: 10.1186/1471-2148-11-27.

[Kassen & Rainey, 2004] R. Kassen, P. B. Rainey, 2004, The ecology and genetics of microbial diversity. *Annu. Rev. Microbiol.* 58:207-231.

[Kauffman & Levin, 1987] S. Kauffman, S. Levin, 1987, Towards a general theory of adaptive walks on rugged landscapes, *Journal of Theoretical Biology,* 128(1):11-45.

[Kim, 2007] Y. Kim, 2007, Rate of adaptive peak shifts with partial genetic robustness, *Evolution* 61(8):1847-1856.

[Kimura, 1977] M. Kimura, 1977, The Neutral Theory of Molecular Evolution and Polymorphism, Scientia, 112: 687-707.

[Kirschner & Gerhart, 1998] M. Kirschner, J. Gerhart,1998, Evolvability, *Proc. Natl. Acad. Sci. USA* 95 (15): 8420-8427.

[Klauck *et al.,* 2007] E. Klauck, A. Typas, R. Hengge, 2007, The sigmaS subunit of RNA polymerase as a signal integrator and network master regulator in the general stress response in *Escherichia coli*, *Sci. Prog.* 90(pt 2-3):103-127.

[Knibbe *et al.,* 2007a] C. Knibbe, O. Mazet, F. Chaudier, J.-M. Fayard, G. Beslon, 2007, Evolutionary coupling between the deleteriousness of gene mutations and the amount of non-coding sequences, *Journal of Theoretical Biology,* 244(4):621-630.

[Knibbe *et al.,* 2007b] C. Knibbe, A. Coulon, O. Mazet, J.-M. Fayard, G. Beslon, 2007, A long-term evolutionary pressure on the amount of noncoding DNA, *Molecular Biology and Evolution,* 24(10):2344-2353.

[Knibbe *et al.,* 2008] C. Knibbe, J.-M. Fayard, G. Beslon, 2008, The topology of the protein network influences the dynamics of gene order : From systems biology to a systemic understanding of evolution, *Artificial Life*, 14(1):149-156.

[Krakauer & Plotkin, 2002] D. C. Krakauer, J. B. Plotkin, 2002, Redundancy, antiredundancy, and the robustness of genomes. *Proc. Natl. Acad. Sci. USA,* 99(3):1405-1409.

[Kuo *et al.,* 2006] P. D. Kuo, W. Banzhaf, A. Leier, 2006, Network topology and the evolution of dynamics in an artificial genetic regulatory network model created by whole genome duplication and divergence, *Biosystems*, 85(3):177-200.

[Kussell & Leibler, 2005] E. Kussell, S. Leibler, 2005, Phenotypic diversity, population growth, and information in fluctuating environments, *Science*, 309(5743):2075-2078.

[Laland *et al.,* 1999] K. N. Laland, F. J. Odling-Smee, M. W. Feldman, 1999, Evolutionary consequences of niche construction and their implications for ecology. *Proc. Natl. Acad. Sci. USA,* 96(18):10242-10247.

[Lalic *et al.,* 2011] J. Lalic, J. M. Cuevas, S. F. Elena, 2011, Effect of host species on the distribution of mutational fitness effects for an RNA virus. *PLoS Genet*. 7:e1002378.

[Lalic & Elena, 2012] J. Lalic, S. F. Elena, 2012, Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. *Heredity* 109(2):71-77.

[Le Gac *et al.,* 2008] M. Le Gac, M. D. Brazas, M. Bertrand, J. G. Tyerman, C. C. Spencer, R. E. Hancock, M. Doebeli, 2008, Metabolic changes associated with adaptive diversification in *Escherichia coli*. *Genetics* 178(2):1049-1060.

[Le Gac *et al.,* 2012] M. Le Gac, J. Plucain, T. Hindré, R. E. Lenski, D. Schneider, 2012, Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *E. coli*. *Proc. Natl. Acad. Sci. USA,* 109(24):9487-9492.

[Lehner, 2008] B. Lehner, 2008, Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol. Syst. Biol.* 4:170.

[Leier *et al.,* 2007] A. Leier, P. D. Kuo, W. Banzhaf, 2007, Analysis of preferential network motif generation in an artificial regulatory network model created by duplication and divergence, *Advances in Complex Systems*, 10(2):155-172.

[Lenski, 2001] R. E. Lenski, 2001, Twice as natural, *Nature*, 414(6861):255

[Lenski, 2004] R. E. Lenski, 2004, Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*, *Plant Breed. Rev.* 24:225-265.

[Lenski *et al.,* 2006] R. E. Lenski, J. E. Barrick, C. Ofria, 2006, Balancing robustness and evolvability. *PLoS Biol*. 4:e428.

[Lerman *et al.,* 2012] J. A. Lerman, D. R. Hyduke, H. Latif, V. A. Portnoy, N. E. Lewis, J. D. Orth, A. C. Schrimpe-Rutledge, R. D. Smith, J. N. Adkins, K. Zengler, B. O. Palsson, 2012, In silico method for modelling metabolism and gene product expression at genome scale. *Nat Commun*. 3:929.

[Lipson & Pollack, 2000] H. Lipson, J. B. Pollack, 2000, Automatic design and Manufacture of Robotic Lifeforms, *Nature*, 406(6799):974-978.

[Locke & Elowitz, 2009] J. C. Locke, M. B. Elowitz, 2009, Using movies to analyse gene circuit dynamics in single cells. *Nat Rev Microbiol* 7(5):383-392.

[Maes, 1987] P. Maes, 1987, Concepts and experiments in computational reflection. *OOPSLA'87*, ACM Press, pp. 147-155.

[Mattiussi & Floreano, 2007] C. Mattiussi, D. Floreano, 2007, Analog genetic encoding for the evolution of circuits and networks. *IEEE Trans. Evol. Comput.* 11(5):596-607.

[Marbach *et al.,* 2009] D. Marbach, C. Mattiussi, D. Floreano, 2009, Replaying the evolutionary tape: biomimetic reverse engineering of gene networks. *Ann. N. Y. Acad. Sci.* 1158:234-245.

[Maynard-Smith, 1992] J. Maynard-Smith, 1992, Byte-sized evolution, *Nature*, 355(6363):772-773.

[Misevic *et al.,* 2012] D. Misevic, A. Frénoy, D. P. Parsons, F. Taddei, 2012, Effects of public good properties on the evolution of cooperation. In: *Proc. of Artificial Life 13*, pp. 218-225

[Moran, 2002] N. A. Moran, 2002, Microbial minimalism: genome reduction in bacterial pathogens, *Cell*, 108(5):583-586.

[Mozhayskiy & Tagkopoulos, 2012] V. Mozhayskiy, I. Tagkopoulos, 2012, Microbial evolution in vivo and in silico: methods and applications, *Integrative Biology*, [Epub ahead of print]

[Mustonen & Lässig, 2009] V. Mustonen, M. Lässig, 2009, From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation, *Trends Genet.,* 25(3):111-119.

[Nellis & Stepney, 2001] A. Nellis, S. Stepney, 2011, Embodied copying for richer evolution. *ECAL 2011*, MIT Press, pp. 597-604.

[Neidhardt, 1996], F. C. Neidhardt, 1996, *Escherichia coli and Salmonella: Cellular and molecular biology*, Second edition. ASM Press, Washington DC.

[O'Neill, 2003] B. O'Neill, 2003, Digital evolution, *PLoS Biology*, 1(1):11-14.

[O'Neill *et al.,* 2010] M. O'Neill, L. Vanneschi, S. Gustafson, W. Banzhaf, 2010, Open issues in genetic programming, *Genetic Programming and Evolvable Machines*, 11(3):339-363.

[Orr, 2009], A. Orr, 2009, Fitness and its role in evolutionary genetics, *Nat. Rev. Genet.,* 10(8):531-539.

[Pal *et al.,* 2006] C. Pál, B. Papp, M. J. Lercher, P. Csermely, S. G. Oliver, L. D. Hurst, 2006, Chance and necessity in the evolution of minimal metabolic networks, *Nature*, 440(7084):667-670.

[Papp *et al.,* 2004] B. Papp, C. Pál, L. D. Hurst, 2004, Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast, *Nature*, 429(6992):661-664.

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

[Pigliucci, 2008] M. Pigliucci, 2008, Is evolvability evolvable? *Nat. Rev. Genet.* 9(1)**:** 75-82.

[Proulx & Phillips, 2005] S. R. Proulx, P. C. Phillips, 2005, The opportunity for canalization and the evolution of genetic networks. *Am. Nat*. 165(2):147-162.

[Proulx *et al.,* 2007] S. R. Proulx, S. Nuzhdin, D. E. Promislow, 2007, Direct selection on genetic robustness revealed in the yeast transcriptome. *PLoS ONE* 2:e911.

[Queitsch *et al.,* 2002] C. Queitsch, T. Sangster, S. Lindquist, 2002, Hsp90 as a capacitor of phenotypic variation. *Nature* 417(6889):618-624.

[Rainey & Travisano, 1998] P. B. Rainey, M. Travisano, 1998, Adaptive radiation in a heterogeneous environment, *Nature* 394(6688):69-72.

[Ray, 1992] T. S. Ray, 1992, An Approach to the Synthesis of Life. *Artificial Life II*, Addison-Wesley, pp. 371-408.

[Reisinger & Miikkulainen, 2006] J. Reisinger, R. Miikkulainen, 2006, Selecting for evolvable representations, In : Proc. of the 2006 Genetic and evolutionary computation conference, ACM Press, pp. 257-264.

[Rennell *et al.,* 1991] D. Rennell, S. Bouvier, L. Hardy, A. Poteete, 1991, Systematic mutation of bacteriophage T4 lysozyme. *J. Mol. Biol.* 222(1):67-87.

[Rodrigo *et al.,* 2012] G. Rodrigo, J. Carrera, V. Ruiz-Ferrer, F. J. del Toro, C. Llave, O. Voinnet, S. F. Elena, 2012, A meta-analysis reveals the commonalities and differences in *Arabidopsis thaliana* response to different viral pathogens. *PLoS ONE* 7:e40526.

[Rosenzweig *et al.,* 1994] R. F. Rosenzweig, R. R. Sharp, D. S. Treves, J. Adams, 1994, Microbial evolution in a simple unstructured environment – genetic differentiation in *Escherichia coli*, *Genetics* 137(4):903-917.

[Rutherford & Lindquist, 1998] S. L. Rutherford, S. Lindquist, 1998, Hsp90 as a capacitator for morphological evolution. *Nature* 396(6709): 336-342.

[Sanjuán & Elena, 2006] R. Sanjuán, S. F. Elena, 2006, Epistasis correlates to genomic complexity, *Proc. Natl. Acad. Sci. USA* 103(39):14402-14405.

[Sanjuán *et al.,* 2004] R. Sanjuán, A. Moya, S. F. Elena, 2004, The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl. Acad. Sci. USA* 101(22):8396-8401.

[Schuster *et al.,* 1994] P. Schuster, W. Fontana, P. F. Stadler, I. L. Hofacker, 1994, From sequences to shapes and back: a case study in RNA secondary structures, *Proc Biol Sci.,* 255(1344):279-284.

[Sims, 1994] K. Sims, 1994, Evolving 3D Morphology and Behavior by Competition, *Proc. of Artificial Life IV*, MIT Press, pp. 28-39

[Sinha & Nussinov, 2001] N. Sinha, R. Nussinov, 2001, Point mutations and sequence variability in proteins: redistributions of preexisting populations. *Proc. Natl. Acad. Sci. USA* 98(6):3139-3144.

[Smarda *et al.,* 2008] P. Smarda, P. Bures, L. Horová, B. Foggi, G. Rossi, 2008, Genome size and GC content evolution of Festuca: ancestral expansion and subsequent reduction. *Ann. Bot.,* 101(3):421-433.

[Sniegowski *et al.,* 1997] P. D. Sniegowski, P. J. Gerrish, R. E. Lenski, 1997, Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387(6634):703-705.

[Snijder & Pelkmans, 2011] B. Snijder, L. Pelkmans, 2011, Origins of regulated cell-to-cell variability. *Nat Rev Mol Cell Biol* 12(2):119-125.

[Souza-Lima et al., 2011] C. A. Souza Lima Jr., C. M. F. Lapa, C. M. do N. A. Pereira, J. J. da Cunha, A. C. M. Alvim, 2011, Comparison of computational performance of GA and PSO optimization techniques when designing similar systems – Typical PWR core case, *Annals of Nuclear Energy*, 38(6):1339-1346.

[Spencer *et al.,* 2007] C. C. Spencer, M. Bertrand, M. Travisano, M. Doebeli, 2007, Adaptive diversification in genes that regulate resource use in *Escherichia coli*. *PLoS Genet.* 3:e15.

[Stanley & Miikkulainen, 2002] K. O. Stanley, R. Miikkulainen, 2002, Evolving neural networks through augmenting topologies, *Evolutionary Computation,* 10(2):99-127.

[Stearns, 2002] S. C. Stearns, 2002, Progress on canalization. *Proc. Natl. Acad. Sci. USA* 99(16):10229-10230.

[Stearns *et al.,* 1995] S. C. Stearns, M. Kaiser, T. J. Kawecki, 1995, The differential genetic and environmental canalization of fitness components in *Drosophila melanogaster*, *J. Evol. Biol.* 8:539-557.

[Stepney, 2008] S. Stepney, 2008, The Neglected Pillar of Material Computation, *Physica D,* 237(9):1157-1164.

[Stepney & Hoverd, 2011] S. Stepney, T. Hoverd, 2011, Reflecting on Open-Ended Evolution, *ECAL 2011,* MIT Press, pp. 781-788.

[Stepney *et al.,* 2013] S. Stepney, K. Alden, P. S. Andrews, J. L. Bown, A. Droop, T. Ghetiu, T. Hoverd, F. A. C. Polack, M. Read, C. G. Ritson, A. T. Sampson, J. Timmis, P. H. Welch,

A. F. T. Winfield, 2013, *Engineering Simulations as Scientific Instruments*, Springer, 2013 (in preparation)

[Stepney *et al.,* 2005] S. Stepney, R. E. Smith, J. Timmis, A. M. Tyrrell, M. J. Neal, A. N. W. Hone, 2005, Conceptual Frameworks for Artificial Immune Systems, *International Journal of Unconventional Computing*, 1(3):315-338.

[Szathmáry & Maynard-Smith, 1994] E. Szathmáry, J. Maynard-Smith, 1997, The major evolutionary transitions, *Nature*, 374(6519):227-232.

[Szathmáry & Maynard-Smith, 1997] E. Szathmáry, J. Maynard-Smith, 1997, From Replicators to Reproducers: the First Major Transitions Leading to Life, *Journal of Theoretical Biology*, 187(4):555-571.

[Taddei *et al.,* 1997] F. Taddei, M. Radman, J. Maynard-Smith, B. Toupance, P.-H. Gouyon, B. Godelle, 1997, Role of mutator alleles in adaptive evolution, Nature 387(6634):700-703.

[Tagkopoulos *et al.,* 2008] I. Tagkopoulos, Y. C. Liu, S. Tavazoie, 2008, Predictive Behavior Within Microbial Genetic Networks, *Science*, 320(5881):1313-1317.

[Takeuchi & Hogeweg, 2008] N. Takeuchi, P. Hogeweg, 2008, Evolution of complexity in RNA-like replicator systems, *Biol Direct*., 3:11. doi: 10.1186/1745-6150-3-11.

[Takeuchi *et al.,* 2011] N. Takeuchi, P. Hogeweg, E. V. Koonin, 2011, On the origin of DNA genomes: evolution of the division of labor between template and catalyst in model replicator systems, *PLoS Comput Biol.,* 7(3):e1002024.

[Tenaillon *et al.,* 2001] O. Tenaillon, F. Taddei, M. Radman, I. Matic, 2001, Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation, *Res. Microbiol* 152(1):11-16.

[Thresh, 2006] J. M. Thresh, 2006, Crop viruses and virus diseases: A global prospective, In:*Virus Diseases and Crop Biosecurity*, J. I. Cooper, T. Kuhne, V. P. Polischuk, Dordrecht (NL): Springer, pp. 9-32.

[Touchon *et al.,* 2009] M. Touchon, C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguénec, M. Lescat, S. Mangenot, V. Martinez-Jéhanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Médigue, E. P. Rocha, E. Denamur, 2009, Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths. *PLoS Genet* 5(1):e1000344.

[Treves *et al.,* 1998] D. S. Treves, S. Manning, J. Adams, 1998, Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Mol. Biol. Evol.* 15(7):789-797.

[van der Vlugt, 2006] R. A. A. van der Vlugt, 2006, Plant viruses in European agriculture: Current problems and future aspects, In:*Virus Diseases and Crop Biosecurity*, J. I. Cooper, T. Kuhne, V. P. Polischuk, Dordrecht (NL): Springer, pp. 33-44.

[van Hoek & Hogeweg, 2007] M. J. van Hoek, P. Hogeweg, 2007, Metabolic adaptation after whole genome duplication, *Mol Biol Evol.,* 24(11):2485-2494.

[van Nimwegen *et al.,* 1999] E. van Nimwegen, J. P. Crutchfield, M. Huynen, 1999, Neutral evolution of mutational robustness, *Proc. Natl. Acad. Sci. USA* 96(17):9716-9720.

[Varela *et al.,* 1991] F. Varela, E. Roch, E. Thompson, 1991, *The embodied mind: Cognitive science and human experience*, MIT Press.

[Veening *et al.,* 2008] J. W. Veening, W. K. Smits, O. P. Kuipers, 2008, Bistability, epigenetics, and bet-hedging in bacteria, *Annu Rev Microbiol*. 62:193-210.

[Velicer & Yu, 2003] G. J. Velicer, Y-t. N. Yu, Y-t., 2003, Evolution of novel cooperative swarming in the bacterium *Myxococcus xanthus*. *Nature,* 425(6953):75-78.

[Velicer & Vos, 2009] G. J. Velicer, M. Vos, 2009, Sociobiology of the myxobacteria, *Annu. Rev. Microbiol.* 63:599-623.

[von Dassow *et al.,* 2000] G. von Dassow, E. Meir, E. M. Munro, G. M. Odell, 2000, The segment polarity network is a robust developmental module, *Nature* 406(6792):188-192.

[Wagner *et al.,* 1997] G. P. Wagner, G. Booth, H. Bagheri-Chaichian, 1997, A population genetic theory of canalization. *Evolution* 51:329-347.

[Wagner, 2005a] A. Wagner, 2005, Robustness, evolvability, and neutrality, *FEBS Lett*. 579(8): 1772-1778.

[Wagner, 2005b] A. Wagner, 2005, *Robustness and evolvability in living systems*. Princeton University Press, 384 p.

[Wagner, 2008] A. Wagner, 2008, Robustness and evolvability a paradox resolved, *Proc. R. Soc. B,* 275(1630): 91-100.

[Wagner, 2012a] A. Wagner, 2012, The role of robustness in phenotypic adaptation and innovation, *Proc Biol Sci.* 279(1732):1249-58.

[Wagner, 2012b] A. Wagner, 2012, Metabolic networks and their evolution, *Adv Exp Med Biol.,* 751:29-52.

[Weimer *et al.,* 2009] W. Weimer, T. Nguyen, C. Le Goues, S. Forrest, 2009, Automatically finding patches using genetic programming. In: Proc. of the Int. Conf. on Soft. Eng. (ICSE), pp. 364-374.

[Widmer & Kubat, 1996] G. Widmer, M. Kubat, 1996, Learning in the presence of concept drift and hidden contexts, *Machine Learning*, 23(1):69-101.

[Wielgoss *et al.,* 2013] S. Wielgoss, J. E. Barrick, O. Tenaillon, M. J. Wiser, W. J. Dittmar, S. Cruveiller, B. Chane-Woon-Ming, C. Médigue, R. E. Lenski, D. Schneider, 2013, Mutation

rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc Natl Acad Sci USA.* 110(1):222-227.

[Wilke *et* al, 2001] C. O. Wilke, J. L. Wang, C. Ofria, R. E. Lenski, C. Adami, 2001, Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* 412(6844): 331-333.

[Wolf *et al.,* 2012] Y. I. Wolf, K. S. Makarova, N. Yutin, E. V. Koonin, 2012, Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer, *Biology Direct,* 4:46

[Wright, 1932] S. Wright, 1932, The roles of mutation, inbreeding, crossbreeding and selection in evolution, *Proc. 6th Int. Cong. Genet.,* 1: 356-366.

[Wright, 1984] S. Wright, 1984, *Evolution and the Genetics of Populations: Genetics and Biometric Foundations*, University of Chicago Press, Vol. 1-4.

[Zhang *et al.,* 2011] Q. Zhang, G. Lambert, D. Liao, H. Kim, K. Robin, C. K. Tung, N. Pourmand, R. H. Austin, 2011, Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments, *Science,* 333(6050):1764-1767.

[Zheng *et al.,* 2004] D. Zheng, C. Constantinidou, J. L. Hobman, S. D. Minchin, 2004, Identification of the , regulon using *in vitro* and *in vivo* transcriptional profiling, *Nucl. Acids Res*, 32(19):5874-5893.

[Zwart *et al.,* 2011] M. P. Zwart, J. A. Daròs, S. F. Elena, 2011, One is enough: in vivo effective population size is dose-dependent for a plant RNA virus, *PLoS Pathog*, 7:e1002122.