World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:3, No:4, 2009

Computational Identification of Bacterial Communities

Eleftheria Tzamali, Panayiota Poirazi, Ioannis G. Tollis and Martin Reczko

***Abstract*—**Stable bacterial polymorphism on a single limiting resource may appear if between the evolved strains metabolic interactions take place that allow the exchange of essential nutrients [8]. Towards an attempt to predict the possible outcome of long-running evolution experiments, a network based on the metabolic capabilities of homogeneous populations of every single gene knockout strain (nodes) of the bacterium *E. coli* is reconstructed. Potential metabolic interactions (edges) are allowed only between strains of different metabolic capabilities. Bacterial communities are determined by finding cliques in this network. Growth of the emerged hypothetical bacterial communities is simulated by extending the metabolic flux balance analysis model of Varma et al [2] to embody heterogeneous cell population growth in a mutual environment.

Results from aerobic growth on 10 different carbon sources are presented. The upper bounds of the diversity that can emerge from single-cloned populations of *E. coli* such as the number of strains that appears to metabolically differ from most strains (highly connected nodes), the maximum clique size as well as the number of all the possible communities are determined. Certain single gene deletions are identified to consistently participate in our hypothetical bacterial communities under most environmental conditions implying a pattern of growth-condition- invariant strains with similar metabolic effects. Moreover, evaluation of all the hypothetical bacterial communities under growth on *pyruvate* reveals heterogeneous populations that can exhibit superior growth performance when compared to the performance of the homogeneous wild-type population.

***Keywords*—**Bacterial polymorphism, clique identification, dynamic FBA, evolution, metabolic interactions.

I. Introduction

UNDERSTANDING bacterial diversity apart from its biological significance is of great applied importance in the area of bio-degradation of pollutants, in food preservation as well as in human health. Pathogens being able to generate extensive variability within populations continue to threaten human life.

In most environmental conditions, bacterial cells have evolved to maximize their growth performance constrained by the physiochemical laws that govern their intracellular processes as well as the dynamic environment in which they grow. On the other hand, mutations occur in nature. If they lead to metabolic improvements, these mutants will dominate the population, otherwise they will vanish. However, coexistence among different mutants may also evolve in the population. Long-term evolution experiments have shown the maintenance of more than a single strain in a simple environment, where the evolved clones significantly differed from one another with respect to their metabolic capabilities such as the growth rate and the uptake rates as well as the gene expression patterns [1]. The actual ecological reasons together with the evolutionary forces that favor more than one competitor in a single limiting resource still comprise an open issue [7]. However, it is argued that stable bacterial polymorphism on a single limiting resource is possible to appear if there are metabolic interactions that allow the exchange of essential nutrients between the diverse strains [8].

This study is based on the following reasoning. If cells of the exact same metabolic capabilities coexist in the same environment, the interactions between them as shaped by their environment will not result in any new experiences; whether they grow on their own or coexist with each other makes no difference. On the other hand, the interactions between cells of various metabolic capabilities (metabolic diversity) that grow in the same environment may give rise to new metabolic capabilities for each participant as they differently shape their environment. Metabolically different cells may satisfy each other's nutritional needs. Whether these interactions might eventually prove beneficial or harmful for the participants depends on their dynamical metabolic constraints.

For certain environmental conditions, a network of potential synergism is reconstructed based on the metabolic capabilities of each single gene knockout strain of the bacterium of *E. coli* where metabolic interactions are investigated among strains with different capabilities. A genome-scale flux balance model of the bacterium E. coli [4] is used to simulate population cell growth in various environmental conditions with different carbon sources.

All fully connected groups of strains that exist in the network are identified using efficient techniques described in [6, 9]. In order to evaluate the bacterial communities that arise

E. Tzamali is with the Institute of Molecular Biology and Biotechnology of the Foundation for Research and Technology-Hellas (FORTH) and the department of Computer Science, University of Crete, Heraklion, Crete, 71110 Greece (phone: 302810391721; fax: 302810391101; e-mail: tzamali@csd.uoc.gr).

P. Poirazi is with the Institute of Molecular Biology and Biotechnology of the Foundation for Research and Technology-Hellas (FORTH) (e-mail: poirazi@imbb.forth.gr).

I. G. Tollis is with the Biomedical Informatics Laboratory (BMI) of the Institute of Computer Science of the Foundation for Research and Technology-Hellas (FORTH) (e-mail: tollis@ics.forth.gr).

M. Reczko is with the Biomedical Sciences Research Center "Alexander Fleming", Varkiza, Greece (e-mail: mreczko@gmail.com).

as cliques in the network, an extended version of the dynamic flux balance model is set up to cope with simulation of heterogeneous cell populations that grow in a mutual environment. In this study, it is further suggested that evolution might support cell communities where metabolically diverse and probably more specialized bacterial cells stably coexist in an attempt to create and maintain more efficient systems. An exhaustive, computational evaluation of all pairs consisting of the wild-type and a single gene knockout cell has shown that metabolic differences are prerequisite for beneficial coexistences [3]. The evaluation of the hypothetical bacterial communities that emerge from the network of potential synergism, as proposed in this study, can identify the existence of efficient, beneficial bacterial communities of any potential size.

II. Network Reconstruction

The active metabolic pathways and consequently the metabolic capabilities (cell phenotype) depend on the specific environmental conditions (food metabolism). Thus, different networks of interactions will arise for each condition. The synergistic networks are constructed independently for each given environmental condition defined by the initial substrate concentrations.

A. Nodes

Different metabolic capabilities for cells of the same species may arise when certain metabolic genes that are actively involved in metabolism are either over- or under-expressed. In this study, single gene knockouts are applied to the wild-type *E. coli* cell, taking into account all genes that are included in the metabolic model. Thus, each gene involved in metabolism is knockout and its growth is simulated on a given environmental condition (nutrition resource). These mutants, as well as the wild-type cell, comprise the nodes of the network and each one carries information about its metabolic capabilities under the same environmental condition. Mutants that cannot grow in the certain environmental condition (lethal mutations) are excluded from the network.

B. Edges

Among all possible interactions between the different strains, those that provide each other with different metabolic capabilities will comprise the edges of the network and should be considered as candidate coexistences for further analysis.

Specifically, during the metabolism of the main source which is initially provided to the population for growth, intermediate byproducts are secreted that might either constrain the fluxes of the metabolic network or prove essential when the main source is exhausted. In that manner, when different mutants grow together these byproducts can play a key role as the fitness landscape is dynamically shaped and new metabolic capabilities are likely to arise. For example, if a participant produces more acetate with the cost of biomass production (specialized participant) then the other participant (if it is greedier) may take advantage of the excess of acetate in the mutual environment and eventually achieve to produce more biomass. The variation of the metabolic capabilities can either be quantitative when certain byproducts are for instance overproduced or qualitative when one potential participant provides to the other novel nutrients. For each node of the network, the maximum concentration values over time for each byproduct are considered. In the quantitative case, only substrates that are eventually consumed by at least one of the two interacting mutants (a strict constraint that can be relaxed) are considered, while in the qualitative case there is no such constraint. Over all byproducts, the maximum absolute relative concentration difference is used as a weight for the interaction between two nodes, leading to the reconstruction of a naturally weighted graph. In that respect, the edges of the reconstructed network are multi-flavor depending on the secondary metabolite in which the metabolic difference between two strains is quantitatively dominant.

Depending on the properties we want to infer from these graphs, a threshold is set that reflects the level above which the concentration differences are considered important. The way in which the interacting candidates will respond to the shaped environment depends on their dynamical metabolic properties (by-production rate, substrate consumption rate, growth rate). The dynamical properties of the community members will ultimately determine the efficiency, by which the shaped environment is utilized in order to produce biomass.

III. Compressed Network Reconstruction

All strains with the exactly same metabolic capabilities under a given initial environmental condition are grouped together. From this partitioning of the nodes to classes, a representative node of each class is arbitrarily chosen to form the compressed network. Each class actually contains all the single gene deletions that have exactly the same effect on cell functioning under a certain condition. It's important to notice that when investigate coexistences among various strains, the environment might be differently shaped and novel substrates may also arise. Therefore, under heterogeneous population growth these gene deletions representing different constraints in the metabolic network might not have the same phenotypic effect. Thus, the members of each class should be one by one evaluated on novel substrates.

The compressed network is reconstructed in order to keep the clique finding problem tractable without losing any valuable information about the ability of the system to form communities of heterogeneous strains.

IV. Building and Evaluating Potential Bacterial Communities

A. Clique Identification

The property of 'difference' as defined in the previous sections is not transitive. When mutants A and B are different and mutants B and C are different, then mutants A and C may or may not be different. Therefore, to form groups of mutants

TABLE I
MUTANT DEGREE DISTRIBUTION OF THE COMPRESSED NETWORKS OF CERTAIN CARBON SOURCES

Carbon	Nodes	Edge density	Node degree distribution										
Arg_L	196	0.1842	k:	16	19	21	26	186	194	195			
			N(k):	7	161	7	2	3	6	10			
4abut	187	0.1635	k:	15	16	183	184	186					
			N(k):	3	168	5	3	8					
Cit	192	0.1875	k:	18	19	20	27	181	186	187	188	189	190	191
			N(k):	8	158	6	1	1	1	2	5	3	2	5
Sucr	192	0.1016	k:	9	10	11	12	187	188	191				
			N(k):	1	176	3	2	1	2	7				
Melib	193	0.1009	k:	7	9	10	11	12	187	188	192			
			N(k):	1	1	178	1	2	1	2	7			
Glc	194	0.1461	k:	10	11	13	14	15	173	182	184	190	192	193
			N(k):	3	3	6	3	164	1	1	3	2	2	6
Ser_D	188	0.1231	k:	8	12	183	185	187						
			N(k):	1	175	4	4	4						
Pyr		0.0834	k:	7	8	182	183	184	186					
			N(k):	2	177	1	2	3	2					
Gly	175	0.0228	k:	2	174									
			N(k):	173	2									
Glyclt	182	0.0110	k:	1	181									
			N(k):	181	1									

that are different in all pairwise comparisons of their metabolic capabilities, cliques (complete sub-graphs) should be identified in the graph. The maximum clique size as well as the number of maximal cliques (cliques not contained in any other clique) of each possible size provides valuable information about the ability of the system to generate variable nutritional environments from metabolic gene knockouts for a given condition.

The maximum clique identification problem is an NP-complete problem, which means that the time required for solving the problem in general increases exponentially with respect to the number of nodes in the graph. The compressed network reconstruction (nodes less than 200) as well as the low edge density (less than 20%) of our graphs, allow the fast identification of all cliques particularly when efficient, exact methods proposed elsewhere are used [9]. Small cliques can be found efficiently using the methods presented in [6].

B. Clique Evaluation

In metabolic simulations, cells are assumed to work greedy in the dynamic environment – optimizing for the best choice at each time step, which simulates adequately a competitive life (competition for food).

This study is based on the dynamic metabolic flux balance analysis model of Varma et al [2]. Flux balance models are constraint-based models that aim to integrate knowledge at different levels in the cascade from genes to proteins and further to metabolic fluxes in a genome-scale metabolic network to describe and understand the overall cellular functions. Flux balance analysis (FBA) models estimate the *optimal* flux distribution of the entire biochemical reacting system, providing a quantitative description of the system when the intracellular fluxes are in balance. When simulating dynamic phenomena within the greedy framework [2], the whole time regime that represents the time of growth in cell populations is properly divided into temporal windows of size δt. The current exchange concentrations ($excC$) that describe the environmental conditions in which populations grow for the certain time interval, properly scaled by the amount of the total biomass (bm) that has been produced

ALGORITHM I

For each time interval δt

For each strain i participating in the heterogeneous population of M strain

maximize μ_i

subject to $\underline{\underline{S}}_i \overrightarrow{Flux_i} = \vec{0}$, $\underline{\underline{S}}_i$: stoichiometric matrix

$\overrightarrow{lb_i} \le \overrightarrow{Flux_i} \le \overrightarrow{ub_i}$, $\overrightarrow{lb_i}$: lower bound, $\overrightarrow{ub_i}$: upper bound

$bm_i(t) = bm_i(t-\delta t)e^{\mu_i \delta t}$

end

$\overrightarrow{excC}(t+\delta t) = \overrightarrow{excC}(t) - \sum_{i=1}^{M} \overrightarrow{excFlux_i} \cdot \frac{bm_i}{\mu_i}\left(1-e^{\mu_i \delta t}\right)$

$bm = \sum_{i=1}^{M} bm_i$

$\overrightarrow{excBounds} = \overrightarrow{excC}/(bm \cdot \delta t)$

end

where

$excFlux_i$ is the $Flux_i$ of the exchange reactions

$excBounds$ is the ub of the exchange reactions

TABLE II
MUTANTS OF MAXIMUM CONNECTIVITY

Carbon source	Single gene deletions						
Arg_L	'b0116'	'b2903'	'b2276'	'b3236'	'b0721'	'b0451'	
4abut	'b0116'	'b2903'	'b2276'	'b3236'		'b0451'	
Cit	'b0116'	'b2903'	'b2276'		'b0721'		'b3731'
Glc	'b0116'	'b2903'	'b2276'	'b3236'	'b0721'		
Glyclt			'b2276'				
Gly			'b2276'				
Melib	'b0116'	'b2903'		'b3236'			'b3731'
Sucr	'b0116'	'b2903'		'b3236'			'b3731'
Ser_D	'b0116'	'b2903'			'b0721'	'b0451'	
Pyr	'b0116'	'b2903'					

TABLE III
METABOLIC INFORMATION OF THE MUTANTS OF MAXIMUM CONNECTIVITY

gene	rxns	subsystems
'b0116'	'2 Oxogluterate dehydrogenase' 'Glycine Cleavage System' 'pyruvate dehydrogenase'	'Citric Acid Cycle' 'Folate Metabolism' 'GlycolysisGluconeogenesis'
'b2903'	'Glycine Cleavage System'	'Folate Metabolism'
'b2276'	'NADH dehydrogenase ubiquinone 8 35 protons ' 'NADH dehydrogenase menaquinone 8 2 protons ' 'NADH dehydrogenase demethylmenaquinone 8 28 protons '	'Oxidative Phosphorylation'
'b3236'	'malate dehydrogenase'	'Citric Acid Cycle'
'b0721'	'succinate dehydrogenase'	'Citric Acid Cycle' 'Oxidative Phosphorylation'
'b0451'	'ammonia reversible transport'	'Transport Extracellular'
'b3731'	'ATP synthase four protons for one ATP '	'Oxidative Phosphorylation'

shape the actual boundaries of the uptake fluxes (*excBounds*). Optimum operation within each time interval is assumed for the system to effectively reach its goal of growth and development.

To evaluate the growth of the possible bacterial communities the model of Varma et al [2] is extended to embody heterogeneous cell population growth in a mutual environment (Algorithm I). In heterogeneous cell populations, each population grows respecting the constraints that its network imposes as well as the dynamically shaped environment. When none of the different populations can grow further in the shaped medium the simulation terminates.

V. RESULTS

Simulations are performed using the genome-scale metabolic model of *E. coli* (*i*JR904) by Reed et al [4]. All simulations are performed using the COBRA toolbox [5] suitably modified to incorporate dynamic growth of interacting heterogeneous populations as described in Algorithm I. The initial biomass (0.003 g/lt) is equally distributed to the strains of the heterogeneous population under study. A carbon source of initial concentration 10mmol/lt is provided to the system for growth. Oxygen and ammonia are assumed to be in excess. In different carbon sources, strains exhibit different metabolic capabilities, thus different networks of potential synergism are produced.

The results that are presented here correspond to the networks obtained after omitting all edges below a threshold of 0.6 (60% relative 'difference') in their corresponding weighted graphs. Structural properties of the reconstructed networks such as the node degree distribution and the number of all and maximal cliques that have been found are presented in order to help our understanding of the metabolically different capabilities of these systems. All the structural properties presented here concern the compressed networks.

Table I presents the number of nodes N(k) of degree k for each different carbon source we have studied. Lethal mutations are environmental specific, thus each network might consist of different mutants. The number of nodes of each of the compressed networks is presented in Table I as well as the graph density. For each environmental condition there are certain mutants with the highest possible node degree (k=nodes-1) - maximum connectivity. The group of mutants with this property actually forms a clique of the size of the group and represents the most metabolically different bacterial community for the certain environmental condition. Certain mutants are identified to consistently participate in these bacterial communities under most environmental conditions implying a pattern of growth-condition-invariant strains that are consistently metabolically different from the rest mutants. Table II presents the strains of maximum connectivity that are commonly present in more than two conditions. The metabolic reactions in which the certain genes participate in as well as the metabolic subsystems in which they are involved are

summarized in Table III.

The clique size distribution is also constructed for each environmental condition. The number of secondary metabolites that can be by-produced during the metabolism of the main source as well as the number of strains that arise to

TABLE IV
UPPER BOUNDS OF DIVERSITY

Carbon source	# byproducts	#Nodes of highest degree	Min maximal size	Max clique size
Arg_L	8	10	14	16
4abut	10	8	11	12
Cit	7	5	10	12
Sucr	9	7	9	10
Melib	9	7	8	10
Glc	7	6	9	10
Ser_D	6	4	7	8
Pyr	4	2	4	6
Gly	2	2	3	3
Glyclt	2	1	2	2

metabolically differ from most strains (highly connected nodes), indicate the maximum clique size that can be found in the network. These properties are summarized in Table IV and actually place the upper bounds of diversity that can be emerged in certain conditions from single-cloned populations of *E. coli*.

All small size-cliques are found to be parts of larger cliques. Whether this fact means that the maximal cliques of saturated capabilities are more beneficial or stable coexistences than their constituent sub-cliques comprises an issue for further investigation.

In the following results from aerobic growth on *pyruvate* are thoroughly presented.

Case study: Aerobic growth on pyruvate

The *pyruvate* graph consists of 383 mutants in its full representation and of 187 nodes in the corresponding compressed representation.

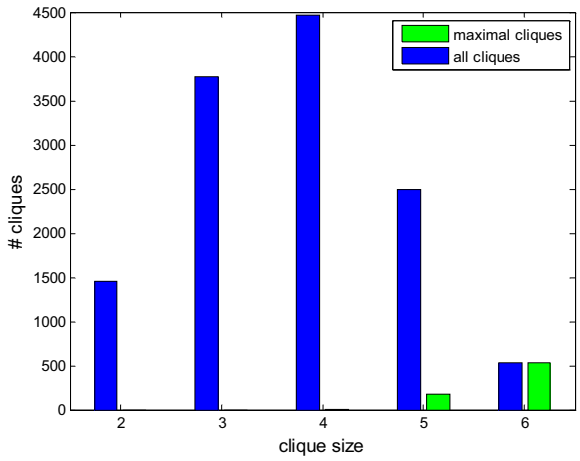

Fig. 1 All and maximal clique-size distribution for growth on *pyruvate*

The compressed network consists of 1451 edges in total for the chosen threshold of 0.6, which is a sparse network with 8% graph density. The node degree distribution of the *pyruvate* network is shown in Table I. As this threshold changes, minor changes arise in the way the node degrees are distributed. The degree distribution (Table I) shows the existence of few nodes of high degree and many nodes of low degree. Actually, this indicates that there are few nodes that consistently differ from all the rest nodes of the graph, therefore these few nodes show high connectivity (preferential attachment to 8 high degree nodes).

The *pyruvate* network is capable of forming bacterial communities with up to 6 mutually different strains. The number of these potentially synergistic communities is shown in Fig. 1. Most cliques are of size 4. Maximal cliques are also presented in Fig. 1.

To quantitatively describe superior performance in heterogeneous bacterial population systems, the overall growth performance of the heterogeneous population is compared with the best among all the performances of the homogeneous strain populations under the same initial conditions. If the heterogeneous system can achieve a better outcome (biomass) utilizing the given limited amount of resources then it is considered more efficient, thus beneficial.

All cliques have been evaluated. From each clique size the most efficient group and its corresponding performance are presented in Table V. The performance of the homogeneous wild-type population is also presented as a reference. The

TABLE V
PERFORMANCE OF THE MOST EFFICIENT CLIQUES

strains	Total biomass	Benefit (%)	Benefit_{WT} (%)
WT	0.23654	- 2.1	0.0
'b3403'	0.24169	0.0	2.1
'b3403b0721'	0.24891	2.9	5.2
'b2903b3403b0721'	0.2435	0.7	2.9
'b2276b2903b3403b0721'	0.2418	0.0	2.2
'b1982b2903b3403b0721'	0.2418	0.0	2.2
'b2276b1982b2903b3403b0721'	0.2403	- 0.5	1.5
'b0116b2276b1982b2903b3403b0721'	0.2342	- 3.1	- 0.9

benefit is calculated as the relative difference of the final biomass that is produced by the group of strains with respect to the best homogeneous population performance, which is the mutant coming from the knockout of the gene 'b3403' for growth on *pyruvate*. However, the benefit with respect to the growth performance of the homogeneous wild-type population is presented as well in Table V.

As Table V shows, none of the hypothetical bacterial communities is maximal; only the last (last row in Table V), which has the maximum clique size and by definition is maximal. Apart from the mutant of the gene deletion 'b3403' all the other mutants of the group are of high degree. This might be expected although it means that these high degree nodes significantly differ from each other as well. The most efficient bacterial community for aerobic growth on *pyruvate* consists of the pair of 'b3403' and 'b0721' mutants. Any new member that is added in this efficient pair provides no further

improvement regarding the benefit. Thus, the pair is saturated with respect to efficiency. However, all cliques except the one of the maximum size (last row in Table V) have better growth performance than the homogeneous wild-type population.

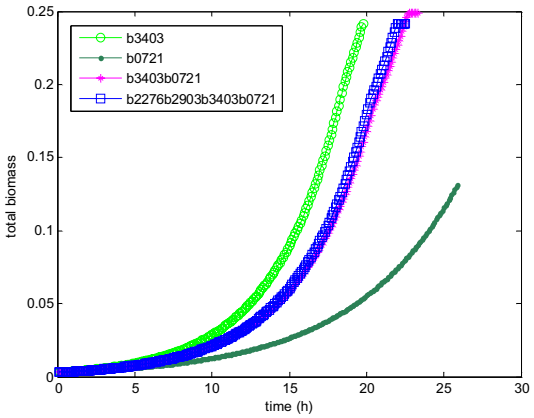

Fig. 2 Total biomass production over time of the homogeneous populations of strains b3403 and b0721 as well as the best cliques of size 2 and 4 for growth on *pyruvate*

Acetate, *formate*, *glycine* and *glycolate* comprise all the intermediate nutrients that can be produced during the *pyruvate* metabolism. The maximum concentrations values

TABLE VI
MAXIMUM CONCENTRATION VALUES OF CERTAIN BYPRODUCTS

KO	$[acetate]_{max}$	$[formate]_{max}$	$[glycine]_{max}$	$[glycolate]_{max}$
'b3403'	2.1830	8.3576	0	0.01208
'b0721'	7.6501	0.4541	0	0
'b2903'	1.4528	9.4621	0.0482	0.01173
'b2276'	5.1896	5.7116	0	0
'b1982'	0.1202	0.7617	0	0.00082
'b0116'	1.4742	10.8052	3 E-05	0.01015
WT	1.4732	9.4604	0	0.01155

that each of the strain that participates in the hypothetical synergistic bacterial community can produce are shown in Table VI.

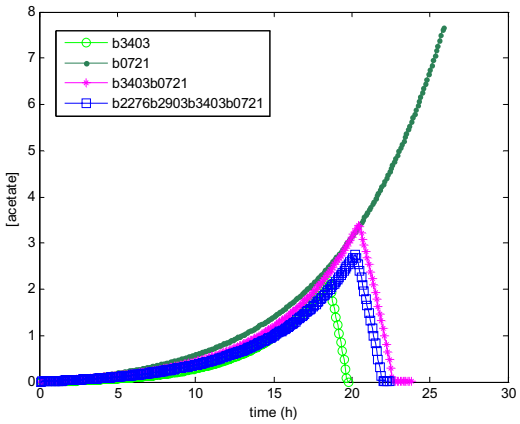

Fig. 3 *Acetate* metabolism of the homogeneous populations of strains b3403 and b0721 as well as the best cliques of size 2 and 4 for growth on *pyruvate*

Glycine is not produced by any of 'b3403' and 'b0721' mutants, which means that this substrate does not play a key role in efficient growth. *Acetate* though, seems to be an essential byproduct. The mutant 'b0721' is an *acetate* specialist (Fig. 3), a specialization that comes with the cost of its growth performance as shown in Fig. 2.

Furthermore, the specific mutant is not capable of consuming the acetate it produces. Fig. 3 shows that the best clique of size 4 eventually produces less *acetate* than the best pair, thus the pair benefits when *pyruvate* is exhausted. The dynamics (by-production rate, consumption rate, growth rate) as shaped by the network constraints and the medium will eventually determine the efficiency of the system to utilize the available nutrients in order to produce biomass.

VI. DISCUSSION

The reconstructed network of potential synergism and the resulting structural sparseness observed, comprise an efficient method for exploring the search space of all possible mutant pairs, triples or any multiples for beneficial growth performance even though the implied clique identification problem is an NP decision problem. Features of the network such as the degree distribution, the maximum clique size, the maximal clique size distribution can be used in order to compare networks for different environmental conditions. Furthermore, these network properties place limits to the diversity that can emerge from single-cloned populations of *E. coli* when single knockouts are allowed.

The network reconstruction is based on the hypothesis that metabolic interactions may explain the emergence and maintenance of heterogeneous bacterial populations. Is it possible from growth simulations of homogenous populations to predict growth performance of heterogeneous populations? The interactions between cells of various metabolic capabilities allow them to satisfy each other's nutritional needs. This change in the supply and demand conditions can shape the environment and the fitness landscape in a different way than the one experienced by homogeneous populations. In that respect, new metabolic capabilities may arise for each of the participants. A qualitative difference (or equivalently a maximum quantitative difference) between two strains appears when novel nutrients are produced by at least one of the participants. Novel nutrients may constrain the fluxes of the metabolic network in an unexpected manner, change the dynamics and produce unexpected phenotypes which are not trivial to predict. Simulation of mutants can be conducted in an environment where novel byproducts are present and tested for changes in their metabolic capabilities. On the other hand, the dynamical metabolic properties of the participants play a major role in the final performance of a population and can be used as features to improve predictability.

While searching for efficient bacterial communities, the necessary but not sufficient condition that arises for all involved participants is to have quantitative differences with each other. This rationale extends the undirected graph we

formed into a directed one where bidirectional relations are considered. Another extension of the current network reconstructions would be the incorporation of nodes that represent over-expressed genes to reflect the natural sources of variability more accurately. Our simulations already show that efficiency might be one explanation to develop and maintain biodiversity. Robustness in changing ecological conditions might also be proved as a source of diversity using the suggested approach.

This study broadens our perspective regarding the emergence of co-existence in bacterial populations from single cells to cell communities, investigates interactions between cells of various metabolic capabilities, and comprises a step towards understanding biodiversity. Furthermore, the identification of heterogeneous bacterial cultures with superior desired properties might further exhibit a broad range of applications in metabolic engineering.

ACKNOWLEDGMENT

This work was supported by the Greek General Secretariat of Research and Technology under the PENED 03ED842, grant "Development of computational methods for genomic data analysis".

REFERENCES

[1] D. Treves, S. Manning, and J. Adams, "Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of Escherichia coli," Mol. Biol. Evol., vol. 15, no. 7, pp. 789-797, 1998.

[2] A. Varma, and B.O. Palsson, "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110," Appl. Environ. Microbiol., vol. 60, no. 10, pp. 3724-31, 1994.

[3] E. Tzamali, and M. Reczko, "The benefit of cooperation: Identifying growth efficient interacting strains of Escherichia coli using metabolic flux balance models," *8th IEEE International conference on bioinformatics and bioengineering*, Greece, 2008.

[4] J.L. Reed, et al., "An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR)," Genome Biol, vol. 4, no. 9, pp. R54, 2003.

[5] Scott A Becker, Adam M Feist, Monica L Mo, Gregory Hannum, Bernhard Ø Palsson & Markus J Herrgard, "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox," Nature Protocols, vol. 2, pp. 727-738, 2007.

[6] Janet M. Six, Ioannis G. Tollis, "Effective Graph Visualization Via Node Grouping," infovis, pp.51-58, 2001 IEEE Symposium on Information Visualization (InfoVis 2001), 2001

[7] Paul Rainey, Angus Buckling, Rees Kassen and Michael Travisano, "The emergence and maintenance of diversity: insights from experimental bacterial populations," Tree, vol. 15, pp. 243-247, 2000.

[8] R. F. Rosenzweig, R. R. Sharp, D. S. Treves, and J. Adams, "Microbial Evolution in a Simple Unstructured Environment: Genetic Differentiation in Escherichia Coli," Genetics, vol. 137(4), pp. 903–917, 1994

[9] Patric R. J. Ostergard, "A New Algorithm for the Maximum-Weight Clique Problem," Electronic Notes in Discrete Mathematics, 6th Twente Workshop on Graphs and Combinatorial Optimization, vol. 3, pp. 153-156, 1999.