

Full Length Research Paper

Insight into microevolution of *Streptomyces rimosus* based on analysis of *zwf* and *rex* genes

Zhenyu Tang¹, Paul R. Herron², Iain S. Hunter², Siliang Zhang¹ and Meijin Guo^{1*}

¹State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai 200237, P. R. China.

²Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, G4 0RE, Scotland, United Kingdom.

Accepted 22 May, 2012

Streptomyces rimosus has greatly influenced human history as a producer of many important secondary polyketide metabolites, such as oxytetracycline (OTC). The traditional screen and mutation program has resulted in a dramatic increase in OTC production. The availability of multiple semi-complete genome sequences of *S. rimosus* facilitates attempts to systematically address basic questions in genome evolution. We refer to such efforts as micro-evolutionary analysis. We report the results of comparative analysis of semi-complete genome sequences of three *S. rimosus* strains from the genealogy map (G7, M4018 and 23383) with different OTC production levels using genome comparison (single nucleotide polymorphisms, SNPs) method. These data were used to assess the influence of microevolution on the physiology, genetics and evolution of *S. rimosus*. Some SNPs were found in primary metabolism related genes which might affect the final OTC production. We further discussed the microevolution of primary metabolite genes (Glucose-6-phosphate dehydrogenase, *zwf*) and regulatory genes (redox regulator, *rex*) in *S. rimosus*. Using SOLEXA sequencing data, the phylogenetic trees of *zwf* and *rex* were constructed. The results indicate that *S. rimosus* is closely related with *Streptomyces albus* and represents a distinct evolutionary lineage compared with other *Streptomyces*. Our research cannot only provide important information for genotyping and evolutionary research of *S. rimosus*, but can also make possible the development of an informed view of genotype and phenotype.

Key words: Microevolution, *Streptomyces rimosus*, single nucleotide polymorphisms (SNPs), glucose-6-phosphate dehydrogenase, redox sensor.

INTRODUCTION

Evolution can be divided into two categories: macroevolution and microevolution. Macroevolution focuses on changes that occur within and among populations, while microevolution refers to smaller evolutionary changes within a species or population, including the genetic composition of a population (Hendry and Kinnison, 2001). Based on this theory, microevolution normally occurs over shorter intervals. In

the past few years, studies of microevolution have made a transition from the evolutionary synthesis to new levels of melioration, and now are the symbol of the evolution outcome.

Microevolution studies tend to focus on polymorphism, which refers to variation within a population. This may evolve a lot of mechanisms, such as mutations, genome rearrangement, gene duplication, transposition, and homologous and non-homologous recombination (Lawrence and Hendrickson, 2003). For mutation, it is defined as an alteration in DNA sequence, including both point mutations where one base pair of DNA is substituted for another, and insertions and deletions

*Corresponding author. E-mail: guo_mj@ecust.edu.cn. Tel: +86 21 64251131. Fax: +86 21 64253702.

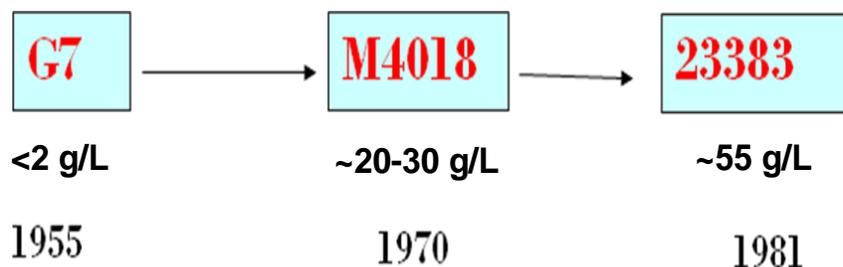


Figure 1. Genealogy map of *S. rimosus*. Original strain G7, OTC 2 g/L, isolated in 1955; type strain M4018, OTC 20-30 g/L, isolated in 1970; production strain 23383, OTC 55 g/L, isolated in 1981.

(Indels) where one or more base pairs of DNA are inserted into, or deleted from, a DNA sequence. These are also called single nucleotide polymorphisms (SNPs). However, SNPs can be divided into two groups, synonymous SNPs (sSNPs) and non-synonymous SNPs (nSNPs). The former refers to the changes in DNA sequence which does not affect the amino acid sequence, while nSNPs can cause the alteration of amino acid sequence and change the function of proteins. nSNPs are the direct factors that affect the biological properties. Genome sequencing has become central to the study of microevolution (Hughes, 1999). The increasing availability of sequencing technologies and bioinformatics tools has accelerated the genome comparison research. Multiple complete or nearly complete genome sequences from very closely related species, and even strains of the same species are now available in the database. The use of such data here to compare the sequences of individual genes allowed for the systematic analysis of several fundamental questions about bacterial genome microevolution and would help us to understand how single gene evolved.

The soil-borne *Streptomyces* has been known for some time to be good producers of many secondary metabolites, such as antibiotics (Lauren et al., 2009). The aromatic polyketide antibiotic, oxytetracycline (OTC), is produced by *Streptomyces rimosus* as an important secondary polyketide metabolite. OTC is used particularly heavily in aquaculture, where 500 kg doses may be used in one treatment (Kerry et al., 1995). Although the clinical use of the tetracyclines has declined in recent years due to the emergence of resistant strains of bacteria, OTC remains the choice of some infection cases, including *Rickettsia*, *Chlamydia*, and mycoplasma in penicillin-sensitive patients incapable of tolerating macrolides (Chopra et al., 1992). Although the traditional screen and mutation program has resulted in a dramatic increase in OTC production from 2 to 55 g/L (Figure 1), what is responsible for this is unknown. Furthermore, the microarray data showed that there were no gene duplications and deletions across the three strains (Kirby et al., 2008). It suggests that point mutations may be the

most probable variation in any changes that allow higher OTC production. Recently, the three *S. rimosus* genomes in the genealogy were all Solexa sequenced by Iain S. Hunter's group in University of Strathclyde.

Glucose-6-phosphate dehydrogenase (G6PDH) is the housekeeping enzyme that is encoded by *zwf* genes (*zwf1* and *zwf2*) in pentose-phosphate pathway (PPP). It catalyzes the reaction of glucose-6-phosphate to 6-phosphogluconatone and generates reducing power NADPH for OTC biosynthesis and other reactions; while *rex* (a redox regulator) was firstly characterized in *Streptomyces coelicolor* and *S. rimosus* (Brekasis and Paget, 2003; Shen et al., 2012). It is a transcriptional regulator that responds directly to the balance of the NADH/NAD⁺ redox system.

In this paper, in order to assess the influence of microevolution on the physiology, genetics and evolution of *S. rimosus*, we report the results of comparative analysis of semi-complete genome sequences of three *S. rimosus* strains from the genealogy map (G7, M4018 and 23383) with different OTC production levels using genome comparison (SNPs) method, including the microevolution analysis of primary metabolite genes (*zwf1* and *zwf2*) and a regulatory gene (*rex*). It does not only help explore the effect of genome sequence changes on the final phenotype, but also understand the molecular mechanisms of multiple gene interactions and network regulations.

METHODOLOGY

Genome sequencing

The whole genome sequencing was performed by applying the SOLEXA Illumina sequencer (GATC, Germany). The sequenced strains were G7, M4018 and 23383.

Bioinformatics analysis

Assembly and comparison of sequences (G7, M4018 and 23383) were achieved by using Maq software which can build mapping assemblies from short reads generated by the next-generation

Table 1. The SNPs related to the primary metabolism in *S. rimosus* 23383.

Primary metabolism related genes	SNPs position	Nucleotide sequence changes	Amino acid changes
Phosphoglycerate mutase	576	G to A	Ser to Asn
Phosphoglycerate kinase	119	G to A	Gly to Asp
Acyl-CoA synthetase	310	G to A	Asp to Asn
Aldehyde dehydrogenase	721	G to A	Gly to Arg
Alcohol dehydrogenase	1055	G to A	Gly to Asp
Glucose-6-phosphate 1-dehydrogenase	139	C to T	Pro to Ser

sequencing machines. Maq first aligns reads to reference sequences (M4018) and then calls the consensus. At the mapping stage, maq performs ungapped alignment. At the assembling stage, maq calls the consensus based on a statistical model. It calls the base which maximizes the posterior probability and calculates a phred quality at each position along the consensus. All the SNPs and Indels generated from G7, M4018 and 23383 were predicated by Maq during the mapping as well.

Phylogenetic analysis

The phylogenetic trees of *zwf1*, *zwf2* and *rex* encoded proteins from *S. rimosus* M4018 were carried out using the Neighbor-Joining algorithm from NCBI website. The neighbor-joining method is a special case of the star decomposition method (Saitou and Nei, 1987). The raw data are provided as a distance matrix and the star tree is the initial tree. Then, a modified distance matrix is constructed in which the separation between each pair of nodes is adjusted on the basis of their average divergence from all other nodes. The tree is constructed by linking the least-distant pair of nodes in this modified matrix. When two nodes are linked, their common ancestral node is added to the tree and the terminal nodes with their respective branches are removed from the tree. This pruning process converts the newly added common ancestor into a terminal node on a tree of reduced size. At each stage in the process, two terminal nodes are replaced by one new node. The process is complete when two nodes remain, separated by a single branch.

Rex protein structure prediction

SWISS-MODEL (<http://swissmodel.expasy.org>) which is a server for automated comparative modeling of three-dimensional (3D) protein structures was used to predict the Rex protein structure. In the 'alignment mode', the modeling procedure is initiated by submitting the Rex sequence (*S. rimosus* M4018) alignment file. We specified which sequence in the given alignment was the target sequence and which one corresponded to a structurally known protein chain from the ExPDB template library. At last, the server would build the model based on the given alignment.

RESULTS

Microevolution analysis of *S. rimosus* with different OTC production levels

Using *S. rimosus* M4018 genome sequence as the reference, 78 and 615 SNPs were found in *S. rimosus* G7

and *S. rimosus* 23383 genomes respectively by Maq software. However, there were no SNPs and Indels identified in OTC gene clusters, while some SNPs were found in primary metabolism related genes in *S. rimosus* 23383 as shown in Table 1. The SNPs positions and amino-acid changes are all listed in the table. These SNPs are all belonged to nSNPs.

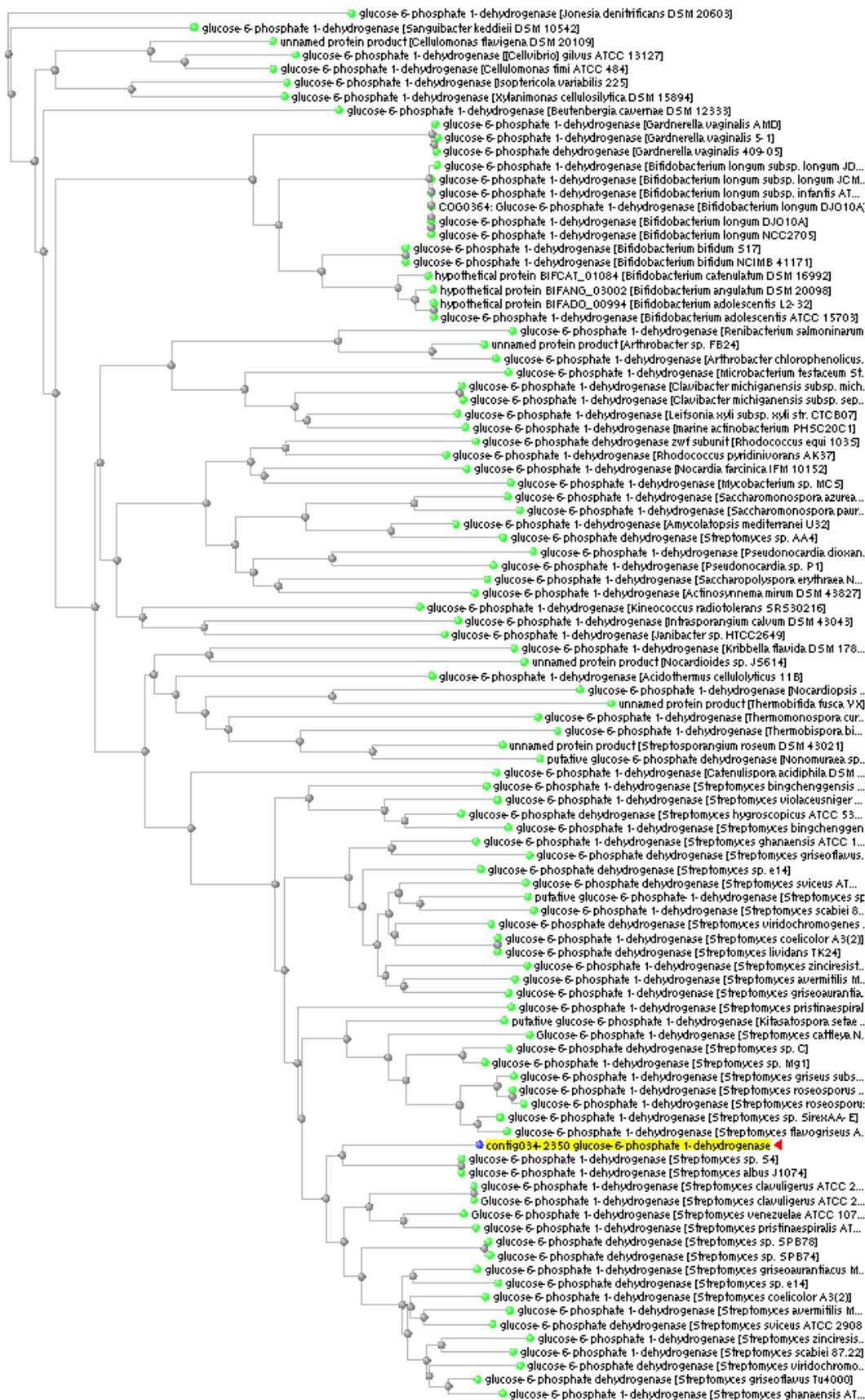
Microevolution analysis of glucose-6-phosphate dehydrogenase gene

In *Streptomyces*, housekeeping genes are normally stable and conserved, so it can well demonstrate the microevolution relationship between different species using phylogeny tree analysis.

As shown in Figure 2, the phylogenetic trees of *zwf1* and *zwf2* form two different groups; while *Streptomyces albus* is located at the adjacent branch of both trees. The GC contents of *zwf1* and *zwf2* in *S. rimosus* are 68 and 73%, respectively. Besides, the ZWF protein sequences share high homology with other *Streptomyces* species by blast analysis, 85 to 90% for *zwf1* and 76 to 86% for *zwf2*.

Microevolutionary analysis of *rex* gene

Rex (Redox regulator, Rex) characterized in *S. rimosus* is a transcriptional regulator that responds directly to the poise of the NADH/NAD⁺ redox (Shen et al., 2012). The *S. rimosus* Rex protein sequence shares high homology with those of four *Streptomyces* species by blast analysis: *S. coelicolor* A3(2) (84%), *Streptomyces avermitilis* MA-4680 (84%), *Streptomyces griseus* (80%) and *Streptomyces lividans* TK24 (71%). Meanwhile, database searches revealed that Rex-related proteins were encoded by the genomes of most Gram-positive bacteria, including *B. subtilis* (38% identity), *Bacillus anthracis* (38% identity), *Listeria monocytogenes* (39% identity). The phylogenetic tree showed that *Sr-Rex* was located at the edge of the tree (Figure 3), which concludes that it is different from other Rex proteins. The protein sequence data also confirms that *Sr-Rex* is much longer than other Rex proteins from homologues species (Shen et al., 2012). Kirby (2008) analyzed SSU ribosomal RNA



a

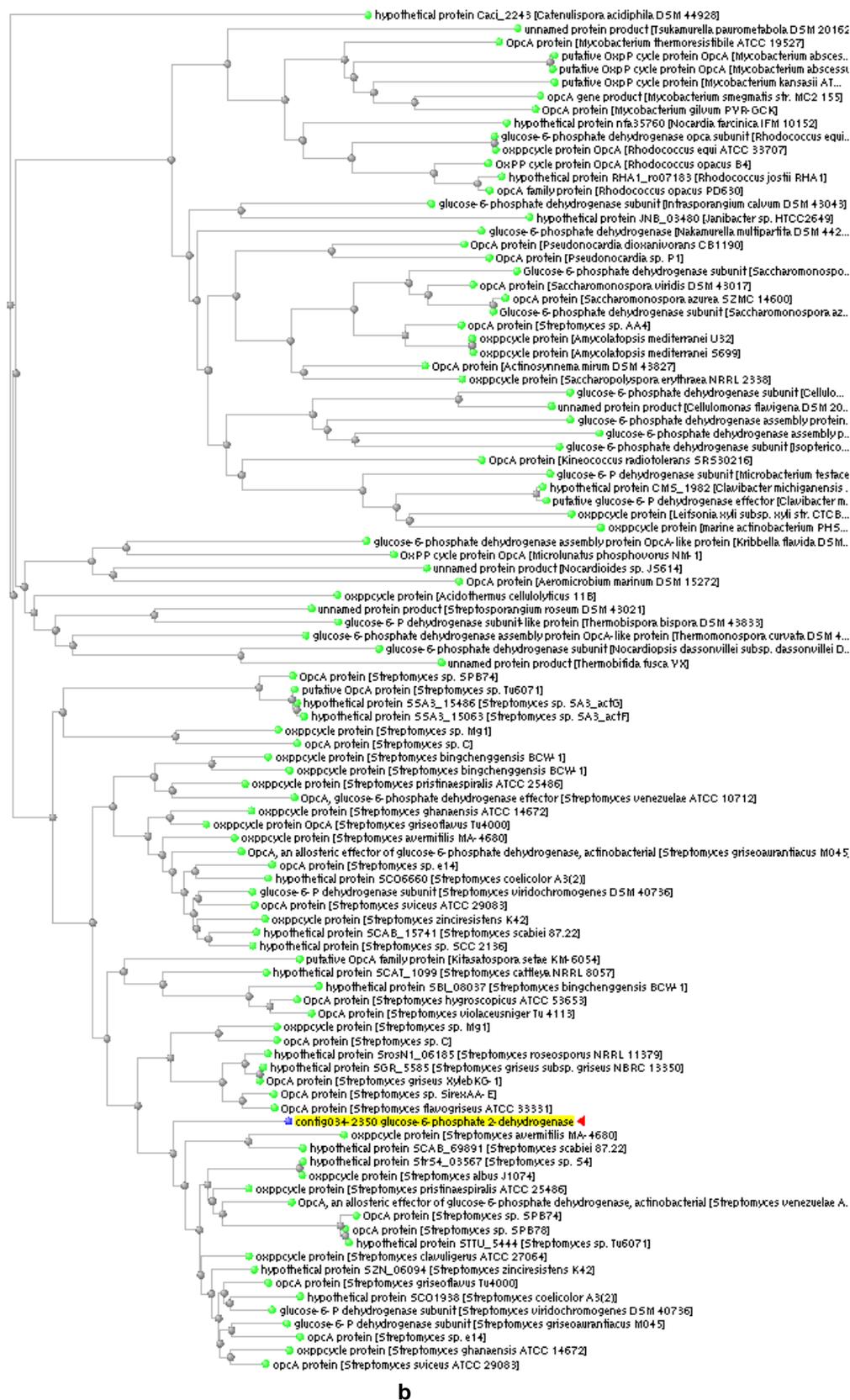


Figure 2. The phylogenetic trees of glucose-6-phosphate dehydrogenase from *S. rimosus* M4018 (a) ZWF1 (b) ZWF2. They were constructed using the Neighbor-Joining algorithm from NCBI website. The ZWF proteins from *S. rimosus* M4018 are highlighted in yellow.

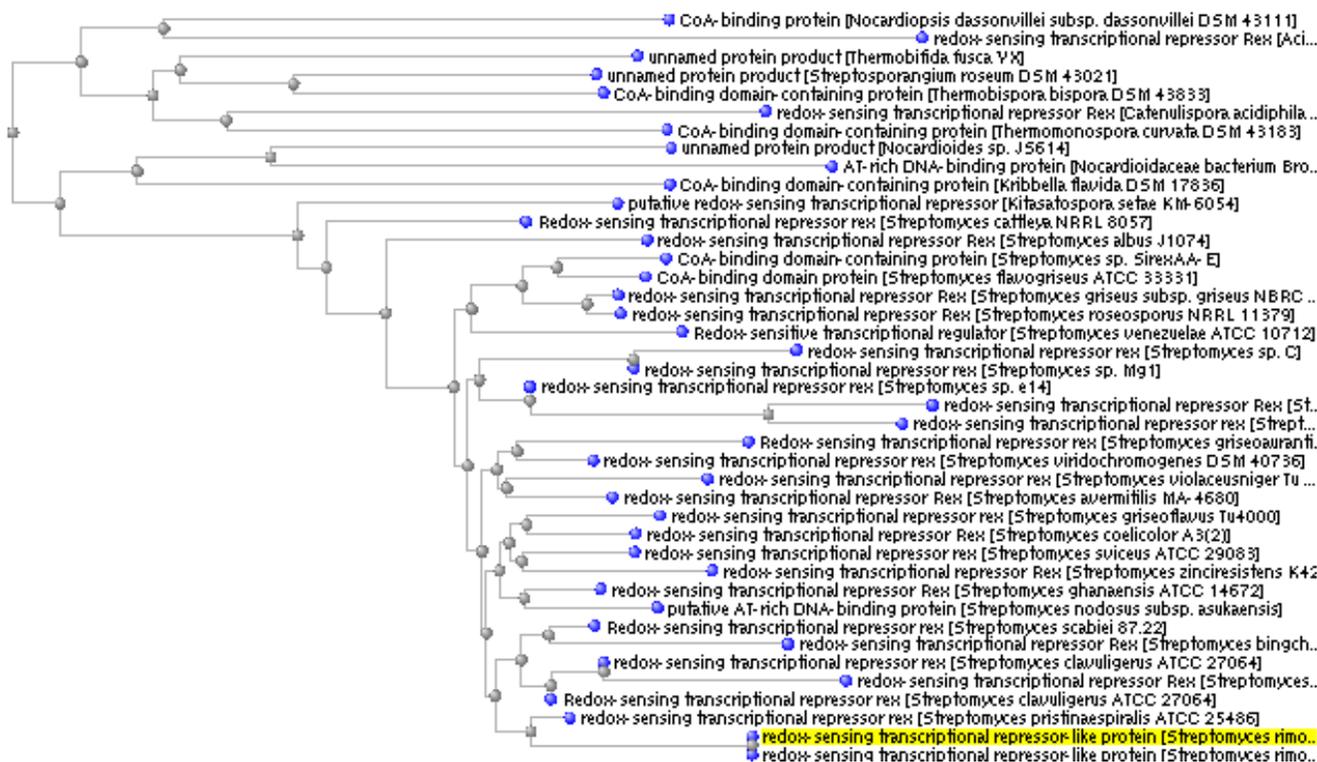


Figure 3. The phylogenetic tree of *rex* gene from *S. rimosus* M4018. It was constructed using the Neighbor-Joining algorithm from NCBI website. The Rex protein from *S. rimosus* M4018 is highlighted in yellow.

phylogeny of *S. rimosus*, the results showed that the species is positioned at the edge of the *Streptomyces* clade as well.

DISCUSSION

With the development of sequencing technology, whole genome sequencing of bacteria becomes more and more popular and advances the microevolution analysis. In this study, we report the results of comparative analysis of semi-complete genome sequences of three *S. rimosus* strains (G7, M4018 and 23383) with different OTC production levels using genome comparison (SNPs) method. Some SNPs were found in primary metabolism related genes. As primary metabolism can provide precursors (e.g. CoenzymA), energy (e.g. ATP), reducing power (e.g. NADPH) for the secondary metabolites (e.g. OTC), we assumed that these changes might affect the final OTC production. The microevolution of primary metabolite genes (glucose-6-phosphate dehydrogenase, *zwf*) and regulatory genes (redox regulator, *rex*) in *S. rimosus* indicate that *S. rimosus* is closely related with *S. albus*. The *zwf* genes of *S. rimosus* and *S. albus* may come from the same ancestor and undergo a long evolution time. The difference GC contents of *zwf1* and *zwf2* suggest that they may

incorporate into the genome in different time. Thus, we can conclude that *S. rimosus* represents a distinct evolutionary lineage compared with other *Streptomyces*.

All the nSNPs found in *S. rimosus* 23383 are involved in the conserved domains of the proteins. For the point mutation in phosphoglycerate mutase (Ser to Asn), it might influence the glycolysis pathway and the reversible interconversion of 3-phosphoglycerate to 2-phosphoglycerate, while the mutation in phosphoglycerate kinase (Gly to Asp) might affect the formation of ATP to ADP. For the nSNP in acyl-CoA synthetase (Asp to Asn), it could change the lipid synthesis, fatty acid catabolism, and remodeling of membranes. The point mutation in alcohol dehydrogenase (Gly to Asp) and aldehyde dehydrogenase (Gly to Arg) might have an impact on the conversion of ethanol to harmless acetic acid. Glucose-6-phosphate 1-dehydrogenase is the first enzyme in PPP and is involved in the production of reducing power NADPH, so the mutation (Pro to Ser) could cause the significant change and affect the OTC production (the last three steps in the biosynthesis of OTC need a lot of NADPHs).

In *S. rimosus*, *zwf1* and *zwf2* are two isoforms of G6PDH. Their amino acids and nucleotide sequences are different, so it may be concluded that they cannot be replaced by each other. By disruption of *zwf1* or *zwf2*, the

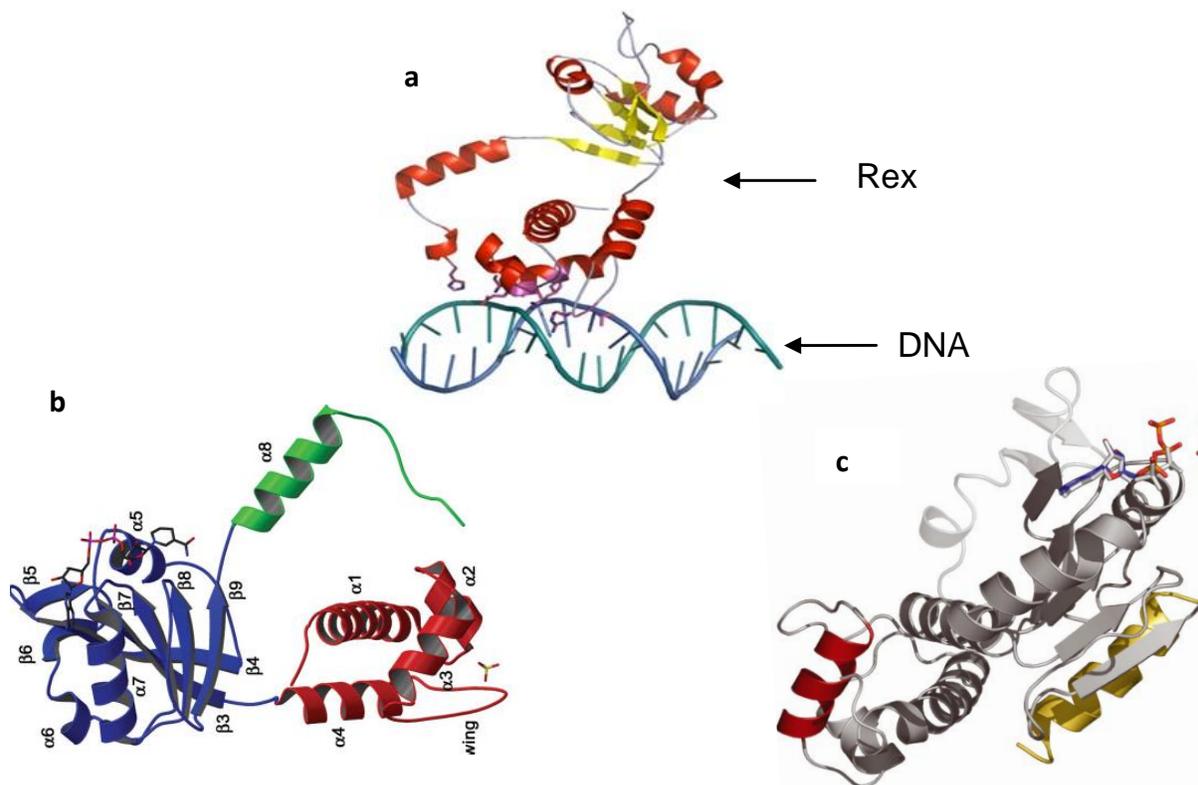


Figure 4. The 3D structure of Rex. (a) A ribbon representation of the 3D structure of *Sr-Rex* was constructed by Swiss-Model, alignment mode. The binding of *Sr-Rex* protein with DNA is presented in this figure as well. (b) Model for *T-Rex*/DNA Recognition (Nakamura et al., 2007). (c) Overall crystal structure of the *B-Rex* (Wang et al., 2008).

specific OTC productivity was increased by 66 or 33% compared with the wild type control. Meanwhile, the biomasses of *zwf* gene intensified mutants (*zwf1*⁺ and *zwf2*⁺) were 20 or 10% more than their *zwf* gene disrupted counterparts (Tang et al., 2011). All these data indicate that *zwf1* contributes more to biomass formation and OTC production than *zwf2* does.

As the production of antibiotics is the outcome of multiple genes interactions and network regulations, regulatory genes are involved as well. *Sr-Rex*, a novel redox-sensitive repressor in *S. rimosus* M4018, which appears to modulate transcription in response to changes in cellular NADH levels, was discovered recently (Shen et al., 2012). The highly conserved phylogenetic sequence suggests common structural mechanisms for redox-dependent gene regulation among Rex family members. Previously, Brekasis and Paget (2003) demonstrated that NADH dissociated a *Sc-Rex*/ROP complex, as well as *B-Rex* homologs in *B. subtilis* (Gyan et al., 2006). Furthermore, there is a common motif GlyXGlyXXGly which is important for the NADH binding in all *rex* genes. We find the same sequence in *Sr-rex* which is Gly₁₀₀-Ile₁₀₁-Gly₁₀₂-Asn₁₀₃-Leu₁₀₄-Gly₁₀₅. By further investigation, it suggests that NADH can inhibit the Rex and ROP binding. However, NAD⁺ has no effect on the REX-ROP

complex formation. This was consistent with the researches in *S. coelicolor* (Brekasis and Paget, 2003) and *B. subtilis* (Ellen et al., 2008). As shown in Figure 4, by Swiss-Model, the final structure of the predicted *Sr-Rex* model is similar to *T-Rex* (Nakamura et al., 2007) and *B-rex* (Wang et al., 2008).

Moreover, we found that some SNPs of regulatory genes existed in *S. rimosus* G7 and *S. rimosus* 23383 genomes. For example, a *Streptomyces* antibiotic regulatory protein (SARP) was found downstream the *otrB* gene in the OTC biosynthesis cluster. Compared with *S. rimosus* 4018, there is a point mutation in this gene which mutates the amino acid from leucine into phenylalanine. Since this leucine is conserved in all the SARPs, it might affect the regulation function of this gene. This work is under study in Paul R. Herron's group in University of Strathclyde.

To understand the role of microevolution in biological control, more data need to be exploited. Thus, it may be possible to guide the evolution into efficient ways. From the DNA sequence alterations (SNPs and Indels), after we get the information of microevolution, all the data related to genotype, phenotype, transcriptional levels will be correlated in an informed view. A selected mutation identified in the production strain (23383) will be

introduced by recombinant techniques to the low producer (G7 or M4018). To date, there is little literature about this. Only with a good understanding of the role of microevolution in *S. rimosus* genealogy, can we minimize the negative factors and maximize the benefits such as OTC production.

ACKNOWLEDGEMENTS

This work was supported by grants from the Scotland-China Higher Education Research Partnership for PhD Studies, the Fundamental Research Funds for the Central Universities (ECUST), and the Open Project Program of the State Key Laboratory of Bioreactor Engineering, ECUST (No. 2060204).

REFERENCES

- Brekasis D, Paget MS (2003). A novel sensor of NADH/NAD⁺ redox poise in *Streptomyces coelicolor* A3(2). *EMBO J.*, 22: 4856–4865.
- Chopra I, Hawkey PM, Hinton M (1992). Tetracyclines, molecular and clinical aspects. *J. Antibiot. Chemother.*, 29: 245–277.
- Ellen W, Mikael CB, Annika R (2008). Structure and functional properties of the *Bacillus subtilis* transcriptional repressor Rex. *Mol. Microbiol.*, 69(2): 466–478.
- Gyan S, Shiohira Y, Sato I, Takeuchi M, Sato T (2006). Regulatory loop between redox sensing of the NADH/NAD⁺ ratio by Rex (YdiH) and oxidation of NADH by NADH dehydrogenase Ndh in *Bacillus subtilis*. *J. Bacteriol.*, 88: 7062–7071.
- Hendry AP, Kinnison MT (2001). An introduction to microevolution: rate, pattern, process. *Genetica*, 112–113: 1–8.
- Hughes AL (1999). *Adaptive Evolution of Genes and Genomes*. Oxford Univ. Press, Oxford, UK.
- Kerry J, Hiney M, Coyne R, Nicgabhainn S, Gilroy D, Cazabon D, Smith P (1995). Fish feed as a source of oxytetracycline-resistant bacteria in the sediments under fish farms. *Aquaculture*, 131: 101–113.
- Kirby R, Gan TK, Hunter IS (2008). The genome of *Streptomyces rimosus* subsp. *rimosus* shows a novel structure compared to other *Streptomyces* using DNA/DNA microarray analysis. *Antonie van Leeuwenhoek*, 94: 173–186.
- Lawrence J, Hendrickson H (2003). Lateral gene transfer: When will adolescence end? *Mol. Microbiol.*, 50: 739–749.
- Nakamura A, Sosa A, Komori H, Kita A, Miki K (2007). Crystal structure of TTHA1657 (AT-rich DNA-binding protein; p. 25) from *Thermus thermophilus* HB8 at 2.16 Å resolution. *Proteins*, 66: 755–759.
- Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 4(4): 406–425.
- Shen J, Tang ZY, Xiao CY, Guo MJ (2012). Cloning and expression of the redox-sensing transcriptional repressor Rex and *in vitro* DNA-binding assay of the Rex and rex operator in *Streptomyces rimosus* M4018. *Acta Microbiol. Sin.*, 52(1): 30–35.
- Tang ZY, Xiao CY, Guo MJ, Zhuang YP, Chu J, Zhang SL, Herron P, Hunter IS (2011). Improved Oxytetracycline Production in *Streptomyces rimosus* M4018 by Metabolic Engineering of House-keeping Enzyme G6PDH Gene in Pentose Phosphate Pathway *Enzym. Microb. Tech.*, 49: 17–24.
- Wang E, Bauer MC, Rogstam A, Linse S, Logan DT (2008). Structure and functional properties of the *Bacillus subtilis* transcriptional repressor Rex. *Mol. Microbiol.*, 69: 466–478.