



GLOBAL JOURNAL OF SCIENCE FRONTIER RESEARCH  
BIO-TECH & GENETICS  
Volume 12 Issue 4 Version 1.0 Year 2012  
Type : Double Blind Peer Reviewed International Research Journal  
Publisher: Global Journals Inc. (USA)  
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

# Exhaustive Sliding-Window Scan Strategy for Genome-Wide Association Study via Pca-Based Logistic Model

By Qingsong Gao , Zhongshang Yuan , Yungang He , Jinghua Zhao Xiaoshuai Zhang, Fangyu Li, Bingbing Zhang & Fuzhong Xue

*Shandong University, China*

**Abstract** - In genome-wide association study (GWAS), various sliding-window scan approaches have been proposed recently. How to determine the optimal window size, which is influenced by the underlying linkage disequilibrium (LD) patterns, minor allele frequency (MAF) of the causal SNP, and others, is crucial for these methods. However, it is difficult to clarify the theoretical relationship between the optimal window size and these factors. In this regard, we proposed exhaustive strategy with ergodic window sizes along the genome matter whatever the relationship is. Simulations are conducted to assess statistical powers under different sample sizes, relative risks, MAF, LD patterns and window sizes, followed by a real data analysis to evaluate its performance. The simulation results suggested that it was difficult to determine the optimal window size because it was influenced by many factors such as MAF and LD pattern. Real data analysis indicated that the p-values with different window sizes were quite different. Furthermore, with the development of multiprocessor computational technique, the proposed exhaustive strategy combined with the cluster computer technique computationally efficient and feasible for analyzing GWAS data. So the exhaustive strategy is a powerful tool for GWAS data analysis regardless of the relationship between the window size and LD.

**Keywords** : *Minor allele frequency; Causal SNP; Cluster computer..*

**GJSFR-G Classification** : *FOR Code: 060407*



*Strictly as per the compliance and regulations of :*



© 2012 By Qingsong Gao , Zhongshang Yuan , Yungang He , Jinghua Zhao Xiaoshuai Zhang, Fangyu Li, Bingbing Zhang & Fuzhong Xue. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License (<http://creativecommons.org/licenses/by-nc/3.0/>), permitting all non commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Exhaustive Sliding-Window Scan Strategy for Genome-Wide Association Study via Pca-Based Logistic Model

Qingsong Gao <sup>α</sup>, Zhongshang Yuan <sup>ο</sup>, Yungang He <sup>ρ</sup>, Jinghua Zhao <sup>ω</sup>, Xiaoshuai Zhang<sup>\*</sup>, Fangyu Li<sup>§</sup>, Bingbing Zhang <sup>x</sup> & Fuzhong Xue <sup>v</sup>

**Abstract** - In genome-wide association study (GWAS), various sliding-window scan approaches have been proposed recently. How to determine the optimal window size, which is influenced by the underlying linkage disequilibrium (LD) patterns, minor allele frequency (MAF) of the causal SNP, and others, is crucial for these methods. However, it is difficult to clarify the theoretical relationship between the optimal window size and these factors. In this regard, we proposed exhaustive strategy with ergodic window sizes along the genome matter whatever the relationship is. Simulations are conducted to assess statistical powers under different sample sizes, relative risks, MAF, LD patterns and window sizes, followed by a real data analysis to evaluate its performance. The simulation results suggested that it was difficult to determine the optimal window size because it was influenced by many factors such as MAF and LD pattern. Real data analysis indicated that the p-values with different window sizes were quite different. Furthermore, with the development of multiprocessor computational technique, the proposed exhaustive strategy combined with the cluster computer technique computationally efficient and feasible for analyzing GWAS data. So the exhaustive strategy is a powerful tool for GWAS data analysis regardless of the relationship between the window size and LD.

**Keywords** : Minor allele frequency; Causal SNP; Cluster computer.

## I. INTRODUCTION

With rapid improvements in high-throughput genotyping techniques, the cost of genome-wide association study. In recent years, sliding-window methods, in which (GWAS) has been greatly reduced and a boom of large studies of common diseases is underway, which results in an increasing need for new analytical methods to the

association mapping study several neighboring single nucleotide polymorphisms (SNPs) together included in a 'window', have been a popular strategy of automated GWAS data analysis (Sha et al., 2009, Manentiet al., 2009, Yanget al., 2009, Liet al., 2007, Tanget al., 2009, Browning, 2006, Lin et al., 2004). In these sliding-window approaches, the candidate region or the whole genome is divided into many contiguous overlapping windows, followed by multi-locus association tests in each window. Sliding-window approach is commonly used with the fixed window size. For example, Manenti et al used a three-SNP sliding window (Manentiet al., 2009), while Yang et al applied multiple moving window sizes: 3, 5, 7, 9 (Yanget al., 2009). One major conc often encountered in these methods, i.e. how to determine the optimal window size. A large window may include too many non-informative markers while a small window may ignore informative markers, both of which will lead to a reduction in testing power (Yanget al., 2006). The potential problem of small windows is that they do not model the untyped, potentially causal, markers so well. The optimal window size is always influenced by the underlying linkage disequilibrium (LD) patterns, which are certainly variable across a large genomic region or the whole genome (Liet al., 2007, Tanget al., 2009). Variable-sized sliding-window approaches with variable window sizes determined by the underlying LD pattern have been proposed in large-scale data analysis (Browning, 2006, Liet al., 2007, Tanget al., 2009). However, how to clarify the potential theoretical relationship between the window size and LD remains unsolved. Furthermore, most of the variable-sized methods have to go through some computationally intensive phasing program to account for uncertain haplotype phases (Sha et al., 2009). Lin et al proposed that an exhaustive search of all possible windows of SNPs at the genome level is not only computationally practical but also statistically sufficient to detect common or rare genetic-risk alleles (Lin et al., 2004). With the development as well as the extensive applications of multiprocessor and multithreading computational technique, the 'exhaustive' methods have been more feasible. At present study, based on the above concerns, we proposed an exhaustive strategy

*Author α ς § χ* : Department of Epidemiology and Health Statistics, School of Public Health, Shandong University, Jinan 250012, China .

*Author ρ* : CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China.

*Author ω* : Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Chinese Academy of Sciences.

*Author v* : MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK.

*Author x* : postal address: 44 Wenhuxi Road, Jinan, China.  
E-mail : xuefzh@sdu.edu.cn.

with ergodic window sizes along the genome no matter whatever the relationship is. Simulations are conducted to assess statistical powers under different window sizes, followed by a real data analysis to evaluate its performance.

Recently, several sliding-window approaches of GWAS have been developed, including PCA-based methods (Shaet al., 2009, Tanget al., 2009, Wanget al., 2009), P-value combination methods (Sun et al., 2009), haplotype-based methods (Trégouët et al., 2009), and data mining methods (Jianget al., 2009). In particular, PCA-based methods have been proved to have better performance (Shaet al., 2009, Tanget al., 2009, Wang et al., 2009). We, therefore, proposed to apply PCA-based logistic model to perform the exhaustive methods.

## II. METHODS

### a) Exhaustive sliding-window procedure

Consider a case-control study with total  $M$  individuals in a data set and assume each individual has been genotyped at  $N$  SNPs. Let  $G_i = (g_{i1}, g_{i2}, \dots, g_{iN})$  ( $i = 1, 2, \dots, M$ ) denote the  $N$  SNP loci of the  $i^{\text{th}}$  individual, where  $g_{ij}$  denote the genotype of the  $i^{\text{th}}$  individual at  $j^{\text{th}}$  SNP. Let  $y_i$

$$\text{Logit}[\Pr(D = 1 | PC_1, PC_2, \dots, PC_k)] = \beta_0 + \beta_1 PC_1 + \dots + \beta_k PC_k$$

where  $\Pr(D = 1 | PC_1, PC_2, \dots, PC_k)$  denotes the probabilities of disease given the first  $k$  PCs.

## III. SIMULATION

To assess statistical powers with different window sizes and illustrate the necessity of exhaustive sliding-windows, we conducted a statistical simulation based on HapMap data under the null hypothesis ( $H_0$ ) and alternative hypothesis ( $H_1$ ). The corresponding steps for the simulation are as follows:

*Step 1*: Download the phased haplotype data of a genome region from the HapMap web site (<http://snp.cshl.org>): we selected a region near the Protein tyrosine phosphatase, non-receptor type 22 (PTPN22) gene to generate the simulating genotype data of CEU population using HapMap Phase 1& 2 full dataset. This region is located at Chr 1: 114021124..114291292, including 96 SNPs. Figure 1 shows their pairwise  $R^2$  structure.

*Step 2*: Based on the HapMap phased haplotype data, we generated large samples with 100,000 cases and 100,000 controls as CEU populations using the software HAPGEN (Marchini et al., 2007). To investigate the performance of the exhaustive sliding window strategy on different causal SNPs with different minor allele frequencies (MAF), we defined two SNPs as the causal variant respectively: the 45<sup>th</sup> SNP (rs1746853, MAF=0.433) and the 46<sup>th</sup> SNP

denote the trait value of the individual  $i$  (1 for cases and 0 for controls). In the exhaustive sliding-window frame, we first set the largest window size  $L$ . Then, we scan the candidate region or the whole genome from the first SNP with sliding-window of all possible sizes  $s$ , ranging from 2 to  $L$ . (Notice that sliding-window approach with window size 1 is the same as single-locus association test.)

### b) PCA-based logistic regression procedure

Let  $w_s^b$  denote the window with  $s$  neighboring SNPs  $\{b, b+1, \dots, b+s-1\}$  beginning from SNP  $b$ . To carry out the PCA in this region, we let  $\Sigma_s^b$  denote the sample variance-covariance matrix of genotypic numerical codes in window  $w_s^b$  and  $\lambda_j^b$  denote the  $j^{\text{th}}$  largest eigenvalue of  $\Sigma_s^b$ . The cumulative contributing proportion of the total variability explained by the first  $k$  principal components (PCs) is  $C = (\lambda_1^b + \lambda_2^b + \dots + \lambda_k^b) / (\lambda_1^b + \lambda_2^b + \dots + \lambda_s^b)$ . The value of  $k$  can be chosen such that  $C$  exceeds a threshold (80% here). Then we can get our PCA-based logistic model as follows:

(rs2185827, MAF=0.208). To assess the indirect association with disease via correlated markers, we removed the causal SNP in the simulation. The SNPs in the simulated region were coded according to the additive genetic model.

*Step 3*: For the remained SNPs, we set the window size from 2 to 15 for each sample. Single-locus association test was also performed for comparison (set window size as 1). To perform the exhaustive strategy, for each defined window size, all of the windows covering the causal SNP were considered. Correspondingly, single-locus association test was conducted on each of the SNPs involved in these windows. For example, for the causal SNPs rs1746853 (45<sup>th</sup>) with window size 4, windows  $\{42, 43, 44, 46\}$ ,  $\{43, 44, 46, 47\}$ ,  $\{44, 46, 47, 48\}$  were tested and the corresponding single-locus test was performed on  $\{42, 43, 44, 46, 47, 48\}$  respectively.

For the exhaustive strategy, overlapping sliding windows and correlated neighboring SNPs were tested, which might lead to the issue of multiple testing. In present work, we employed simulations under  $H_0$  to construct the null distribution of this strategy, rather than correction methods, to solve the multiple comparison problem. Such simulations have been widely used to

establish significance levels while accounting for multiple tests (Zondervan & Cardon, 2004, Deng et al., 2009). For each set of parameters and a given false-positive error rate ( $\alpha = 0.05$ ), 10000 replications were first generated to construct the null distribution and to determine the critical P value over the simulated region, that is, the smallest P-value of each replication over the simulated region were collected to form the null distribution. Based on the established critical values, we then assessed the power to detect the disease association under different relative risk levels (RR= 1.1, 1.2, 1.3, 1.4 and 1.5 per allele).

To investigate how the optimal window size depended on the underlying LD pattern, the density of genotyped SNPs was used as a surrogate for underlying LD of the region (i.e. the higher LD would be where every SNP is counted as a marker, and lower LD could come from only considering every three SNP).

Specifically, we chose the 45<sup>th</sup> SNP as the causal SNP again, and selected every three SNP of the original data set, i.e. the 3<sup>rd</sup>, 6<sup>th</sup>, ..., 96<sup>th</sup> SNPs. For this particular subset, we conducted the same simulation procedure as above.

*Step 4* : For each window, we sampled the simulation data from the population and performed the PCA-based logistic regression under different sample sizes  $N$  ( $N/2$  cases and  $N/2$  controls,  $N = 1000, 2000, \dots, 5000$ ) using the R package *Design* (<http://cran.rproject.org/web/packages/Design/index.html>).

#### IV. APPLICATION

The proposed method was applied to rheumatoid arthritis (RA) data from GAW16 Problem 1. The data consisted of 2062 Illumina 550k SNP chips from 868 RA patients and 1194 normal controls collected by the North American Rheumatoid Arthritis Consortium (NARAC) (Plenge et al., 2007). At present study, only 1493 females (641 cases and 852 controls) were analyzed to avoid potential bias with the fact that rheumatoid arthritis is two to three times more common in women than in men (Firestein, 2003). We only analyzed chromosome 1 of the data.

Before the sliding-window approach, we excluded data from SNPs that had extensive missingness (missingness > 10%), deviations from Hardy-Weinberg equilibrium (< 0.00001), and low minor allele frequency (< 0.2%) using the software PLINK (Purcell et al., 2007). After this quality control filtering, 38829 SNPs remained. No individuals were excluded for missingness. Then, we applied MACH to impute the missing data (Liet et al., 2009).

## V. RESULTS

### a) Data Simulation

#### i. Critical values under null distribution

Table 1, Table S1 and Table S2 display the critical values for the three cases (different MAF or different LD patterns) based on the given significant level of  $\alpha = 0.05$  over the simulated region. For different sample sizes, the critical values for the same case and the same window size are almost identical. However, for different cases or different window sizes, the critical values are different.

#### ii. Power

Under the case of defining the 45<sup>th</sup> SNP (MAF=0.433) as the causal variant including every SNP in the region, Figure 2 shows the powers with different window sizes under different sample sizes at the given relative risk of 1.3, while Figure 3 shows the powers with different window sizes under different relative risks at the given sample size of 1000. As expected, the powers are monotonically increasing functions of sample sizes and the relative risk levels for each window size. With fixed sample size 2000 and the relative risk 1.3, Figure 4 shows the powers of PCA-based logistic model under different window sizes compared with the corresponding results of single-locus test. Generally, the sliding-window approach is more powerful than the corresponding single-locus test except for window size 11, and the optimal window size is 10.

Under the case of defining the 46<sup>th</sup> SNP (MAF=0.208) as the causal variant including every SNP in the region, Figure 5 shows the powers of PCA-based logistic model under different window sizes compared with the corresponding result of single-locus test. The optimal window size is 3. Nevertheless, the single-locus tests are more powerful for other window sizes.

Under the case of defining the 45<sup>th</sup> SNP (MAF=0.433) as the causal variant including every three SNP in the region, which creates different LD pattern between SNPs by adjusting the density of SNPs, Figure 6 shows the powers of PCA-based logistic model under different window sizes compared with the corresponding results of single-locus test. In this case, the sliding-window approach is less powerful than the corresponding single-locus test except for window size 4.

These simulation results indicate that the powers of both sliding-window approach and single-locus test are influenced by the minor allele frequency of the causal SNP as well as the LD pattern between SNPs, and it is difficult to decide the optimal window size.

### b) Application

Figure 7 shows the exhaustive results to the chromosome 1 of the RA data with window sizes from 1 to 20. The 'win=1' panel denotes the scan results from single-locus association test, while the other panels (from win=2 to win=20) denote the results from PCA-

based logistic model. It is clear that the rs2476601 SNP within PTPN22 gene region was detected at  $10^{-7}$  level ( $p$ -value= $2.30 \times 10^{-8}$ ) by single-locus association test, which has been identified association with RA ( Killberget al., 2007, Begovich et al., 2004, Carlton et al., 2005). However, when the sliding window size was from 2 SNPs to 9 SNPs, no region showed significant at  $10^{-7}$  level, while the same significant region with the rs2476601 SNP involved was re-detected when the sliding window size was from 10 SNPs to 20 SNPs at  $10^{-8}$  level. On the other hand, the  $p$ -values were similar when window sizes ranged from 10 to 20.

## VI. DISCUSSION

As the potential theoretical relationship between the optimal window size and LD is difficult to clarify, and the LD varies across the whole genome, the LD-based sliding-window approaches (Browning, 2006, Li et al., 2007, Tang et al., 2009) may not be always the optimal strategy. We, therefore, proposed an exhaustive strategy with ergodic window sizes along the genome no matter whatever the relationship is. Simulation results suggest that, although the powers are monotonically increasing functions of sample sizes and the relative risk levels for each window size, the powers are also influenced by various factors, including MAF of the causal SNP, LD pattern between the SNPs and window sizes. From Figure 4 and Figure 5, the sliding-window approach seems more powerful than the corresponding single-locus test, but the optimal window sizes are various and heavily depend on the data. Further comparison between Figure 4 and Figure 5 shows that sliding-window approach is generally much more powerful than the corresponding single-locus test when the MAF of the causal SNP is higher. However, this may be attributed to different LD patterns which can be affected by the MAF of causal SNP. To investigate whether the powers are influenced by the LD patterns, we design additional simulations by adjusting the density of SNPs under the same MAF of the causal SNP (0.433). Clearly, the optimal window size changes from 10 to 1 (i.e. single-locus test) when the LD is lower. All the simulation results suggest that it is difficult to determine the optimal window size because it is influenced by so many factors. We, therefore, propose the exhaustive sliding-window strategy to detect various associated genome region with the disease. Real data analysis (Figure 7) results indicate that the  $p$ -values with different window sizes are also quite different. In particular, single-locus association test ( $win=1$ ) identifies a significant SNP (rs2476601) at  $10^{-7}$  level ( $p$ -value= $2.30 \times 10^{-8}$ ), which has been detected as a RA-associated variation by different methods ( Killberget al., 2007, Begovich et al., 2004, Carlton et al., 2005). However, when the sliding window size is from 2 SNPs to 9 SNPs, no region shows significant at  $10^{-7}$  level, while the same significant region with the rs2476601 SNP involved is re-detected when

the sliding window size is from 10 SNPs to 20 SNPs at  $10^{-8}$  level. In practice, it is difficult to capture whole information in genome using the approaches with fixed window size (Yanget al., 2009, Manentiet al., 2009) or with variable window sizes (Liet al., 2007, Tang et al., 2009, Browning, 2006) by the specific algorithms. Our exhaustive strategy does not require a prior knowledge of the optimal window size and genetic factors, such as MAF, LD patterns. On the other hand, both fixed and variable sliding-window approach are just special cases of the exhaustive strategy. Thus, our exhaustive sliding-window strategy is reasonable and essential in GWAS.

Recently, cluster computer, usually known as a multiprocessor based on the chip multithreading architecture, has been widely used for scientific applications. It is very fast because several central processing units (CPUs) inside it can each execute a task's instructions independently of the others. In this article, we used this cluster computer system, and run each sliding-window scan task with one CPU simultaneously. When exhaustively scanning the chromosome 1 of the RA data, it took less than half an hour, so it would only take about 10 hours if scanning the whole genome. Therefore, the proposed exhaustive strategy combined with the cluster computer technique is computationally efficient and feasible for GWAS data analysis.

There are several limitations about the proposed method. First, the proposed exhaustive strategy is still less powerful when the effect of the causal SNP is minor (e.g. relative risk is less than 1.2), and only one causal SNP is considered in present work. Second, the frequencies of both causal SNPs selected are higher than 0.05, so it is hard to decide whether the proposed method is powerful for rare variants. Further work to solve such problems will certainly be warranted.

## VII. ACKNOWLEDGEMENTS

This work was supported by the grant from National Natural Science Foundation of China (30871392). We thank NARAC for supplying us with the data.

## REFERENCES RÉFÉRENCES REFERENCIAS

1. Begovich, A., Carlton, V., Honigberg, L., Schrod, S., Chokkalingam, A., Alexander, H., Ardlie, K., Huang, Q., Smith, A. & Spoerke, J., 2004. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *The American Journal of Human Genetics*. 75, 330-337.
2. Browning, S.R., 2006. Multilocus association mapping using variable-length Markov chains. *American Journal of Human Genetics*. 78, 903-913.
3. Carlton, V., Hu, X., Chokkalingam, A., Schrod, S., Brandon, R., Alexander, H., Chang, M., Catanese, J., Leong, D. & Ardlie, K., 2005. PTPN22 genetic

- variation: evidence for multiple variants associated with rheumatoid arthritis. *The American Journal of Human Genetics*. 77, 567-581.
4. Deng, H.W., Guo, Y.F., Li, J., Bonham, A.J. & Wang, Y.P. , 2009. Gains in power for exhaustive analyses of haplotypes using variable-sized sliding window strategy: a comparison of association-mapping strategies. *Eur J Hum Genet*.17, 785-792.
  5. Firestein, G.S. , 2003. Evolving concepts of rheumatoid arthritis. *Nature*. 423, 356-61.
  6. Jiang, R., Tang, W., Wu, X. & Fu, W. , 2009. A random forest approach to the detection of epistatic interactions in case-control studies. *BMC bioinformatics*. 10, S65.
  7. Killberg, H., Padyukov, L., Plenge, R., R Nnelid, J., Gregersen, P., Van Der Helm-Van Mil, A., Toes, R., Huizinga, T., Klareskog, L. & Alfredsson, L. ,2007. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. *The American Journal of Human Genetics*. 80, 867-875.
  8. Li, Y., Sung, W.K. & Liu, J.J. ,2007. Association mapping via regularized regression analysis of single-nucleotide-polymorphism haplotypes in variable-sized sliding windows. *Am J Hum Genet*. 80, 705-15.
  9. Li, Y., Willer, C., Sanna, S. & Abecasis, G. ,2009. Genotype imputation. *Annual review of genomics and human genetics*. 10, 387.
  10. Lin, S., Chakravarti, A. & Cutler, D. ,2004. Exhaustive allelic transmission disequilibrium tests as a new approach to genome-wide association studies. *Nature genetics*. 36, 1181-1188.
  11. Manenti, G., Galvan, A., Pettinicchio, A., Trincucci, G., Spada, E., Zolin, A., Milani, S., Gonzalez-Neira, A. & Dragani, T.A. ,2009. Mouse genome-wide association mapping needs linkage analysis to avoid false-positive Loci. *PLoS Genet*. 5, e1000331.
  12. Marchini, J., Howie, B., Myers, S., Mcvean, G. & Donnelly, P. ,2007. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 39, 906-13.
  13. Plenge, R., Seielstad, M., Padyukov, L., Lee, A., Remmers, E., Ding, B., Liew, A., Khalili, H., Chandrasekaran, A. & Davies, L. ,2007. TRAF1-C5 as a risk locus for rheumatoid arthritis--a genomewide study. *New England Journal of Medicine*. 357, 1199-1209.
  14. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M., Bender, D., Maller, J., Sklar, P., De Bakker, P. & Daly, M. ,2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*. 81, 559-575.
  15. Sha, Q., Tang, R. & Zhang, S. ,2009. Detecting susceptibility genes for rheumatoid arthritis based on a novel sliding window approach. *BMC Proc*. 3 Suppl 7, S14.
  16. Sun, Y., Jacobsen, D., Turner, S., Boerwinkle, E. & Kardia, S. ,2009. Fast implementation of a scan statistic for identifying chromosomal patterns of genome wide association studies. *Computational statistics & data analysis*. 53, 1794-1801.
  17. Tang, R., Feng, T., Sha, Q. & Zhang, S. ,2009. A variable-sized sliding-window approach for genetic association studies via principal component analysis. *Ann Hum Genet*. 73, 631-7.
  18. Trégouët, D., König, I., Erdmann, J., Munteanu, A., Braund, P., Hall, A., Gro Hennig, A., Linsel-Nitschke, P., Perret, C. & Desuremain, M. ,2009. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nature genetics*. 41, 283-285.
  19. Wang, X., Qin, H. & Sha, Q. ,2009. Incorporating multiple-marker information to detect risk loci for rheumatoid arthritis. *BMC Proc*. 3 Suppl 7, S28.
  20. Yang, H.C., Liang, Y.J., Wu, Y.L., Chung, C.M., Chiang, K.M., Ho, H.Y., Ting, C.T., Lin, T.H., Sheu, S.H., Tsai, W.C., Chen, J.H., Leu, H.B., Yin, W.H., Chiu, T.Y., Chen, C.I., Fann, C.S., Wu, J.Y., Lin, T.N., Lin, S.J., Chen, Y.T., Chen, J.W. & Pan, W.H. ,2009. Genome-wide association study of young-onset hypertension in the Han Chinese population of Taiwan. *PLoS One*. 4, e5459.
  21. Yang, H.C., Lin, C.Y. & Fann, C.S. ,2006. A sliding-window weighted linkage disequilibrium test. *Genet Epidemiol*. 30, 531-45.
  22. Zondervan, K.T. & Cardon, L.R. ,2004. The complex interplay among factors that influence allelic association (vol 5, pg 89, 2004). *Nat Rev Genet*. 5, 238-238.

**Table 1 :** Empirical critical values ( $\alpha=0.05$ ) of the exhaustive sliding-window approach for MAF=0.433 including every SNP under different window sizes and their corresponding single locus test.

Window size	Sample size				
	1000	2000	3000	4000	5000
win=2	0.053(0.042)	0.049(0.040)	0.051(0.041)	0.048(0.037)	0.047(0.038)
win=3	0.044(0.018)	0.039(0.015)	0.040(0.016)	0.038(0.016)	0.038(0.015)
win=4	0.039(0.016)	0.036(0.014)	0.039(0.014)	0.038(0.014)	0.037(0.014)
win=5	0.034(0.013)	0.032(0.011)	0.032(0.011)	0.032(0.011)	0.028(0.010)
win=6	0.032(0.013)	0.030(0.011)	0.030(0.011)	0.031(0.011)	0.027(0.011)
win=7	0.031(0.012)	0.028(0.011)	0.028(0.011)	0.028(0.011)	0.024(0.009)
win=8	0.029(0.012)	0.026(0.010)	0.027(0.010)	0.027(0.010)	0.026(0.009)
win=9	0.027 (0.012)	0.025(0.010)	0.024(0.010)	0.024(0.010)	0.025(0.010)
win=10	0.026(0.012)	0.023(0.010)	0.022(0.010)	0.023(0.010)	0.022(0.009)
win=11	0.033(0.011)	0.029(0.009)	0.028(0.009)	0.028(0.009)	0.028(0.010)
win=12	0.033(0.009)	0.029(0.008)	0.029(0.008)	0.029(0.008)	0.028 (0.007)
win=13	0.032(0.009)	0.028(0.008)	0.027(0.007)	0.027(0.007)	0.027(0.007)
win=14	0.031(0.008)	0.027 (0.007)	0.026(0.007)	0.026 (0.007)	0.027(0.007)
win=15	0.033(0.008)	0.028(0.007)	0.029(0.007)	0.028(0.007)	0.028(0.007)

*Note: the number in ( ) denotes the results of the corresponding single locus test*

**Table 2 :** Empirical critical values ( $\alpha=0.05$ ) of the exhaustive sliding-window approach for MAF=0.208 including every SNP under different window sizes and their corresponding single locus test.

Window size	Sample size				
	1000	2000	3000	4000	5000
win=2	0.052(0.025)	0.049(0.025)	0.056(0.025)	0.057(0.025)	0.045(0.023)
win=3	0.030(0.014)	0.032(0.014)	0.025(0.012)	0.022(0.012)	0.027(0.015)
win=4	0.036(0.012)	0.035(0.011)	0.035(0.010)	0.032(0.009)	0.034(0.011)
win=5	0.035(0.012)	0.032(0.011)	0.031(0.010)	0.031(0.009)	0.031(0.011)
win=6	0.028(0.012)	0.025(0.011)	0.026(0.010)	0.023(0.009)	0.027(0.010)
win=7	0.030(0.011)	0.027(0.010)	0.027(0.009)	0.025(0.008)	0.026(0.009)
win=8	0.029(0.011)	0.026(0.009)	0.026(0.009)	0.023(0.007)	0.025(0.009)
win=9	0.027 (0.011)	0.025(0.009)	0.025(0.009)	0.021(0.007)	0.022(0.009)
win=10	0.028(0.011)	0.026(0.009)	0.026(0.009)	0.023(0.007)	0.026(0.008)
win=11	0.032(0.009)	0.030(0.008)	0.028(0.007)	0.025(0.006)	0.026(0.006)
win=12	0.032(0.008)	0.029(0.007)	0.028(0.006)	0.025(0.005)	0.024 (0.005)
win=13	0.031(0.007)	0.029(0.006)	0.027(0.006)	0.025(0.005)	0.025(0.005)
win=14	0.027(0.007)	0.026 (0.006)	0.024(0.006)	0.021 (0.005)	0.022(0.005)
win=15	0.031(0.007)	0.029(0.006)	0.026(0.006)	0.022(0.005)	0.022(0.005)

*Note: the number in ( ) denotes the results of the corresponding single-locus test.*

**Table 3 :** Empirical critical values ( $\alpha=0.05$ ) of the exhaustive sliding-window approach for MAF=0.433 including every three SNP under different window sizes and their corresponding single locus test.

Window size	1000	2000	3000	4000	5000
win=2	0.041(0.041)	0.035(0.035)	0.027(0.027)	0.023(0.023)	0.021(0.021)
win=3	0.032(0.020)	0.026(0.015)	0.020(0.012)	0.018(0.010)	0.014(0.008)
win=4	0.020(0.009)	0.015(0.006)	0.009(0.004)	0.007(0.003)	0.006(0.002)
win=5	0.024(0.008)	0.017(0.006)	0.011(0.004)	0.009(0.002)	0.007(0.002)
win=6	0.021(0.007)	0.015(0.005)	0.010(0.003)	0.007(0.002)	0.006(0.002)
win=7	0.021(0.007)	0.015(0.005)	0.009(0.003)	0.007(0.002)	0.005(0.002)
win=8	0.019(0.006)	0.013(0.004)	0.008(0.002)	0.006(0.002)	0.004(0.002)
win=9	0.022 (0.005)	0.015(0.003)	0.009(0.002)	0.007(0.002)	0.005(0.001)
win=10	0.019(0.005)	0.013(0.003)	0.008(0.002)	0.006(0.001)	0.005(0.001)
win=11	0.019(0.004)	0.012(0.003)	0.008(0.002)	0.006(0.001)	0.005(0.001)
win=12	0.018(0.004)	0.012(0.003)	0.008(0.002)	0.006(0.001)	0.005 (0.001)
win=13	0.017(0.004)	0.012(0.003)	0.008(0.002)	0.005(0.001)	0.004(0.001)
win=14	0.017(0.004)	0.012 (0.002)	0.008(0.001)	0.006 (0.001)	0.004(0.001)
win=15	0.017(0.003)	0.012(0.002)	0.008(0.001)	0.006(0.001)	0.004(0.001)

Note: the number in ( ) denotes the results of the corresponding single-locus test.

### FIGURE LEGENDS

Figure 1. Pairwise R<sup>2</sup> structure for selected region

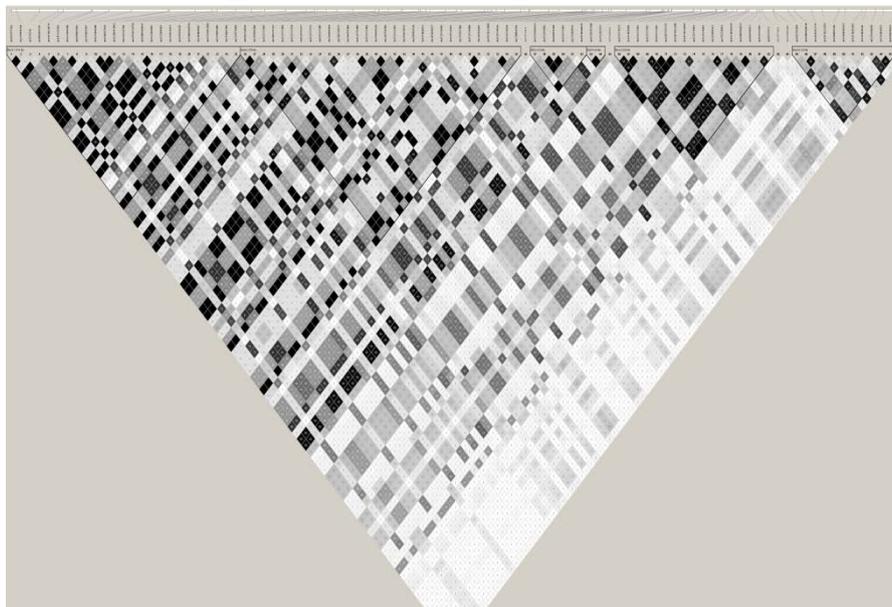
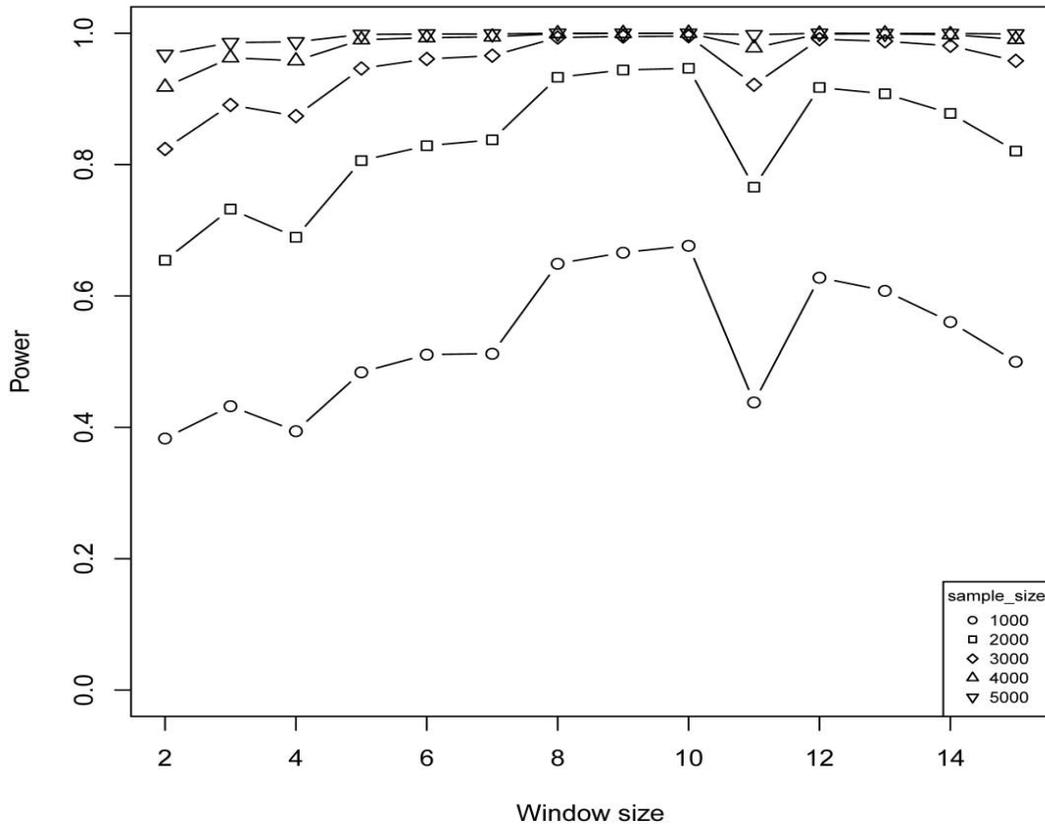


Figure 1 : Pairwise R<sup>2</sup> among the SNPs in the selected region.

**Figure 2. Power for MAF=0.433 and RR=1.3**



*Figure 2 :* The powers of the exhaustive sliding-window approach under different window sizes (2-15) and different sample sizes. The horizontal axis denotes the window sizes and the vertical axis denotes the powers of PCA-based logistic regression model with different window sizes.



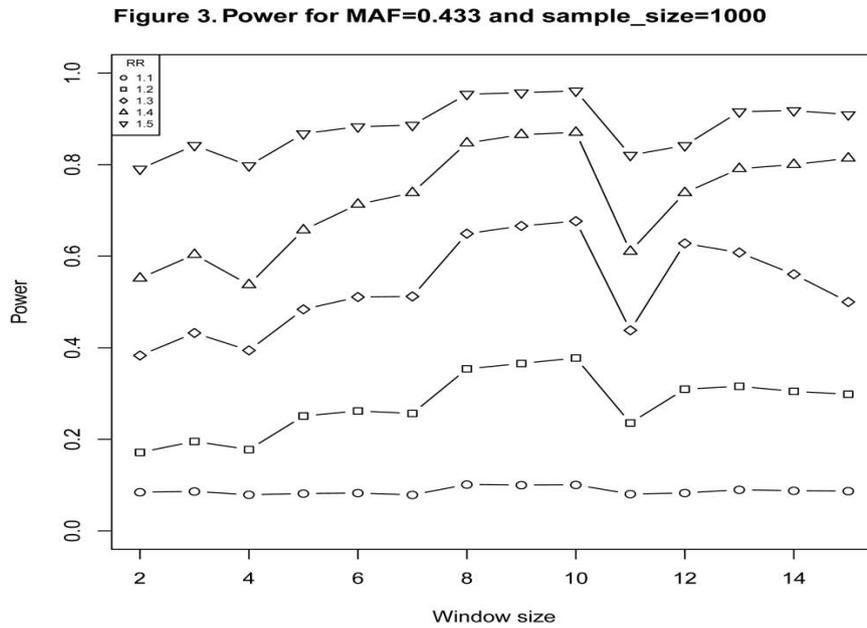


Figure 3 : The powers of the exhaustive sliding-window approach under different window sizes (2-15) and different relative risks. See Figure 2 for the figure legends.

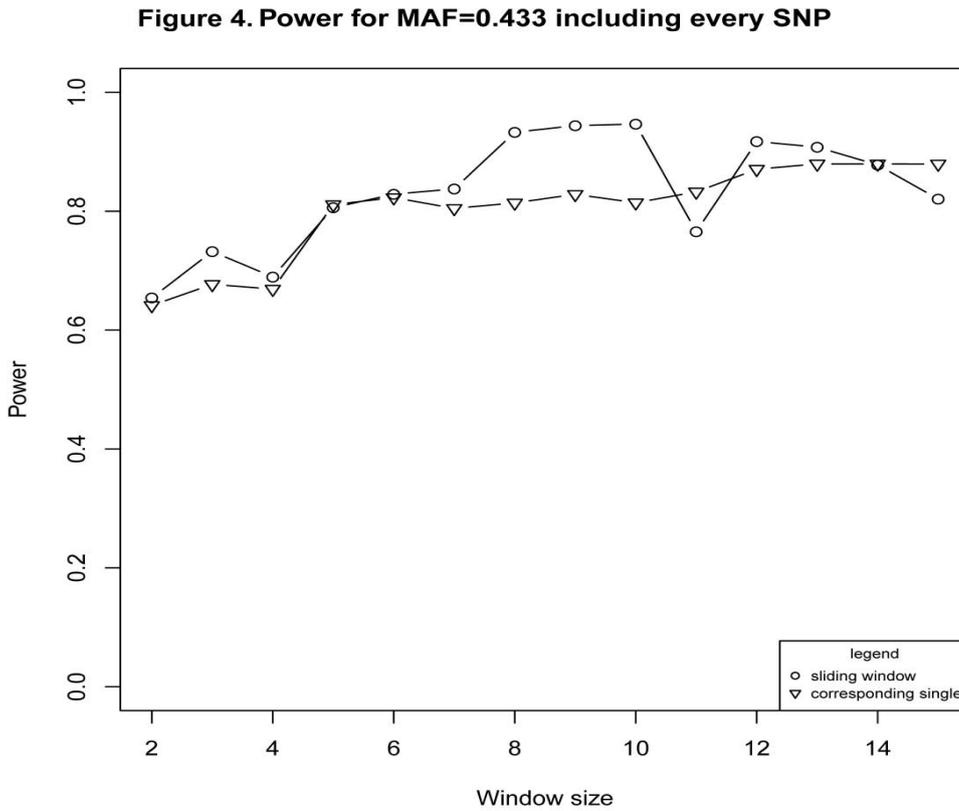
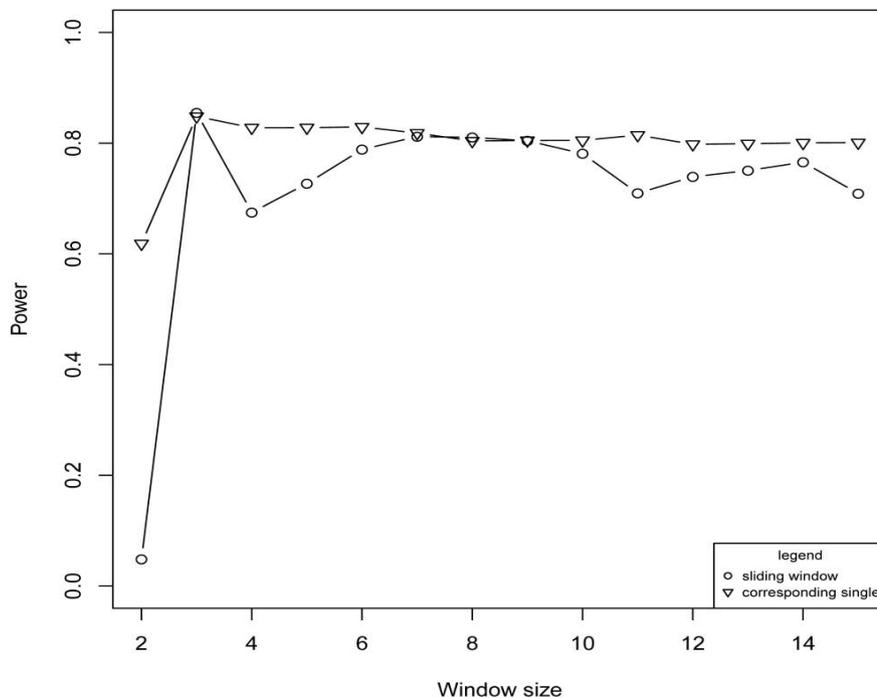


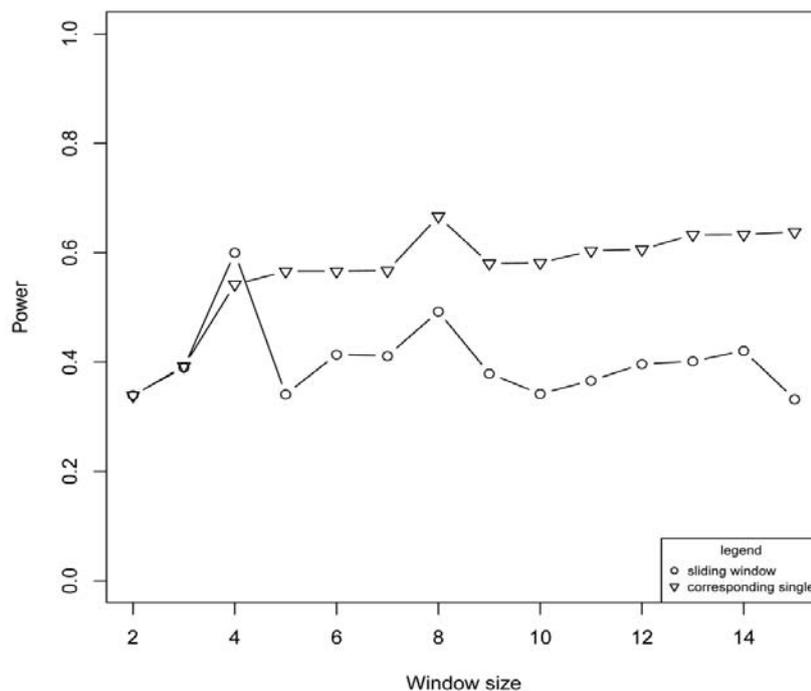
Figure 4 : The powers of the exhaustive sliding-window approach for MAF=0.433 including every SNP under different window sizes and their corresponding single locus test. See Figure 2 for the figure legends.

**Figure 5. Power for MAF=0.208 including every SNP**



*Figure 5 :* The powers of the exhaustive sliding-window approach for MAF=0.208 including every SNP under different window sizes and their corresponding single locus test. See Figure 2 for the figure legends.

**Figure 6. Power for MAF=0.433 including every three SNP**



*Figure 6 :* The powers of the exhaustive sliding-window approach for MAF=0.433 including every three SNP under different window sizes and their corresponding single locus test. See Figure 2 for the figure legends.

Figure 7. Manhattan plot

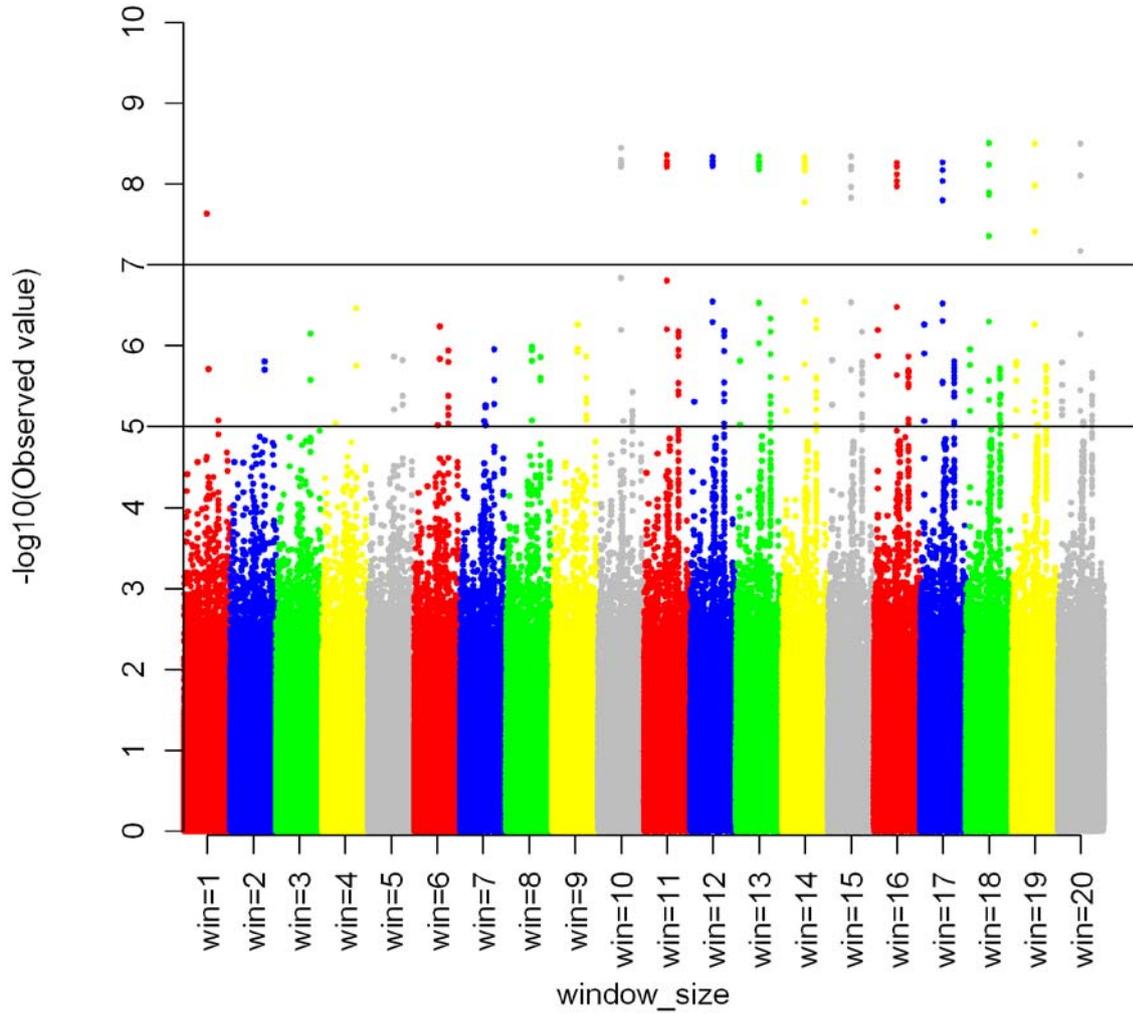


Figure 7: The results for chromosome 1 of the RA data using single-locus association test (the 'win=1' panel) and exhaustive sliding-windows (from panel 'win=2' to panel 'win=20'.)

This page is intentionally left blank



GLOBAL JOURNAL OF SCIENCE FRONTIER RESEARCH  
BIO-TECH & GENETICS  
Volume 12 Issue 4 Version 1.0 Year 2012  
Type : Double Blind Peer Reviewed International Research Journal  
Publisher: Global Journals Inc. (USA)  
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

# Bioleaching of Copper Concentrate and Pyrite by Using Native Bacterium Acidithiobacillus Ferrooxidans IRL.8F and Evaluating the LPS Role in Bioleaching Process

By Dr. Ali Mohammad Latifi , Ahmadi . M & Olad .G

*Baqiyatallah Medical Sciences University*

**Abstract** - This study was performed to evaluate the ability of native bacterium to extract copper and iron from their ores. This bacterium was isolated from iron mineral springs in Iran's Larzan region and was named Acidithiobacillus ferrooxidans IRL.8F based on morphological and physiological characteristics and 16S rRNA molecular analyses. The results from bioleaching of copper concentrate showed that the amount of extracted copper and iron was 71.4% and 29.3%, respectively. Furthermore, in comparison with control samples, these amounts increased by 93.5% and 92%, respectively. In the control samples minor amount of metals were extracted due to spontaneous leaching. To assess the importance of bacterial lipopolysaccharides(LPS) role, LPS of bacterium was removed. When ethylenediaminetetraacetic acid (EDTA) in concentrations of 5 and 10% was used during the bioleaching process of pyrite, process efficiency decreased to 61 and 70%, respectively. The cells lacking LPS were led to 59.4 % decrease in the amount of bacterial leaching, in contrast to whole cells. Therefore, it can be concluded that: 1. EDTA causes a drastic reduction in the efficiency of leaching process, 2. Bacterial LPS have a key role in attachment to particles of ore and 3.This bacterium is capable of leaching metals through the direct mechanism.

**Keywords** : Bioleaching, LPS, Ore, Copper, Pyrite.

**GJSFR-G Classification** : FOR Code: 060501



*Strictly as per the compliance and regulations of :*



RESEARCH | DIVERSITY | ETHICS

# Bioleaching of Copper Concentrate and Pyrite by Using Native Bacterium *Acidithiobacillus Ferrooxidans* IRL.8F and Evaluating the LPS Role in Bioleaching Process

Dr. Ali Mohammad Latifi<sup>α</sup>, Ahmadi . M<sup>σ</sup> & Olad .G<sup>ρ</sup>

**Abstract** - This study was performed to evaluate the ability of native bacterium to extract copper and iron from their ores. This bacterium was isolated from iron mineral springs in Iran's Larzan region and was named *Acidithiobacillus ferrooxidans* IRL.8F based on morphological and physiological characteristics and 16S rRNA molecular analyses. The results from bioleaching of copper concentrate showed that the amount of extracted copper and iron was 71.4% and 29.3%, respectively. Furthermore, in comparison with control samples, these amounts increased by 93.5% and 92%, respectively. In the control samples minor amount of metals were extracted due to spontaneous leaching. To assess the importance of bacterial lipopolysaccharides(LPS) role, LPS of bacterium was removed. When ethylenediaminetetraacetic acid (EDTA) in concentrations of 5 and 10% was used during the bioleaching process of pyrite, process efficiency decreased to 61 and 70%, respectively. The cells lacking LPS were led to 59.4 % decrease in the amount of bacterial leaching, in contrast to whole cells. Therefore, it can be concluded that: 1. EDTA causes a drastic reduction in the efficiency of leaching process, 2. Bacterial LPS have a key role in attachment to particles of ore and 3.This bacterium is capable of leaching metals through the direct mechanism.

**Keywords** : Bioleaching, LPS, Ore, Copper, Pyrite.

## I. INTRODUCTION

Bioleaching is a general term used to refer to the conversion of insoluble to soluble metals (usually in sulfated form) through biological oxidation by using microorganisms (Rawlings., 2002; Makita et al., 2004).

Bacteria of the genus *Thiobacillus*, like *Thiobacillus ferrooxidans* retrieve the energy from ores via enzymatic oxidation. Biological oxidation of sulfide ores and the electron transport occur in three forms, including direct (or enzymatic or contact), indirect (mediated by compounds such as Fe<sup>3+</sup> ions) and cooperative (which includes both direct and indirect ) mechanisms. In the indirect mechanism, Fe<sup>3+</sup> iron plays major role, while in the direct mechanism, the

bacterium should have access to the ore, bind it and then the reaction will occur at the ore-water interface (Donati and Sand, 2007; Wolfgang and Edgardo, 2007). In this mechanism, the microbial attachment to the ore surface is necessary for the bioleaching process. As the micro-organism approaches.

The mineral, the cell surface changes and this accompanies with expression of extracellular polymeric substances (EPS) which lead to the attachment (Scobar et al., 1997; Clausen, 2003). EPS forms chemical bonds with the surface and mediates or promotes respiration and nutritional chemical reactions (Scobar et al., 1997). These bonds are made stronger with the attachment of microorganism to the ore and the reactions are followed by oxidation of reduced mineral compounds or reduced Fe<sup>2+</sup> or sulfate ions. Attaching to the ores may be mediated by forming EPS on the surface of solid particles such as lipopolysaccharide (LPS), phospholipids or other macromolecules like the polypeptides in the outer membrane of the bacterium. These compounds are released by the organism when it is in contact with the ores (Donati and Sand., 2007; Scobar et al., 1997). The Mechanism of the electron transport from pyrite to molecular oxygen has been identified in detail. The primary stages occur in the EPS, in which electrons are extracted by means of the Fe<sup>3+</sup> ion in complex with glucuronic acid (Rangin and Basu., 2004). Attachment to hydrophobic substrates such as sulfur is mediated by van der Waals forces, while for binding to charged substrates like pyrite, cations or molecules which act as Lewis acids accept the uncharged electron pair of the pyrite sulfur, followed by formation of a complex between different iron species and the exopolysaccharide and finally the attachment of the bacterium to the substrate (Gehrke et al., 1998)

## II. MATERIALS AND METHODS

### a) Media

Types and compositions of the media used for culturing, isolating and screening included: (1)SF or T.F.

**Author α σ ρ** : Baqiyatallah Medical Sciences University  
E-mail : amlatify@yahoo.com

medium containing: Solution A:  $K_2HPO_4$  (0.5g/l),  $(NH_4)_2SO_4$  (0.5g/l),  $MgSO_4$  (0.5g/l),  $H_2SO_4$  0.5M (5 ml/l) and D.W (1000ml); Solution B:  $FeSO_4 \cdot 7H_2O$  (167g/l),  $H_2SO_4$  . 0.5M (50 ml) and D.W (1000ml). One unit volume of solution B is mixed with four unit volumes of solution A and the pH is adjusted on 2-2.5 by using  $H_2SO_4$  . 0.5M. (2) TSB medium containing:  $KH_2PO_4$  (3g/l),  $(NH_4)_2SO_4 \cdot 7H_2O$  (0.4 g/l),  $MgSO_4 \cdot 7H_2O$  (0.5 g/l),  $CaCl_2 \cdot 2H_2O$  (0.25 g/l),  $FeSO_4 \cdot 7H_2O$  (0.01g/l),  $Na_2S_2O_3 \cdot 5H_2O$  (5 g/l), Agar powder (16g/l) and D.W (1000 ml). (3) TTB medium which contains (g/l) : (  $(NH_4)_2SO_4$  0.3;  $K_2HPO_4$  , 0.5;  $MgSO_4 \cdot 7H_2O$ , 0.5; and 0.5M  $H_2SO_4$  . After autoclaving, sterilizing and cooling the medium, 5% sulfur powder separately sterilized in an aluminum foil, was added (Chen and Lin, 2000; Sasaki et al., 2009). During the preparation of these media, the iron sulfate was sterilized with a (0.22 $\mu$ ) filter and added to the solution. The cells were collected from 10-day media centrifuged in 50 ml falcon tubes at 15000 rpm for 20 min. (Elzekey an

### III. THE BACTERIUM

The bacterium used in this study was isolated from mineral springs in Larzan, Qazvin province, Iran. With this purpose, the mixed samples of water and precipitations deposited at the bottom of the spring were collected, transferred to the laboratory and incubated into 250 ml Erlenmeyer flasks containing 50 ml of broth TF and TT media (respectively containing elemental iron and sulfur as the sole sources of electron and energy). The samples were placed in shaker incubator at 30°C and in 200 rpm for 7 days, and then recultured in fresh media. The oxidation power of Fe and S elements were evaluated. During the cell culture period, essential parameters including the daily measure of pH, titration of produced acid, solution rate of elemental sulfur in medium, the amount of oxidized Fe, macroscopic and microscopic study of samples and counting and calculating the cell concentration were also considered or evaluated.

### IV. OXIDATION OF $Fe^{2+}$ TO $Fe^{3+}$

The oxidation of  $Fe^{2+}$  by the bacterium was investigated in a 250 ml Erlenmeyer flask containing 50 ml SFB medium. With this purpose, 5 ml of 14-day bacterium culture (comprising  $\sim 9 \times 10^8$  cells) was inoculated into the medium and incubated at 30°C at 200 rpm. The control was without bacterium inoculation. The initial pH was adjusted to 2.5 using 0.5M sulfuric acid. As  $Fe^{2+}$  is oxidized to  $Fe^{3+}$  iron, the medium turns from lime green to yellow, brown and brick red. Orthophenanthroline method and atomic adsorption spectroscopy analysis systems were used to analyze the iron. The total iron content ( $Fe^{2+} + Fe^{3+}$ ), the converted  $Fe^{2+}$  iron to  $Fe^{3+}$  and the  $Fe^{3+}$  content of the medium were measured.

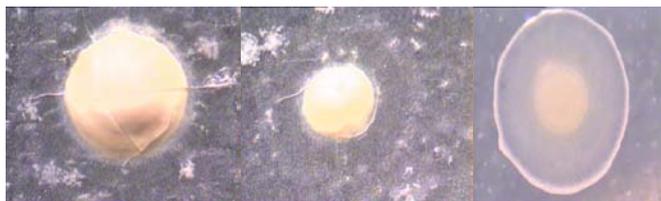
### V. SULFUR OXIDATION AND SULFURIC ACID PRODUCTION

The medium in this study was TTB. The inoculation and growth conditions were similar to those of iron oxidation. The initial pH was adjusted to 4.5. Sulfur oxidation, pH reduction and sulfuric acid production were measured. The control was prepared in a similar way without bacterium inoculation. The sulfuric acid content of the medium was measured after the drastic decrease of pH by titration using 0.1 M NaOH.

### VI. BACTERIA IDENTIFICATION BY 16S RRNA

To identify the bacteria, the sequencing of 16S rRNA gene fragments was applied. Considering that these bacteria have a slow growth rate and are extremophilic species, the alkaline lysis and lysosyme methods were used in hybrid to extract their genomes. At first, the bacterial genomes were purified. With this purpose, 500 ml of 7-day culture of bacteria was prepared in TTB medium. The sample was centrifuged at 15000 rpm and the bacterial biomass was obtained. One hundred microliter of the SET cold buffer was added to the bacteria and 100 $\mu$ l of lysosyme was added to the above mixture and it was vortexed thoroughly. The mixture was incubated at 37°C for 30 min, and then, 200 $\mu$ l of lysis buffer (NaOH (5M), SDS (10%),  $H_2O$ ) was added to the mixture and placed in ice for 10 min. In the next stage, as much phenol as the volume of the solution in the tube was added. The mixture was blended thoroughly and centrifuged at 10000 rpm and 4°C for 3 min. Furthermore, the supernatant was transferred to another tube and as much as its volume, phenol-chloroform (1:1) was added. The mixture was centrifuged again in a similar way as mentioned above. The supernatant was transferred to another tube, chloroform was added as much as its volume and the sample was again centrifuged as above. The supernatant was removed and isopropanol stored at -20°C was added as much as 0.6 of volume of the supernatant. This solution was stored at -20°C for 1 h. In the next stage, the sample was centrifuged at 14000 rpm in 4°C for 15 min. Isopropanol was immediately removed and 1 ml of 70% alcohol was added and then the sample was centrifuged at 14000 rpm at 4°C for 10 min. The sample was drought at room temperature and 20 to 30 $\mu$ l of TE buffer or distilled water and 3 to 5 $\mu$ l of RNase A was added. The tube containing the sample was stored at 37°C for 1 h and the sample was then stored at 4°C (Ohba and Owa., 2005). After electrophoresis, the PCR was performed. The primers required to identify the bacteria were universal primers with the following sequences (Hong et al., 2006 ; Yeats et al 1998): Forward: FORB: 5' AGAGTTTGATCCTGGCTCAG3'

*reverse* : REVB: 5' GGTTACCTTGTTACGACT3'. Using the purified genome as the template and the *Taq* polymerase, the 16S rRNA gene fragment was amplified as the defined program and the final product was investigated on agarose 1% gel. After confirming the quality of the PCR product, the samples were sequenced using Genetic analyzer 31030 - Accessories Applied Biosystem.



**Figure 1** : A. ferrooxidans IRL.8F micro-colony and production of exopolysaccharide (white corona) on TSA medium (magnified  $\times 100$  – by A.M. Latifi). From right to left: fresh colony, semi- dried colony, dried colony.

The ore

The pyrite and copper ore were used in this study. The ore was powdered using the crusher or mortar-and-pestle, and the samples were prepared using special sieves with appropriate gridding. The elemental and compositional analysis of the ore powder

## VII . BIOLEACHING MEDIA AND CONDITIONS

To measure the metal produced from ore, 10 g of ore powder in flasks containing 250ml medium without any of energy resources (iron, sulfur, etc) was used. The base medium for all samples was 100ml water. The gridding of the particles was 200 and their

## VIII . STUDY OF THE EFFECT OF EDTA ON BIOLEACHING PROCESS

To study the effect of EDTA on bioleaching process, four samples were prepared as followed: Sample 1: leaching medium without bacterium inoculation (as control), Sample 2: leaching medium inoculated with bacterium without EDTA, Sample 3: leaching medium inoculated with bacterium and EDTA (5%) and Sample 4: leaching medium inoculated with bacterium and EDTA (10%) test samples, the 0.2M EDTA solution was used.

All samples were placed in a shaker-incubator at 28°C and 200 rpm. Bacterium compatibility with new media was investigated by measuring pH of media during the bioleaching process. The initial pH at the start time of the process was also recorded. During its growth period, the bacterium reduces the pH and produces sulfuric acid through oxidation of the sulfur ore.

## IX . LPS REMOVAL IN ACIDITHIOBACILLUS FERROOXIDANS IRL.8F

To remove the bacterial LPS, 0.2M EDTA and Tris-HCl at pH4.5 were used. (Ramadas et al., 1991; Scoabar et al.,1997). Bacterial biomass was collected from 50 ml of bacterial suspension comprising a 10-day culture. The biomass was converted to a homogenous suspension in the EDTA and Tris-HCl solution and incubated at 37°C for 1 h. The tube containing the sample was then centrifuged at 11000 rpm and the supernatant containing EDTA, Tris-HCl and lipopolysaccharide (LPS) was removed. Bacterial cells lacking LPS were extracted from the solution containing EDTA and LPS by centrifugation at 11000 rpm for 10 min, and were inoculated with the bioleaching medium previously prepared .

## X. ASSESSING THE ACTIVITY OF LPS-LACKING BACTERIA IN OXIDATION OF $Fe^{2+}$

Two samples containing TF medium with  $Fe^{2+}$  as the source of energy, were inoculated as follows: (1) The control in which the normal *A. ferrooxidans* IRL.8F was inoculated into the medium without any treatment, and (2) The test sample with LPS-lacking *A. ferrooxidans* IRL.8F inoculated into the medium.

To investigate the restoring of LPS production ability of the bacteria, they were collected from 10-day culture of the second sample and inoculated into the fresh TF medium.

## XI. ASSESSING THE ACTIVITY OF LPS-LACKING BACTERIA IN BIOLEACHING OF THE IRON FROM PYRITE SOIL

To assess such an activity, the pyrite ore with mesh of 200 and mesh size of 0.074mm was used and 10g/l of the ore was added to each of Erlenmeyer flasks. The base medium of all samples was water. pH of all samples was adjusted on 4.5. Prepared samples included: leaching medium inoculated with normal bacteria (having LPS), leaching medium inoculated with LPS-lacking bacteria and leaching medium without any bacterium.

The samples were placed in a shaker-incubator at 200 rpm, and 25°C for 14 days, and after precipitation the supernatant was used to analyze the amount of

## XII . RESULTS

During screening stage, we could isolate a bacterial strain with remarkable enzymatic ability to oxidize the iron and sulfur as its sole energy and electron source. It is noteworthy that the mineral spring



from where the bacterium was isolated have a brick-red solution and fawn deposits. This results from the natural activity of the bacterium in oxidation of  $Fe^{2+}$  iron in the nearby soils to  $Fe^{3+}$ , leading to color change and generation of jarosite (iron hydroxide) deposits. The isolated bacterium produces small colonies similar to fried egg in TSA agar medium which are hardly visible with naked eyes. Applying an innovative method using optical microscope and simultaneous lighting from up and down in this study, we could produce high-quality pictures of bacterial colonies (Fig. 1).

Results obtained from morphological, physiologic and molecular identification based on 16S rRNA revealed that this bacterium is mostly similar to *A. ferrooxidans* strain. Therefore, the bacterial strain was named *A. ferrooxidans* IRL.8F. It is chemautotroph and uses  $CO_2$  in the air as its carbon source. Fig 2 shows what was obtained from extraction of bacterial genome.

### XIII . FE OXIDATION

The medium was observed to turn from lime green to brick red (confirming the conversion of  $Fe^{2+}$  to  $Fe^{3+}$ ) (Fig. 3). After 18 h, the medium turns to yellow as a result of bacterial activity and the brick red color observed within 48 to 72 h represents complete oxidation of  $Fe^{2+}$  to  $Fe^{3+}$ . Results of cell counting showed that bacterial cell concentration has begun to increase when the color changes started and it increased from  $9 \times 10^8$  cells per ml in the first 24-h period to  $18 \times 10^8$  in the second day. Within this period, pH of the medium decreased from 2.5 to 2. Deposits in fawn color were observed on the wall of flask, which increased in amount daily. . They can trap the leached metals in the solution in their lattices and thereby disturb the bioleaching process. Meanwhile, this problem can be overcome by retaining the low pH.

This experiment demonstrates the bacterial capability to leaching the minerals containing iron compounds.

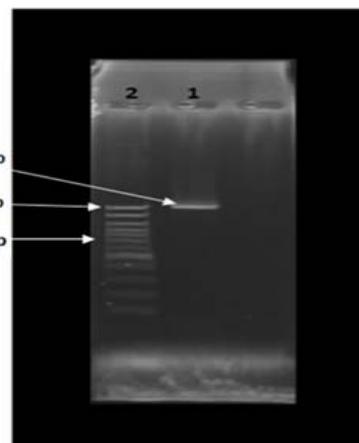


Figure 2: Electrophoresis of the PCR product of IRL.8F bacterium on 1% agarose gel. 1. PCR product, 2. DNA Ladder.



Figure 3: Oxidation of  $Fe^{2+}$  to  $Fe^{3+}$  in SF broth medium and iron oxide particles formed on the colonies surface in the solid TSA medium (magnified  $\times 100$  – by A.M. Latifi).

### XIV . OXIDATION OF SULFUR AND PRODUCTION OF SULFURIC ACID

Results obtained from this experiment (Figure 4) revealed the high capability of the bacterium to produce acid, reduce the pH and make strong acidic conditions in TTB medium, such that in the third day the pH reached 1.5, in tenth day it decreased below 1 and in the 18<sup>th</sup> day it was 0.75. As the pH decreases, the number of bacterial cells progressively increased, such that it doubled (to  $6.13 \times 10^8$ ) with pH decrease from 4 to 1.6 and it triples with pH decrease from 1.6 to 0.9. For the fact that the most populated cell colony is observed in 14<sup>th</sup> day, we used the 14-day suspension to produce the bacterial seed. The maximum acid production rate in 25<sup>th</sup> day is 20 g/l. The pH changes and the sulfuric acid production are shown in the plot(Figure 4) .

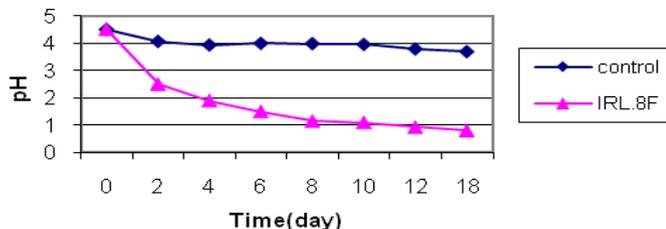


Figure 4: Decrease of pH by the bacterium in TT broth medium.

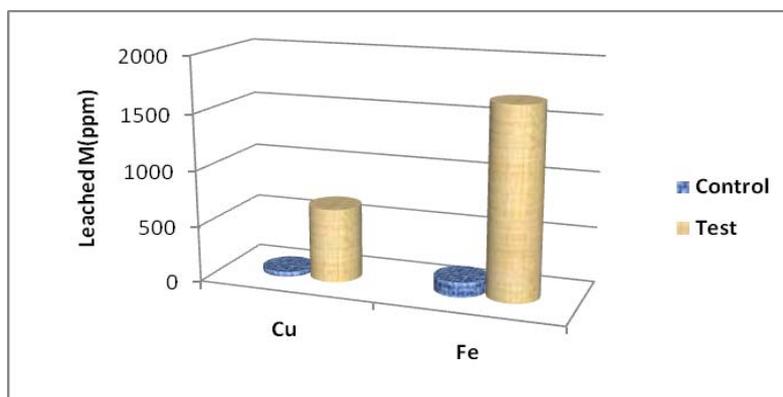


Figure 5: Bioleaching of Cu and Fe from Copper concentrate.

## XV. BIOLEACHING OF IRON AND COPPER FROM ITS CONCENTRATE

X-ray fluorescence (XRF) and X-ray diffraction (XRD) analyses of pyrite ore showed that it contains 23.91% iron, 23.11% copper, 0.052% manganese, 0.001% nickel, 5.55% gold, 0.014% molybdenum, 22.79% sulfur.

Bioleaching of this ore for copper and iron elements showed that this mine is a highly appropriate medium for growth and activity of this bacterium. The bacterial cells consume and oxidize the sulfur element available in the mineral soil to sulfuric acid and drastically decrease the pH to as low as 1.98; thereby they provide.

Appropriate conditions for extraction of insoluble metals in the mineral soil. Analysis of the leaching solution showed 1690 and 663ppm rates of iron and copper extraction, which equal 29.3 and 71.4%, respectively. The values showed 92 and 93.5% increase in comparison with control samples, respectively (Fig.5).

The small amounts of extracted metals in control samples have resulted from spontaneous leaching. Note that in comparison with control samples (without inoculated bacteria), bioleaching medium of copper concentrate came in green and with development process the intensity of color was increased (Fig.6).



Figure 6: Color change due to bacterial activity and extraction of Cu in medium.

### Bioleaching of pyrite ore

XRF and XRD analyses of pyrite ore showed that it contains 23.91% iron and 23.11% copper. The mineralogical analysis showed its composition as  $\text{CaCO}_3$  (Calcite),  $\text{FeS}_2$  (Pyrite),  $\text{ZnS}$  (Sphalerite) and  $\text{CaMg}(\text{CO}_3)_2$  (dolomite). Bioleaching of this type of ore demonstrated the decrease of pH as a consequence of sulfur consumption, resulting in the efficient metal extraction. The amount of metal extracted from the test sample showed a 60 to 70% increase in comparison with control sample (without bacterium).

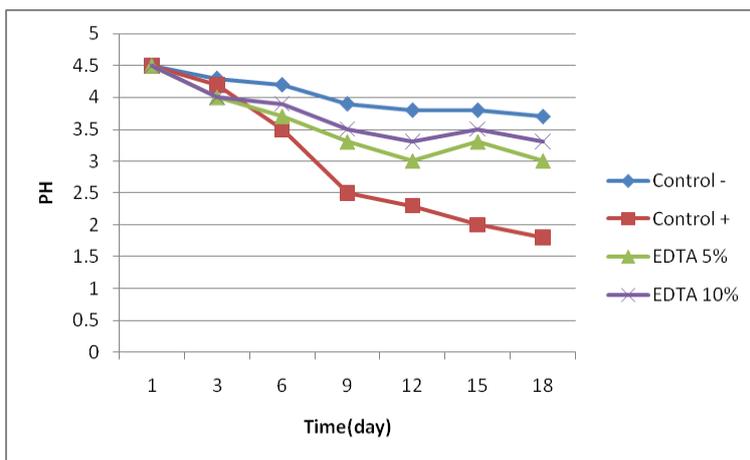


Figure 7: pH changes in the investigation of EDTA effect on bacterial activity of *A. ferrooxidans* IRL.8F and rate of  $Fe^{3+}$  extraction from pyrite ore.

### XVI . STUDY OF EDTA EFFECT ON LPS IN THE OUTER MEMBRANE OF *A. FERROOXIDANS* IRL.8F AND ON RATE OF $Fe^{3+}$ EXTRACTION FROM THE PYRITE ORE

To investigate the effect of EDTA on bioleaching through damaging the bacterial cell membrane, 5 and 10% concentrations of EDTA were simultaneously added to the leaching medium (the

base medium of water). Decrease of pH in the sample without EDTA was observed to follow a slower slope and it directly depends on rate of bioleaching process (Fig.7). The results also confirm the severe reductive effect of EDTA on efficiency of the bioleaching process, such that the bioleaching rate in 5 and 10% concentrations of EDTA decreased by 61 and 70%, respectively in comparison with the sample without EDTA (Fig.8).

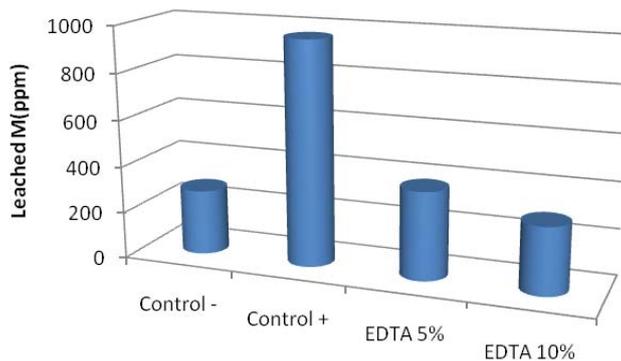


Figure 8: EDTA effect on extraction of iron from pyrite ore  
Control - = leaching medium without bacterium and EDTA  
Control + = leaching medium + bacterium without EDTA.

### XVII . STUDY OF THE ACTIVITY OF LPS-LACKING BACTERIA AND OXIDATION OF $Fe^{2+}$ IRON IN TF MEDIUM

For this purpose, the bacterial LPS was first removed and the LPS-lacking bacteria were inoculated into the TFB medium. No color change was observed after 10 days, which represents the inability of the bacteria in the oxidation of  $Fe^{2+}$  to  $Fe^{3+}$ , whilst in the control sample with normal bacterium, the color began

to change in the 5 day and it remarkably turned from green to red after the 10<sup>th</sup> day (Fig.9, right). To ensure that the EDTA + Tris-HCl treatment has not killed the bacteria and the cells just have lost their LPS, in the second stage, the LPS-lacking bacteria used in this experiment were transferred to a fresh TF medium. After 10 days, the bacterium turned the medium from lime green to red, indicating that the bacteria have restored the ability of LPS synthesis and have oxidized  $Fe^{2+}$  to  $Fe^{3+}$  (Fig.9, left).

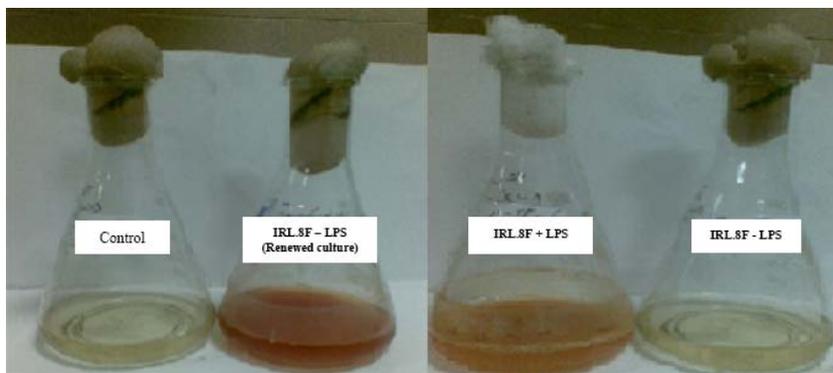


Figure 9 : Two stages of bacterial cell culture in TF medium. Right: the initial culture of LPS-lacking bacteria in TF medium, Left: re-culturing the LPS-lacking bacteria in the fresh TF medium.

### XVIII . STUDY OF THE CAPABILITY OF LPS-LACKING BACTERIA IN LEACHING OF $Fe^{3+}$ FROM PYRITE OR

The samples were analyzed after 14 day from beginning of the process and the amount of

leached iron was obtained as shown in Figure 10. As it can be observed from the figure, LPS-lacking bacteria have remarkably lost their ability to leaching the iron.

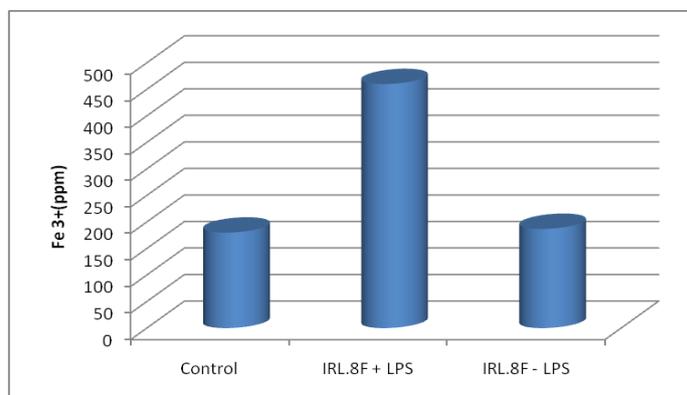


Figure 10 : Concentration of extracted  $Fe^{3+}$  from pyrite ore, in presence of LPS-lacking *A. ferrooxidans* IRL.8F.

## XIX . DISCUSSION

In the present study with the aim of evaluating the ability of native bacterium to extract copper and iron from their ores, an acidophilic strain was isolated from an iron mineral spring in Larzan, Iran. The isolated strain shows a remarkable enzymatic activity in Fe and S oxidation and is highly capable with bioleaching the copper and pyrite.

The results from bioleaching of copper concentrate showed that the amount of extracted copper and iron was 71.4 and 29.3%, respectively. Furthermore, in comparison with control samples, without bacteria, these amounts increased by 93.5 and 92%, respectively. In the control samples, however, a minor amount of metals were extracted due to spontaneous leaching.

One of the factors influencing the quality and quantity of the bioleaching process is the bacterial ability to attach to the mineral surface. The microbial contact with the ore surface stimulates the expression and production of extracellular polymers which entrap the bacterium at the side of the ore and attach it to the mineral surface (Dispirito et al., 1983; Bagdigian and Meyerson, 1986).

In addition, EPS can form chemical bonds with the mineral surface and mediate or promote respiration and nutritional reactions (Scobar., 1997; Ehrlich and Brierley., 1990). The molecules constituting the EPS can be made of LPS, phospholipids or other macromolecules such as poly-peptides. These compounds are released by the organism when it is attached to the mineral (Scobar, 1997).

There are various techniques to isolation of LPS from bacteria, Such as phenol-water method or by the phenol-chloroform-petroleum ether extraction methods, but these methods usually cause cellular damage or death. (Ramadas et al., 1991). The purpose of this study was to isolate and remove bacterium LPS without causing bacterial cell damage or death. Studies show that EDTA treatment negatively affects the adherence of the cell to mineral by the loss of part of LPS, without cell lysis. (Arredondo et al., 1994; Scobar et al 1997).

Scobar et al (1997) investigated the effect of EDTA on iron extraction from the pyrite ore. They believed that this substance removes the LPS from bacterial outer membrane and this leads to a remarkable decrease in the attachment of bacterium to its substrate. In the investigation of chalcopyrite and pyrite ores, they observed 85 and 77% decrease in attachment, respectively for bacterial cells treated with EDTA (Scobar, 1997). Such substances as EDTA absorb bivalent cations attached to phosphate groups in LPS and transform it from natural form to aggregated form, which obstruct the subsequent reactions (Rangin and Basu, 2004). Results of the

present study revealed that samples with EDTA treatments show a remarkable decrease in bioleaching rate of the metal of interest. In direct mechanism where contact and attachment of the bacterium to the mineral surface is mediated by releasing exopolymers (Vandevivere and Kirchman., 1993) , EDTA removes part of this exopolymer and thereby, to a great extent decreases the efficiency of iron extraction from pyrite ore (Arredondo et al., 1994; Scobar et al 1997).

Since the bioleaching drastically decreased with LPS removal, in the present study, it can be concluded that the most amount of metal has been extracted through direct mechanism. The bacterium secretes such substances as LPS when approaching the mineral surface in order to be able to attach to the mineral surface; however, when LPS is removed the attachment cannot occur and the leaching by the bacterium will decrease (Pogliani and Donati., 1999; Arredondo et al., 1994; Scobar et al 1997). In the next experiments, to ensure that the decrease in bioleaching has resulted from LPS removal by EDTA, the bacterial LPS was removed by use of EDTA treatment and LPS-lacking bacteria were transferred to the leaching medium. Cultures of these bacteria in leaching media containing pyrite soil also significantly showed the decrease in extraction of  $Fe^{3+}$ . These bacterial cells were also cultured in iron-containing TF medium. As it was expected, the treated bacteria with EDTA could not oxidize the iron, whereas the iron oxidation was observed in the Erlenmeyer flask containing non-treated bacteria.

Furthermore, EDTA in the leaching medium may act as a chelator absorbing the iron cations and decrease the oxidation of iron from  $Fe^{2+}$  to  $Fe^{3+}$ ; therefore, the bacterial LPS may be of no role in decrease of leaching. LPS removal and inoculation of LPS-lacking bacteria into the leaching medium led to a 60.1% decrease in extracted metal. The less decrease in metal in comparison with when EDTA was used can be attributed to three possible reasons: 1) Some bacterial cells have restored their ability to produce LPS, 2) EDTA has acted as an iron chelator, or 3) EDTA has decreased the enzymatic oxidation of the iron.

In bacterial bioleaching, Thiobacillus Thiooxidans is used together with T. ferrooxidans, for the following reasons: It can release metallic elements by oxidation of reduced and semi-reduced sulfur compounds of the minerals and can promote the leaching of metals by producing the sulfuric acid as an oxidant. In addition, it provides the optimal acidic conditions for growth and activity of T. ferrooxidans. In bioleaching processes, that bacterial strain is of the greater importance which produces more amount of acid ( Qiu et al., 2005).

In conclusion, results obtained from the present study indicate that: 1) EDTA drastically

decreases the efficiency of the bioleaching process, 2) LPS in the isolated bacterial strain in this study has a key role in bacterial attachment to mineral particles, and 3) the bioleaching process in this case promotes through the direct mechanism.

In Fe oxidation in TFB, pH of the medium decreased from 2.5 to 2, Probably because of consumption of sulfur compounds in the medium. Deposits in fawn color were observed on the wall of flask, which increased in amount daily; probably the iron hydroxide(jarosite) appeared in pH<2 ( Qiu et al., 2005).

## REFERENCES RÉFÉRENCES REFERENCIAS

- Arredondo R, Garcia A, Jerez C (1994). Partial Removal of Lipopolysaccharide from *Thiobacillus ferrooxidans* Affects Its Adhesion to Solids. *Applied And Environmental Microbiology*. 60: 2846-2851.
- Bagdigian RM, Meyerson AS (1986). The adsorption of *Thiobacillus ferrooxidans* on coal surfaces. *Biotechnol Bioeng*, 28: 467-479.
- Chen S, Lin JG, (2000). Influence of solid content on bioleaching of heavy metals from contaminated sediment by *Thiobacillus* sp. *Chemistry technology biotechnology*. 75: 649-659.
- Clausen C (2003). Reusing Remediated CCA-Treated Wood, Special Seminar sponsored by American Wood-Preservers' Association Utility Solid Waste Activities Group.
- DiSpirito AA, Dugan PR, Tuovinen OH (1983). Sorption of *Thiobacillus ferrooxidans* to particulate material. *Biotechnol Bioeng*. 25: 1163-1168.
- Donati ER, Sand W (2007). *Microbial Processing of Metal Sulfides*. Springer book.
- Ehrlich EC, Brierley CL (1990). *Microbial Mineral Recovery*. McGraw-Hill, New York.
- Elzaky M, Attia YA, (1989). Bioleaching of gold pyrite tailing with adapted bacteria. *Hydrometallurgy*. 151-159.
- Gehrke T, Telegdi J, Thierry D, and Sand W (1998). Important of extracellular polymeric substances from *Thiobacillus ferrooxidans* for bioleaching. *Appl. Environ. Microbiol*. 2743-2747.
- Hong P., Yang Y., Li X., Qiu G., Liu X., Huang J. and Hu Y. Structure analysis of 16s rDNA sequences from strains of *Acidithiobacillus ferrooxidans*. *Journal of biochemistry and molecular biology* 39: 178-182, 2006.
- Makita M, Esperón M, Pereyra B, Lopez A, Orrantia E (2004). Reduction of arsenic content in a complex galena concentrate by *Acidithiobacillus ferrooxidans*. *BMC Biotechnology*. 1-23.
- Ohba H, Owa N (2005). Isolation and identification of sulfur-oxidizing bacteria from the buried layer containing reduced sulfur compounds of a paddy field on sado island in niigata prefecture. *Nigata Unit*. 5: 55-61,.
- Pogliani C, Donati E (1999). The role of exopolymers in the bioleaching of a non-ferrous metal sulphide. *Journal of Industrial Microbiology & Biotechnology*. 22:88-92.
- Qiu M, Xiong S, Zhang W, Wang G (2005). A comparison of bioleaching of chalcopyrite using pure culture or a mixed culture. *Mineral Engeneering*. 18: 987-990.
- Ramadas U, Carlson RW, Busch M, Mayer H, (1991). Distribution and Phylogenetic Significance of 27-Hydroxy-Octacosanoic Acid in Lipopolysaccharides from Bacteria Belonging to the Alpha-2 Subgroup of *Proteobacteria*. *International Journal Of Systematic Bacteriology*. 41: 213-217.
- Rangin M. and Basu A (2004). Lipopolysaccharide Identification with Functionalized Polydiacetylene Liposome Sensors. *American Chemical Society*. 126: 5038-5039.
- Rawlings DE (2002). Heavy metal mining using microbes. *Annual Review Microbiology*. 56: 65-91.
- Sasaki K, Nakamuta Y, Hirajima T, Tuovinen OH (2009). Raman characterization of secondary minerals formed during chalcopyrite leaching With *Acidithiobacillus ferrooxidans*. *Hydrometallurgy*. 95:153-158.
- Scobar B, Huerta G and Rubio J (1997). Short communication :influence of lipopolysaccharides on the attachment of *Thiobacillus ferrooxidans* to minerals. *W.J of Microbiol & Biotechnol*. 13: 593-594.
- Vandevivere P, Kirchman DL (1993). Attachment stimulates exopolysaccharide synthesis by a bacterium. *Appl Environ Microbiol*. 59: 3280-3286..
- Wolfgang S, Edgardo RD (2007). *Microbial Processing of Metal Sulfides*. University of La Plata, Argentina. 169-191.
- Yeats C, Gillings M, Davison A, Altavilla N, Veal DA (1998). Methods for microbial DNA extraction from soil for PCR amplification. *Biological Procedures Online*. 1: 40-45.

This page is intentionally left blank