

DETECTION OF MOLECULAR MARKERS BY COMPARATIVE SEQUENCE ANALYSIS OF ENZYMES FROM *MYCOBACTERIA* SPECIES

SABA ASAD¹, MADEEHAHUSSAIN¹, AYESHA SIDDEQUA¹, QURRAT UL AIN¹ AND HABIB BOKHARI^{1*}

¹Department of Biosciences, COMSATS Institute of Information Technology,
Park Road, Chak Shahzad, Islamabad, Pakistan.

*Correspondence author e-mail: habib@comsats.edu.pk; Telephone +92-300-5127684; Fax 0092-051-4442805

Abstract

Mycobacterial species are one of the most important pathogens and among these members of non-tuberculous mycobacteria (NTM) and mycobacterial tuberculosis complex (MTC) are the causative agent of a relatively milder form of Tuberculosis. Traditional methods for identification of these groups of pathogens are time consuming, lack specificity and sensitivity and furthermore lead to the misidentification due to high similarity index. Therefore, more rapid, specific and cost-effective methods are required for the accurate identification of *Mycobacterium species* in routine diagnostics. In our study, we identified molecular markers in order to differentiate closely related cousin species of genus *Mycobacterium* including *M. bovis*, *M. avium*, *M. leprae* and *M. tuberculosis*. The nucleotide sequences of selected unique markers, i.e., enzymes (used previously in various biochemical tests for the identification of *M. species*) were selected and their ORFs were retrieved and selected functional proteins of respective biosynthetic pathways were compared *in-silico*. Result suggested that the variations in nucleotide sequences of the selected enzymes can be directly used for *M. species* discrimination in one step PCR test. We believe that the *in-silico* identification and storage of these distinctive characteristics of individual *M. species* will help in more precise recognition of pathogenic strains and hence specie specific targeted therapy.

Introduction

Mycobacteria are slow growing, strictly aerobic, and form slightly curved to straight rods with a width of 0.2–0.6 µm and length of 1.0–10 µm. The *Mycobacterium* genus has been the subject of a significant research effort and about 100 species are now recognized. They cause a variety of illnesses in humans and animals that differ in severity and public health importance (Wolinsky, 1983).

Identification of *Mycobacteria* to the species level is important for epidemiological investigations, public health and therapeutic reasons. Early detection and identification of *M. tuberculosis* is particularly important because it can be transmitted from one person to another, and because diseases due to non-tuberculous mycobacterium (NTM) require adapted treatment regimens. Furthermore, misidentification may occur, because different species may have indistinguishable morphological and biochemical profiles. Newly discovered species are even more difficult to identify because of the no characteristic phenotypic patterns. The culture-dependent laboratory procedures may take 7 to 10 days on most recently developed liquid media and 4 to 6 weeks on solid media. It is necessary for the clinical laboratory to consider alternative strategies to address these limitations of the conventional approach (Wolinsky, 1983; Maeda *et al.*, 2000; Fukushima *et al.*, 2003; Turi *et al.*, 2012). Moreover, most of the NTM infections do not respond to conventional existing anti-tuberculosis treatment and are misdiagnosed as drug resistant strains of *M. tuberculosis* due to lack of species identification (David, 1970; Ratanasuwana *et al.*, 2002).

In order to address the above-mentioned limitations of traditional laboratory diagnosis of TB, the current study has resulted in the emergence of novel approaches for rapid and reliable detection of *Mycobacterial species*. Since long, 16S rRNA has been used as an authenticated and reliable marker for identification of *Mycobacterial*

species. However, various discrepancies have been noticed in the results obtained through comparison of 16S rRNA sequences from different databases, hence leading to erroneous conclusions (Bono *et al.*, 1998; DeSantis, 2006). In this study, we have identified useful molecular markers for rapid identification of members of NTM and MTB complex.

Materials and Methods

Biochemical testing: Delineation of the biochemical activities of a microbial isolate is the most convenient way to narrow the search path towards the identity of an unknown strain. The comparative sequence analysis of genes encoding the enzymes involved in the outcome of the biochemical identification tests is a novel approach for the identification of species. Genetic variations in their DNA sequences (single nucleotide polymorphism) are directly used as a marker to differentiate closely related *Mycobacterial species*. We took these biochemical tests results for identification of selected *Mycobacterial species* from Bergey's manual (Fraser, 2007) and picked the enzymes linked to the functional distinguished phenotypes as a unique marker.

Essential genes retrieval: Integrated Microbial Genome Database (IMG) (<http://img.jgi.doe.gov/cgi-bin/pub/main.cgi>) was used to retrieve KEGG entries for biosynthetic pathways, in order to find out the unique genes and set of enzymes involved in particular biosynthetic pathways and collected from biochemical tests (Victor *et al.*, 2006). It was preferred to retrieve the KEGG entries instead of taking out the entries directly from the KEGG as it displays information in a well-organized way. The KEGG (Kyoto Encyclopedia of Genes and Genomes) resource provides a reference knowledge base for linking genomes to biological systems (<http://www.genome.jp/kegg/>).

Unique ORFs retrieval: In order to find the unique ORFs to the genes having role in biosynthetic pathways only, Genomes to Protein (GTOP) (Fukuchi *et al.*, 2009) structures and functions database was considered (<http://spock.genes.nig.ac.jp/~genome/gtop.html>) (Dupanloup & Kaessmann, 2006).

Results and Discussions

Essential biosynthetic enzymes: *M. avium* and *M. leprae* have highly variable genomes as compared to *M. bovis* and *M. tuberculosis*. Hence for nucleotide comparison, those enzymes were collected which are common in all four species having unique sequences. *M. bovis* was found to have certain pathways that are not present in any other closely related species. The unique differences can be exploited to develop a molecular marker to detect medically important Mycobacterial pathogens. Table 1 enlists the enzymes selected as markers and among these few enzymes was found to be completely conserved e.g. Aconitase, short chain enolyl hydratase, Glycerol-3-

phosphate dehydrogenase, Nucleoside diphosphate kinase, etc. Therefore, they cannot be a good marker for differentiation of closely related *Mycobacterial species* in a simple PCR based test whereas other enzymes (DNA polymerase II Beta subunit and Pyruvate dehydrogenase subunit I) showed differences ranging from 2 SNPs to 3 SNPs that can be useful molecular marker for discriminating closely related *Mycobacterial species* i.e. *M. tuberculosis* H37Rv and *M. bovis* AF2122/97.

Unique ORFs: Unique ORFs which have biosynthetic function were filtered out from selected *Mycobacteria species* as shown in Table 2. Moreover, the *in-silico* expression analysis revealed that all the unique ORFs are expressed and hence are considered to play active role in the life cycle of the bacterium. Based on this they were recruited as potential molecular markers as well possibly used along with the other common markers in multiplex PCR reaction for the diagnosis of important *Mycobacterium species*.

Table 1. Unique biosynthetic enzymes functioning in biosynthetic pathways¹.

Biosynthetic enzymes	Main biosynthetic pathway	Sub-pathway
*Aconitase	Carbohydrate metabolism	Citric acid cycle
DNA polymerase III beta subunit	Nucleotide metabolism	Purine metabolism
*Short chain enolyl hydratase	Lipid metabolism	Fatty acid elongation in mitochondria
*Glycerol-3-phosphate dehydrogenase	Lipid metabolism	Glycerophospholipid biosynthesis
*Nucleoside diphosphate kinase	Nucleotide metabolism	Pyrimidine metabolism
Pyruvate dehydrogenase subunit I	Amino acid synthesis	Alanine and aspartate metabolism

¹The enzymes denoted by * are 100% conserved in all *Mycobacterial species* used for study. The other two enzymes show variations as SNP and can be used as important markers to distinguish between *M. bovis* and *M. tuberculosis*.

Table 2. Unique ORFs of selected *Mycobacterial species*.

Species	Gene (ORFs)	Function
<i>M. avium</i>	Gene : <i>cobQ2</i>	Cobyric acid synthase
	Gene : <i>gltB</i>	Glutamate synthase
	Gene : <i>pheA</i>	Prephenatedehydratase
	Gene : <i>ponA.1</i>	Penicillin binding protein transpeptidase
	Gene : <i>asd</i>	Aspartate-semialdehyde dehydrogenase
<i>M. bovis</i> AF2122/97	Gene : <i>cpsY</i>	UPD-glucose-4-epimerase
	Gene : <i>adk</i>	Adenylate kinase
	Gene : <i>coaA</i>	Pantothenate kinase
<i>M. tuberculosis</i> H37Rv	Gene : <i>thyX</i>	Thymidylate synthase
	Gene : <i>CAA17392.1</i>	Regulatory protein
	Gene : <i>CAA17444.1</i>	Nicotinatephosphoribosyltransferase
	Gene : <i>CAB06687.1</i>	Transcriptional regulatory protein
	Gene : <i>Pks15</i>	Polyketide synthase
<i>M. leprae</i> TN	Gene : <i>nadB</i>	L-aspartate oxidase
	Gene : <i>nadC</i>	Nicotinate-Nucleotide pyrophosphorylase
	Gene : <i>folK</i>	Hydroxymethyldihydropterine pyrophosphokinas
	Gene : <i>hisB</i>	Imidazole glycerol phosphate dehydratase
	Gene : <i>thiE</i>	Thyamine phosphate pyrophosphorylase
	Gene : <i>lysS</i>	LysyltRNAsynthetase
	Gene : <i>trpC</i>	Indole-3-glycerol phosphate synthase
	Gene : <i>alr</i>	Alanine racemase
	Gene : <i>guaA</i>	BifunctionalGMP synthase
	Gene : <i>guaB2</i>	Inositol-5-monophosphate dehydrogenase
	Gene : <i>ribG</i>	Riboflavin specific deaminase/reductase
	Gene : <i>purK</i>	Phosphoribosylimidazol carboxylase
	Gene : <i>trpS</i>	TryptophenyltRNAsynthetase
	Gene : <i>bccA</i>	Acetyl coA carboxylase
	Gene : <i>folB</i>	Dihydroneotasealdolase

Single nucleotide polymorphism: Moreover, common ORFs from the *Mycobacterial species* were shortlisted and further screened for the possible differences in their nucleotide sequences. We found total 15 common ORFs from all the selected *Mycobacterial species* (*M. bovis*, *M. avium*, *M. leprae* and *M. tuberculosis*). However, only the sequences from *M. bovis* and *M. tuberculosis* were aligned for comparison. Out of 15 common conserved ORFs, single nucleotide polymorphism (SNP) was found in *opd*, *aroF*, *glmU* and *leuS* genes (Table 3).

Furthermore, the *in-silico* restriction fragment length polymorphism (RFLP) pattern suggested that *opd*, *glmU*, *leuS* and *aroF* can be a useful candidates due to the distinguished restriction patterns while using same restriction enzymes on DNA of *M. tuberculosis*, *M. bovis*, *M. avium* and *M. avium* subsp. *paratuberculosis* respectively. In nutshell these differences can then be utilized as an additional marker to differentiate closely related *Mycobacterial species* (Table 4). The linked enzymes and associated ORFs were stored in the form of a table along with respective similarities and differences (Table 5).

There is a strong need that knowledge base solutions to improve the quality of life can be best developed by the developing countries themselves keeping in mind their national needs. This is mainly because of the fact that

companies and industries based in developed world will likely to focus on the big market potentials rather than to produce products e.g. Health products, for people in the poorer parts of the world. This is clearly seen by the lack of new drugs to treat infectious diseases such as multi-drug resistant Tuberculosis (MDR-TB). Pakistan, being one of the largest country populations wise (> 170 Millions) where every year large number of Tuberculosis cases are reported has to come up with some in house solutions not in understanding its dynamics and spread at least.

In this scenario we believe that small studies like ours where we have identified useful molecular markers for rapid identification of members of NTM and MTB complex using *in-silico* tools can be useful to solve indigenous health issues. Biosynthetic pathways are dependent on the genes that are essential for bacterial growth and they are found to be diverted, either due to an ancient transfer of genes or deep paralogy and subsequent retention of the genes in unrelated lineages (Coenye *et al.*, 2005; Dupanloup & Kaessmann, 2006; Mahmood *et al.*, 2010). In our study we discovered a number of ORFs which are linked to biosynthetic pathways, the minor change or variation in their nucleotide sequence can be exploited to develop molecular marker in identification of *Mycobacterial species*.

Table 3. Identification of single-nucleotide polymorphism (SNP) and restriction mapping in unique linked ORFs. These variations were obtained by finding the differences in sequence alignments.

Name of ORF	Alignment results	Restriction mapping
Gene: <i>opd</i>	One SNP (single nucleotide polymorphism) found	Different fragment patterns for <i>MTB</i> , <i>M. bovis</i> , <i>M. avium</i> & <i>M. avium</i> subsp. <i>paratuberculosis</i>
Gene: <i>aroF</i>	One SNP (single nucleotide polymorphism) found	No difference found except <i>M. avium</i> subsp. <i>paratuberculosis</i>
Gene: <i>aroK</i>	Completely conserved	No difference found
Gene: <i>cysS</i>	Completely conserved	No difference found
Gene: <i>pdxH</i>	Completely conserved	No difference found
Gene: <i>purF</i>	One SNP (single nucleotide polymorphism) found	No difference found
Gene: <i>purl</i>	Completely conserved	No difference found
Gene: <i>purQ</i>	Completely conserved	No difference found
Gene: <i>prsA</i>	Completely conserved	No difference found
Gene: <i>aspS</i>	One SNP (single nucleotide polymorphism) found	No difference found
Gene: <i>hiss</i>	Completely conserved	No difference found
Gene: <i>glgC</i>	One SNP (single nucleotide polymorphism) found	No difference found
Gene: <i>alaS</i>	One SNP (single nucleotide polymorphism) found	No difference found
Gene: <i>glmU</i>	Two SNP (single nucleotide polymorphism) found	Different fragment patterns for <i>M. bovis</i> & <i>M. avium</i> from others
Gene: <i>leuS</i>	Two SNP (single nucleotide polymorphism) found	Different fragment patterns for <i>M. avium</i> subsp. <i>Paratuberculosis</i> & <i>M. avium</i> from others

Conclusions

The incidence of Tuberculosis caused by members of MTC is increasing worldwide in particular in the developing countries. However, the percentage contribution of members of NTM in causing active TB is also increasing. The conventional techniques used for the identification of *Mycobacterial species* were either time consuming or tedious. We identified novel and distinguishable molecular markers for quick identification of these closely related *Mycobacterial species* using *in-silico* approach. The *in-silico* approach identified single nucleotide polymorphism (SNP) in expressed open reading frames (ORFs) and their linkage with important

biosynthetic pathways and enzymes essentially present in all closely related species. SNPs found in DNA polymerase III beta subunit and pyruvate dehydrogenase subunit I of *M. tuberculosis* and *M. bovis* and more than 15 SNPs found for *M. laprae* and *M. avium* can be well exploited for their selective identification in simple PCR test. The rationale of the study was to see whether the small possible difference(s)/polymorphism in nucleotide sequences of these enzymes can act as potential markers for discriminating selected closely related *Mycobacterial species*. Our study in this regard provides a prototype, which uses set of unique approaches in order to classify the closely related species of *Mycobacteria* as well as other microbes.

Table 4. *Mycobacteria species* showing difference in restriction mapping along with their respective fragments

Organism/Accession No.	Gene	Enzymes	Restriction site	No. of fragments	Size(s) bp
MTB/gi 50953765	<i>opd</i>	Sau96I	G [^] GNC_C	10	237,255,291,384,626,647,759,804,953, 1016
MTB F11/ gi 148719718	//	//	//	9	237,255,291, 384 ,626,647,759,804,953
<i>M. bovis</i> BCG/ gi 121491530	//	//	//	8	237,255,291,626,647,759,804,953
<i>M. bovis subsp/ gi 31616762</i>	//	//	//	8	237,255,291,626,647,759,804,953
<i>M. avium/ gi 118163506</i>	//	//	//	8	159,270,271 ,292,627,663, 906,907
<i>M. avium. subsp. paratuberculosis/ gi 41400296</i>	//	//	//	10	11,12 ,255,291,626,647,648,759,952,953
		Sau96I	G [^] GNC_C	17	131,187,222,241,247,454,472,473,572,606,907,1222,1435,1441,1442,1455,1461
<i>M. tuberculosis/ gi 57116681</i>	<i>glmU</i>	MspAII	CMG [^] CKG	7	80,482,708,759,966,1140,1466
		PvuII	CAG [^] CTG	2	708,759
		Sau96I	G [^] GNC_C	17	131,187,222,241,247,454,472,473,572,606,907,1222,1435,1441,1442,1455,1461
<i>M. tuberculosis/ gi 50953765</i>		MspAII	CMG [^] CKG	7	80,482,708,759,966,1140,1466
	//	PvuII	CAG [^] CTG	2	708,759
		Sau96I	G [^] GNC_C	17	131,187,222,241,247,454,472,473,572,606,907,1222,1435,1441,1442,1455,1461
<i>M. bovis</i> BCG/ gi 121635883		MspAII	CMG [^] CKG	7	80,482,708,759,966,1140,1466
	//	PvuII	CAG [^] CTG	2	708,759
		Sau96I	G [^] GNC_C	18	131,187,222,241,247,454,472,473,572,606,753,907,1222,1435,1441,1442,1455,1461
<i>M. bovis/ gi 31791177</i>		MspAII	CMG [^] CKG	8	80,482,708, 750 ,759,966,1140,1466
	//	PvuII	CAG [^] CTG	3	708, 750 ,759
		Sau96I	G [^] GNC_C	17	24,116 ,207,226,232,439,458,557,591,1069,1216,1248,1261, 1302 ,1426,1427,1446
<i>M. avium/ gi 118462219</i>		MspAII	CMG [^] CKG	7	600 ,693,735,1030,1067, 1125,1213,1278 ,1451
	//	PvuII	CAG [^] CTG	2	600 ,693,735
		Sau96I	G [^] GNC_C	17	24,116,207,226,232,439,458,557,591,1069,1216,1248,1261,1302,1426,1427,1446
<i>M. avium. subsp. paratuberculosis/ gi 41406098</i>	//	MspAII	CMG [^] CKG	9	600,693,735,1030,1067,1125,1213,1278,1451
		PvuII	CAG [^] CTG	3	600,693,735
	<i>leuS</i>	Sau96I	G [^] GNC_C	22	22,110,128,129,174,318,410,447,644,1013,1490,1518,1767,1912,1933,2176,2308,2333,2582,2622,2683,2832
<i>M. tuberculosis/ gi 50953765</i>		MspAII	CMG [^] CKG	11	19,193,658,1344,1516,1523,1575,1917,2313,2373,2619
		PvuII	CAG [^] CTG	--	

Table 4. (Cont'd.).

<i>M. bovis</i> / gi 31791177	//	Sau96I	G^GNC_C	22	22,110,128,129,174,318,410,447,644,1013,1490,1518,1767,1912,1933,2176,2308,2333,2582,2622,2683,2832
		MspAII	CMG^CKG	11	19,193,658,1344,1516,1523,1575,1917,2313,2373,2619
		PvuII	CAG^CTG	--	
<i>M. bovis</i> BCG/ gi 121635883	//	Sau96I	G^GNC_C	22	22,110,128,129,174,318,410,447,644,1013,1490,1518,1767,1912,1933,2176,2308,2333,2582,2622,2683,2832
		MspAII	CMG^CKG	11	19,193,658,1344,1516,1523,1575,1917,2313,2373,2619
		PvuII	CAG^CTG	--	
<i>M. avium</i> /gi 118462219	//	Sau96I	G^GNC_C	31	50,165,197,243,257,264,479,516,531,642,713,1014,1113,1248,1481,1830,1886,1975,1982,1996,2036,2100,2239,2370,2371,2492,2727,2752,2769,2946,2952
		MspAII	CMG^CKG	10	154,182,339,513,857,1980,2237,2376,2436,2667
		PvuII	CAG^CTG	2	182,513
<i>M. avium</i> .subsp.paratuberculosis /gi 41406098	//	Sau96I	G^GNC_C	29	96,128,174,188,195,410,447,462,573,668,945,1044,1179,1761,1817,1906,1913,1927,1967,2031,2170,2302,2423,2499,2658,2683,2700,2877,2883
		MspAII	CMG^CKG	9	85,113,270,444,788,1911,2307,2367,2598
		PvuII	CAG^CTG	2	113,444
<i>M. tuberculosis</i> /gi 57116681	//	<i>aroF</i> Sau96I	G^GNC_C	14	117,322,352,353,519,651,658,671,799,864,958,991,992,1092
		MspAII	CMG^CKG	9	285,388,465,490,663,873,1069,1098,1187
		PvuII	CAG^CTG	2	465,1098
<i>M. bovis</i> / gi 31791177	//	Sau96I	G^GNC_C	14	117,322,352,353,519,651,658,671,799,864,958,991,992,1092
		MspAII	CMG^CKG	9	285,388,465,490,663,873,1069,1098,1187
		PvuII	CAG^CTG	2	465,1098
<i>M. bovis</i> BCG/ gi 121635883	//	Sau96I	G^GNC_C	15	4,120,220,254,344,413,520,541,552,553,837,859,873,890,1095
		MspAII	CMG^CKG	11	146,303,342,433,448,725,750,760,827,930,972
		PvuII	CAG^CTG	2	303,750

Table 5. List of ORFs of four *Mycobacterial* species with their linked biosynthetic pathways and corresponding functions.

Specie name	Unique ORFs	Linked enzymes	Linked biosynthetic pathways
<i>M. bovis</i>	<i>cpsY</i>	UPD-glucose-4-epimerase	Galactose metabolism, purine metabolism, Pyrimidine metabolism
	<i>adk</i>	Adenylate kinase	Purine metabolism
<i>M. avium</i>	<i>gltb</i>	Glutamate synthase	Lysine degradation
	<i>asd</i>	Aspartate-semialdehyde dehydrogenase	Glycine, serine and threonine metabolism, Lysine biosynthesis
<i>M. tuberculosis</i>	<i>thyX</i>	Thymidylate synthase	Nitrobenzene degradation, Tryptophan metabolism, Type II secretion system, Ubiquinone biosynthesis, Histidine metabolism, Androgen and estrogen metabolism, Tyrosine metabolism, Aminophosphonate metabolism, Selenoamino acid metabolism
	<i>nadB</i>	L-aspartate oxidase	Alanine and aspartate synthesis, Nicotinate and Nicotinamide synthesis
	<i>nadC</i>	Nicotinate-nucleotide pyrophosphorylase	Nicotinate and nicotinamide synthesis
	<i>folK</i>	Hydroxymethyl dihydropterine pyrophosphokinase	Folate biosynthesis
	<i>lysS</i>	Lysyl tRNA synthetase	Lysine biosynthesis, Aminoacyl tRNA synthesis
	<i>alr</i>	Alanine racemase	Alanine and aspartate metabolism, D-alanine metabolism
	<i>guaA</i>	Bifunctional GMP synthase	Purine metabolism, Glutamate metabolism
	<i>guaB2</i>	Inositol-5-monophosphate dehydrogenase	Purine metabolism
	<i>purK</i>	Phosphoribosyl imidazol carboxylase	Purine metabolism
	<i>trpS</i>	Tryptophenyl tRNA synthetase	Tryptophan metabolism, Aminoacyl tRNA synthesis
<i>M. leprae</i>	<i>bccA</i>	Acetyl coA carboxylase	Fatty acid biosynthesis
	<i>folB</i>	Dihydroneotase aldolase	Folate biosynthesis

Acknowledgments

We would like to thank Department of Biosciences and COMSATS Institute of Information Technology for their kind support.

References

- Bono, H., H. Oqata, S. Goto and M. Kanehisa. 1998. Reconstruction of amino acid biosynthesis pathways from the complete genome sequence. *Genome Res.*, 8: 203-210.
- Coenye, T., D. Gevers, Y. Van de Peer, P. Vandamme and J. Swings. 2005. Towards a prokaryotic genomic taxonomy. *FEMS Microbiol. Rev.*, 29: 147-167.
- David, H.L. 1970. Probability distribution of drug-resistant mutants in unselected populations of *Mycobacterium tuberculosis*. *Appl. Microbiol.*, 20: 810-14.
- DeSantis, T.Z., P. Hugenholtz, K. Keller, E.L. Brodie, N. Larsen, Y.M. Piceno, R. Phan and G.L. Andersen. 2006. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucl. Acid. Res.*, 34: W394-9.
- Dupanloup, I. and H. Kaessmann. 2006. Evolutionary simulations to detect functional lineage-specific genes. *Bioinformatics*, 22: 1815-1822.
- Fraser, C., W.P. Hanage and B.G. Spratt. 2007. Recombination and the nature of bacterial speciation. *Science.*, 315: 476-480.
- Fukuchi, S., K. Homma, S. Sakamoto, H. Sugawara, Y. Tateno, T. Gojobori and K. Nishikawa. 2009. The GTOP database in 2009. Updated content and novel features to expand and deepen insights into protein structures and functions. *Nucl. Acids Res.*, 37: D333-7.
- Fukushima, M., K. Kakinuma, H. Hayashi, H. Nagai, K. Ito and R. Kawaguchi. 2003. Detection and identification of *Mycobacterium* species isolates by DNA microarray. *J. Clin. Microbiol.*, 41: 2605-2615.
- Maeda, K.M., J.T. Rhee, T.R. Gingeras, H. Salamon, J. Drenkow, N. Smittipat and P.M. Small. 2001. Comparing genomes within the species *Mycobacterium tuberculosis*. *Genome Res.*, 11: 547-554.
- Mahmood, T., S. Muhammad and Z.K. Shinwari. 2010. Molecular and morphological characterization of *Caralluma* Species. *Pak. J. Bot.*, 42(2): 1163-1171.
- Ratanasuwan, W., W. Techasathit, V. Chuenarom, S. Suwanagool, T. Anekthanont, J. Jearanaisilavong and A. Chaiprasert. 2002. Infection due to nontuberculous *Mycobacterium* other than MAC in AIDS patients at Siriraj hospital during 1998-2000: saprophyte vs pathogen. *J. Med. Assoc. Thai.*, 85: 886-93.
- Turi, N.A., Farhatullah, M.A. Rabbani and Z.K. Shinwari. 2012. Genetic diversity in the locally collected *Brassica* species of Pakistan based on Microsatellite Markers. *Pak. J. Bot.*, 44(3): 1029-1035.
- Victor, M.M., I.M. Chen, K. Palaniappan, K. Chu, E. Szeto, Y. Grechkin, A. Ratner, B. Jacob, J. Huang, P. Williams, M. Huntemann, I. Anderson, K. Mavromatis., N.N. Ivanova and N.C. Kyrpides. 2006. IMG: the integrated microbial genomes database and comparative analysis system. *Nucl. Acid. Res.*, 34: 344-348.
- Wolinsky E. The role of scotochromogenic mycobacteria in human disease. 1983. *Ann. New York Acad. Sci.*, 28: 67-71.