Original Article

Molecular based identification and phylogenetic relationship by using cytochrome b gene of *Pangasius pangasius*

Identificação de base molecular e relação filogenética usando o gene citocromo b de *Pangasius pangasius*

S. Hassan^a, M. Naeem^a, M. Farhan Nasir^b, P. Riaz^a, M. N. Khan^c and I. Atiq^d

^aBahauddin Zakariya University, Institute of Zoology, Multan, Punjab, Pakistan

^bUniversity of Education, Department of Zoology, Division of Science & Technology, Lahore, Punjab, Pakistan ^cUniversity of Jhang, Department of Zoology, Jhang, Punjab, Pakistan

^dGhazi University Dera Ghazi Khan, Department of Zoology, Punjab, Pakistan

Abstract

Molecular appraoch for identification of unknown species by using Cytochrome b gene is an effective and reliable as compared with morphological based identification. For DNA barcoding universal molecular genes were used to identify the species. Cytochrome b is a specific gene used for identification purpose. DNA barcoding is a reliable and effective method compared to the different traditional morphological methods of specie identification. So,in the present study which was conducted to identify the species, a total of 50 fish samples were collected from five different sites. DNA was extracted by using the Phenol Chloroform method from muscle tissue. Five sequences were sequenced (one from each site), analyzed, and identified specific species as *Pangasius pangasius*. Identified sequences were variable in length from 369 bp (Site 1), 364 bp (Site 2), 364 bp (Site 3), 352 bp (Site 4), and 334 bp (Site 5). Identity matches on the NCBI database confirmed the specific species as *P. pangasius*. A distancing tree was drawn to show maximum likelihood among the same and different species. Yet, in many cases fishes on diverse development stages are difficult to identify by morphological characters. DNA-based identification methods offer an analytically powerful addition or even an alternative tool for species identification and phylogenetic study. This work intends to provide an updated and extensive overview on the DNA based methods for fish species identification by using Cytochrome b gene as targeted markers for identification purpose.

Keywords: identification, phylogenetic relation, Cytochrome b, DNA barcoding, P. pangasius.

Resumo

A abordagem molecular para identificação de espécies desconhecidas usando o gene citocromo b é eficaz e confiável em comparação com a identificação baseada na morfologia. Códigos de barras de DNA de genes moleculares universais foram usados para identificar as espécies. O citocromo b é um gene específico usado para fins de identificação. O código de barras de DNA é um método confiável e eficaz em comparação com os diferentes métodos morfológicos tradicionais de identificação de espécies. Assim, no presente estudo, que foi realizado para identificar as espécies, um total de 50 amostras de peixes foram coletadas em cinco locais diferentes. O DNA foi extraído usando o método Fenol Clorofórmio do tecido muscular. Cinco sequências foram sequenciadas (uma de cada local), analisadas e identificadas espécies específicas, como Pangasius pangasius. As sequências identificadas tinham comprimento variável de 369 bp (Local 1), 364 bp (Local 2), 369 bp (Local 1), 364 bp (Local 3), 352 bp (Local 4) e 334 bp (Local 5). As correspondências de identidade no banco de dados do NCBI confirmaram a espécie específica como P. pangasius. Uma árvore de distanciamento foi desenhada para mostrar a máxima probabilidade entre elas e diferentes espécies. No entanto, em muitos casos, peixes em diversos estágios de desenvolvimento são difíceis de identificar por caracteres morfológicos. Os métodos de identificação baseados em DNA oferecem uma adição analiticamente poderosa ou mesmo uma ferramenta alternativa para identificação de espécies e estudo filogenético. Este trabalho pretende fornecer uma visão geral atualizada e abrangente sobre os métodos baseados em DNA para identificação de espécies de peixes usando o gene citocromo b como marcadores direcionados para fins de identificação.

Palavras-chave: identificação, relação filogenética, Citocromo b, DNA barcoding, P. pangasius.

*e-mail: dr_naeembzu@yahoo.com Received: September 19, 2022 – Accepted: October 10, 2022

 \bigcirc

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Morphological characters useful in identification include the shape of a head, mouth shape, size, number of scales, and gill rakers (Talwar and Jhingran, 1991). Morphological variation in fishes may provide a record of population structuring. Many well-documented studies provide evidence for stock discrimination based on traditional morphometric characters. In this context, the DNA barcoding method proved to be a useful approach based on the variation of morphometric and meristic characters for species identification (Hubert et al., 2008). Molecular identification, which identifies species using molecular markers, is widely used today. Among the various molecular approaches used for species molecular identification, DNA barcoding based on mitochondrial DNA (mtDNA) is one of the most suitable tools for specieslevel identification. In addition, mtDNA-based molecular identification has several advantages over morphological approaches. Species identification does not require complete specimens; however, a tiny piece of tissue such as muscle, skin, fin, or teeth is acceptable for DNA extraction, DNA is more stable than morphological characters and is more resistant to degradation (Endo et al., 2022). Fish species identification has been easier because of the use of markers for identification. One of the most often used procedures for species identification is the sequencing and amplification of a core section of DNA barcoding or Cytochrome C oxidase subunit I, which is among the most useful methods for identification (Taylor and Harris, 2012).

The aim of the present study wasto carry out molecularbased identification of *P. pangasius* using Cytochrome b from Pakistan.

2. Materials and Methods

2.1. Species identification and sample collection

50 fish specimens were collected from different fish farms as; Mirza Rizwan Ahmed Fish farm choki Narool, Tehsil Kabir Wala (Site 1); Tariq Khan Faheem Fish Farm Mouza Mumbhal, Kabirwala (Site 2); Nursery unit, Fisheries Department Muzaffargarh (Site 3); Tawakal Fish Hatchery Muzaffargarh (Site 4) and Al Raheem fish Farm Muzaffargarh (Site 5). Specimens were frozen and transported to the fisheries laboratory, Institute of Zoology, Bahauddin Zakariya University, Multan, Pakistan. Samples of fish muscle tissue were taken, preserved in 95 percent ethanol, and stored at 4°C for later use. Formalin was used to preserve voucher specimens.

2.2. Morphological analysis

Different morphometric measurements has been done by takingthe total length and width. Meristic characteristics like scales, fin and lateral line count of the fish samples were also analyzed (Talwar and Jhingran, 1991).

2.3. DNA extraction and amplification

20 mg of muscle tissue have been cut and collected for DNA extraction. Gently homogenized with a solution

of 500 µl of TNES buffer (1M Tris base, 0.5M EDTA, 0.5M NaCl). Then 10µl of Proteinase K was added and incubated at 56 °C for two hours. Phenol, chloroform, and isoamyl alcohol with a ratio of 25:24:1 were added and centrifuged at 13000 rpm for 10 min. Then its supernatant was collected and transferred to a new MCT. Added the same volume of chloroform: isoamyl alcohol and centrifuged at 13000 rpm for 10 min. The supernatant was collected in another micro-centrifuge tube MCT and added chilled ethanol. Put MCT at 20°C overnight for precipitation. Then centrifuged at 5000 rpm for 10-15 min. DNA pellet was visualized at the tube base. After that 100µl of 70% ethanol was added and centrifuged again at 5000 rpm for 15 min. PCR amplification was completed by using the fish primer Fish F1 and Fish R1 (Ward et al., 2005).

Fish F1 CO1 5' AGCCTACGAAAAACCCACCC 3'and Fish R1 CO1 5' AAACTGCAGCCCCTCAGAATGATATTTGTCCTC 3' was sequences by these universal Cyt b primers.

For PCR reaction, total volume was 25 μ l with 1.5 μ l DNA template, 12.5 μ l TaqNova Red mastermix (BLIRT S.A.), 0.1 μ l forward and reverse primers, and 10.8 μ l deionized water (nuclease free). The initial denaturation of 95 °C for 2 min, followed by 30 cycles of thermal cycling conditions The final extension (denaturation at 95 °C for the 30s, annealing at 54 °C for 40s and extending at 72 °C for 1 min) Seven min at 72 °C. The temperature of the annealing was therefore 52 °C for the sample solutions.

2.4. Sequencing, phylogenetic, and statistical analysis

PCR purified products were sent for sequencing to First BASE was shipped to purified PCR products, sequencing laboratories, located in Malaysia. Data of chromatogram sequence has been in FASTA format. The nucleotide sequence was BLAST for identification of homology with its best matches through NCBI.

2.5. Quantification, visualization, and PCR amplification of extracted DNA

Analysis of isolated concentration and purity was examined using NanoDrop Nanophotometer quantification and PCR amplification. The DNA concentration and purity were derived from the Nanophotometer (Implen) quantification absorbance (OD) A_{260}/A_{280} .

2.6. Statistical analysis

BioEdit software 7.0 was used for the removal of ambiguous sequences. MEGA X software was used to measure evolutionary relationships (Kumar et al., 2018). Kimura 2 Parameter of inter and intra specie distance model by Kimura (1980).

3. Results

3.1. Identification

3.1.1. Morphometric identification

The body waslaterally compressed and elongated. Rounded snout obtusely and unpolished upper side of the head was observed. With a moderate mouth gap, the lower jaws were smaller than the upper jaws. A serrated dorsal spine anteriorly was found. Two pairs of barbels, and dorsal spines were comparatively weaker than pectoral and internally serrated. Deeply forked caudal fin and complete lateral line was present. Silver color abdomen, a golden tinge on the head, whitish-grey above the lateral line, on flank silvery purple color, and the back side was yellowish dark/green (Figure 1).

3.1.2. Fin formula

Dorsal.1/7; Pectoral 1.1/12; Pectoral 2.6; Anal. 26-29/3-4 (Rahman, 1989)

B. ix-x; Dorsal i.1/7; Dorsal 2.0; Pectoral. 1/12; Ventral.6; a.31-34 (27-29/4-5) (Bhuiyan, 1964)

Dorsal.I 6-7; Anal iv-v 26-29; Pl 12-13; Ventral i5 (Talwar and Jhingran, 1991).

3.1.3. Molecular identification

In the present work, DNA was extracted from the muscles of *Pangasius pangasius*. A total of fifty samples was collected from five different farming site for identification purpose.



Figure 1. Pangasius pangasius.

Five samples were used to extract DNA, each sample represented a separate selection site. After sampling, DNA extraction was proceeded to extract quality DNA. Extracted DNA was visualized on gel electrophoresis, After the positive result of the polymerase chain reaction, samples were sequenced. PCR products were sent to 1st BASE Laboratory, Malaysia for barcode sequence. Gel documentation of PCR is shown in Figure 2.

3.2. BLAST analysis

As a result of sequencing, a sequenced sample of *P. pangasius* was analyzed, size of the sequence varies from 334 to 369 base pair as; Sequence1 (369 bp), Sequence 2 (364 bp) Sequence 3 (364 bp) Sequence 4 (352 bp) and Sequence 5 (334 bp) (Figure 3-7). Samples were blast on NCBI to identify the specific sequence with its exact match. Blast analysis similarity results match the inquiry specimen as *P. pangasius*. Genbank data record verified that the studied species was *P. pangasius* as similarity index match was found maximum. After confirmation through NCBI blast analysis, these sequences were submitted on Genbank NCBI bankit and the Accession Number of these sequences are OM938452 (Specimen 1), OM938453 (Specimen 2), OM938454 (Specimen 3), OM938455 (Specimen 4) and OM938456 (Specimen 5).

3.3. Pairwise genetic distance analysis

Pairwise genetic distance among species was found maximum of 0.796% between specimens 4 and 5, while a minimum distance was 0.00% between specimens 1 and 2 (Table 1). The evolutionary relationship among *Pangasius pangasius* species was based on the neighborhood joining method. Evolutionary tree analysis showed that close



Figure 2. Gel documentation of extracted DNA (left) and PCR (Right).



relationships with similar species and shared common ancestors, genetically different species were bunched under separate nodes. Specimen 1 showed 99% similarity with specimen 2; 95% among samples 3 and 4, while a separate node for specimen 5 with 4.7 genetic distance (Figure 8).

ACGACGCACTAATTGACCTTCCTGCCCACCAATATTTCCCGCATGATGAAACTTTGGTTCCCTACTAT TATTATGCCTTATAGTACAGACCGCACTACTACCAGCACATTATACCTCAGACATCTTA CTGCCTTCTATCGTGACGCCCACACTCGTCGAGGAGTGTAAATTAGCGATGAGACTTACATG CCAACGGAGCTTCATTCTTTTTCATCTGTATTTACCTACACATCGGACGAGGATTATTATGGCTCTA CTTATTATAAGAAACCTGAAATATTGGGAGTAGTACTTCTCCTATTAGTTATAATAACCCGCTTTCGTCGG ATATGTTTTACCATGGAGCAAA

Figure 3. Identification sequence of *P. pangasius* from Fish farm choki Narool, Kabir Wala, Pakistan (Site 1).

Figure 4. Identification sequence of *P. pangasius* Faheem fish farm, Mumbhal. Kabir Wala, Pakistan (Site 2).

Figure 5. Identification sequence of *P. pangasius* from Nursery unit Fisheries Department Muzaffargarh, Pakistan (Site 3).

CCTTCCTGCCCCATCAATATTTCCCCATGATGAAACTTTGGTTCCCTACTATTATTATGCCTTATAGTA CAGATCCTAACAGGACTTTTCCTAGCCATACATTATACCTCAGACATTCTACTGCCTTCTACTGCCT GCCCACATCTGTCGAGATGTAAATTACGGATGAGTCATCCGCAACTGACATCGAGCTCATT CTTTTTCATCTGTATTTACCTACACACGGACGAGGATTATATTATGGCTCTTACTATAAAGAAAC CTGAAATATTGGGATGATACTTCTCCTATTAGTTATAATAACCGCTTTGGTCGGATATGTTTTACCATG AGGACAA

Figure 6. Identification sequence of *P. pangasius* from Tawakal Fish Hatchery Muzaffargarh, Pakistan (Site 4).

Figure 7. Identification sequence of *P. pangasius* from Al Raheem fish Farm Muzaffargarh, Pakistan (Site 5).

4. Discussion

DNA barcoding is recently under consideration for fish species (Günther et al., 2018). DNA barcoding is essential for identification, evolutionary relationship, and taxonomy. DNA barcoding technology has been adopted as a good approach for species identification (Pereira et al., 2008). This genetic variation and divergence may be due to habitat effects, spatial variations (Doherty, 1991), and physical variables which have a great influence on the fish distribution and other aquatic organisms (Lasram et al., 2010). In the present study identification process was adoptedfor accurate identification of the studied fish (*P. pangasius*)and an evolutionary tree is also constructed for genetic distancing among the species identified.

Meristic or morphometric-based features were used for the identification of species (Musikasinthorn, 2000), Morphometric based methods have many controversies and complexities (Miyan et al., 2014). Variability of morphological characters is often imperfectly recognized for older species descriptions that create taxonomic lineage (Ndiaye et al., 2014). The proper way to resolve taxonomic uncertainties is important to obtain data from both molecular and morphological basis (Padial et al., 2010). Meristic or morphometric features were studied based on different characteristics described by (Bhuiyan, 1964; Rahman, 1989; Talwar and Jhingran, 1991), contrary to morphometric character-based methods, molecular identification method is used for identification purpose.

Variation in the concentration of DNA may be due to considerable variation in the sample size. Time, Temperature, proteinase K, and RNAse concentration

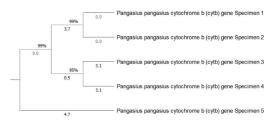


Figure 8. UPGMA method for distancing among the species of *P. pangasius.*

Table 1. Pairwise genetic distance among identified species.

	Specimen 1	Specimen 2	Specimen 3	Specimen 4	Specimen 5
Pangasius pangasius cytochrome b (cytb) gene Specimen 1					
Pangasius pangasius cytochrome b (cytb) gene Specimen 2	0.000000000				
Pangasius pangasius cytochrome b (cytb) gene Specimen 3	0.728021978	0.728021978			
Pangasius pangasius cytochrome b (cytb) gene Specimen 4	0.744318182	0.744318182	0.715909091		
Pangasius pangasius cytochrome b (cytb) gene Specimen 5	0.754491018	0.754491018	0.772455090	0.796407186	

are also considered very important for the high-quality isolation of DNA (Wasko et al., 2003). Similarly, the purity of isolated DNA also relies on numerous factors comprising sample timing (Nielsen et al., 1999) storage of the sample, sample size, and extraction methods (Bauer and Patzelt, 2003). The findings of the present study wereaccording to the need for isolation of quality DNA.

In the present study, a total of 369 bp was used for the identification of species as for identification of species 650 bp was used by (Lohman et al., 2009). Molecular identification is used in animal species as a molecular marker for taxonomy (Luo et al., 2011). These sequences are also used to determine/identify species and divergence among the population. Molecular identification is used to distinguish different animals based on morphological characters. Yet fish species are hard to identify because different species are hard to identify due to the complexities in external morphology and similarity with members of other species (Khan et al., 2013). Genetic or molecular research of the DNA barcoding effectively resolved various taxonomic issues (Barman et al., 2018). Thus, molecular identification can help local fishermen to understand adaptive radiation in fishes in near future.

5. Conclusion

The results of the study revealed that DNA extraction by the phenol chloroform method is suitable for the extraction of quality DNA from muscle tissue for DNA barcoding. DNA barcoding is an effective and precise method for identification purposes. Moreover, GenBank database results revealed similar matches of species as *P. pangasius*. As, Cytochrome b gene was used for identification purposes, hence, *P. pangasius* was successfully identified by using it.

References

- BARMAN, A.S., SINGH, M., SINGH, S.K., SAHA, H., SINGH, Y.J., LAISHRAM, M. and PANDEY, P.K., 2018. DNA Barcoding of freshwater fishes of Indo-Myanmar biodiversity hotspot. *Scientific Reports*, vol. 8, no. 1, pp. 8579. http://dx.doi.org/10.1038/ s41598-018-26976-3. PMid:29872224.
- BAUER, M. and PATZELT, D., 2003. A method for simultaneous RNA and DNA isolation from dried blood and semen stains. *Forensic Science International*, vol. 136, no. 1-3, pp. 76-78. http://dx.doi. org/10.1016/S0379-0738(03)00219-6. PMid: 12969623.
- BHUIYAN, A.L., 1964. Fishes of Dacca. Asiatic Society of Pakistan, vol. 1, no. 3, pp. 90-91.
- DOHERTY, P.J.1991. Spatial and temporal patterns in recruitment. In: P. F.SALE, ed. *The ecology of fishes on coral reefs.San Diego*: Academic Press, pp. 261-293. http://dx.doi.org/10.1016/B978-0-08-092551-6.50015-5.
- ENDO, S., HIBINO, Y. and MOCHIOKA, N., 2022. Identification of first recorded ophichthid larvae of *Ophichthus celebicus* and *O. macrochir* (Anguilliformes; Ophichthidae) from Japan, based on morphometric and genetic evidence.*Ichthyological Research*, vol. 69, no. 3, pp. 393-398. http://dx.doi.org/10.1007/s10228-021-00848-9.
- GÜNTHER, B., KNEBELSBERGER, T., NEUMANN, H., LAAKMANN, S. and MARTÍNEZ ARBIZU, P., 2018. Metabarcoding of marine

environmental DNA based on mitochondrial and nuclear genes.*Scientific Reports*, vol. 8, no. 1, pp. 14822. http://dx.doi. org/10.1038/s41598-018-32917-x. PMid:30287908.

- HUBERT, N., HANNER, R., HOLM, E., MANDRAK, N.E., TAYLOR, E., BURRIDGE, M., WATKINSON, D., DUMONT, P., CURRY, A., BENTZEN, P., ZHANG, J., APRIL, J. and BERNATCHEZ, L., 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One*, vol. 3, no. 6, pp. e2490. http://dx.doi.org/10.1371/ journal.pone.0002490. PMid:22423312.
- KHAN, M.A., MIYAN, K. and KHAN, S., 2013. Morphometric variation of snakehead fish, *Channa punctatus*, populations from three Indian rivers *Journal of Applied Ichthyology*, vol. 29, no. 3, pp. 637-642. http://dx.doi.org/10.1111/j.1439-0426.2012.02058.x.
- KIMURA, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, vol. 16, no. 2, pp. 111-120. http://dx.doi.org/10.1007/BF01731581. PMid:7463489.
- KUMAR, S., STECHER, G., LI, M., KNYAZ, C. and TAMURA, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, vol. 35, no. 6, pp. 1547-1549. http://dx.doi.org/10.1093/molbev/ msy096. PMid:29722887.
- LASRAM, F., GUILHAUMON, F., ALBOUY, C., SOMOT, S., THUILLER, W. and MOUILLOT, D., 2010. The Mediterranean Sea as a 'cul-de-sac' for endemic fishes facing climate change. *Global Change Biology*, vol. 16, no. 12, pp. 3233-3245. http://dx.doi. org/10.1111/j.1365-2486.2010.02224.x.
- LOHMAN, D.J., PRAWIRADILAGA, D.M. and MEIER, R., 2009. Improved COI barcoding primers for Southeast Asian perching. *Molecular Ecology Resources*, vol. 9, no. 1, pp. 37-40. http:// dx.doi.org/10.1111/j.1755-0998.2008.02221.x. PMid:21564563.
- LUO, A., ZHANG, A., HO, S.Y., XU, W., ZHANG, Y., SHI, W., CAMERON, S.L. and ZHU, C., 2011. Potential efficiency of mitochondrial genes for animal DNA barcoding: a case study using eutherian mammals.*BMC Genomics*, vol. 12, no. 1, pp. 1-13. http://dx.doi. org/10.1186/1471-2164-12-84.
- MIYAN, K., KHAN, A.M. and KHAN, S., 2014. Stock structure delineation using variation in otolith chemistry of snakehead, *Channa punctata* (Bloch, 1793), from three Indian rivers. *Journal* of Applied Ichthyology, vol. 5, no. 5, pp. 881-886. http://dx.doi. org/10.1111/jai.12479.
- MUSIKASINTHORN, P., 2000. Channa aurantimaculata, a new channid fish from Assam (Brahmaputra River basin), India, with designation of a neotype for C. amphibious (McClelland, 1845). Ichthyological Research, vol. 1, no. 1, pp. 27-37. http:// dx.doi.org/10.1007/BF02674310.
- NIELSEN, E.E., HANSEN, M.M. and LOESCHCKE, V., 1999. Analysis of applications DNA from old scale samples: technical aspects, and perspectives for conservation.*Hereditas*, vol. 130, no. 3, pp. 265-276. http://dx.doi.org/10.1111/j.1601-5223.1999.00265.x.
- PADIAL, J.M., MIRALLES, A., DE, L., RIVA, I. and VENCES, M., 2010. The integrative future of taxonomy.*Frontiers in Zoology*, vol. 7, no. 1, pp. 16. http://dx.doi.org/10.1186/1742-9994-7-16. PMid:20500846.
- PEREIRA, F., CARNEIRO, J. and AMORIM, A., 2008. Identification of species with DNA-based technology: current progress and challenges.*Recent Patents on DNA & Gene Sequences*, vol. 2, no. 3, pp. 187-199. http://dx.doi.org/10.2174/187221508786241738. PMid:19075956.
- RAHMAN, A.K.A., 1989. Freshwater fishes of Bangladesh.Dhaka: Zoological Society of Bangladesh, 364 p.

- TALWAR, P.K. and JHINGRAN, A.G., 1991. Inland fishes of India and adjacent countries.Rotterdam: Balkema, vol. 2.
- TAYLOR, H.R. and HARRIS, W.E., 2012. An emergent science on the brink of irrelevance: a review of the past 8years of DNAbarcoding. *Molecular Ecology Resources*, vol. 12, no. 3, pp. 377-388. http:// dx.doi.org/10.1111/j.1755-0998.2012.03119.x. PMid:22356472.
- WARD, R.D., ZEMLAK, T.S., INNES, B.H., LAST, P.R. and HEBERT, P.D., 2005. DNA barcoding Australia's fish species. *Philosophical*

Transactions of the Royal Society of London. Series B, Biological Sciences, vol. 360, no. 1462, pp. 1847-1857. http://dx.doi. org/10.1098/rstb.2005.1716. PMid:16214743.

WASKO, A.P., MARTINS, C., OLIVEIRA, C. and FORESTI, F., 2003. Non destructive genetic sampling in fish. An improved method for DNA extraction from fish fins and scales.*Hereditas*, vol. 138, no. 3, pp. 161-165. http://dx.doi.org/10.1034/j.1601-5223.2003.01503.x. PMid:14641478.