Predict Functionally Important Residues Responsible for Estrogen Receptor Subtype Divergence

Feng Wang[1,2] wangfeng@cc.saga-u.ac.jp
Hirohisa Kishino[2,3] kishino@lmb.ab.a.u-tokyo.ac.jp
Yasuhiko Wada[1,2] ywada@cc.saga-u.ac.jp

[1] Faculty of Agriculture, Saga University, Saga 840-8502, Japan
[2] BIRD, Japan Science and Technology Agency, Tokyo 102-8666, Japan
[3] Laboratory of Biometrics and Bioinformatics, Graduate School of Agriculture and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan

Keywords: estrogen receptor, functional divergence, functional residues, structural mapping

1 Introduction

The estrogen receptor (ER) is a ligand-activated transcription factor that mediates the physiological effects of the female sex steroid hormone 17 beta estradiol (E2), and regulates the expression of genes involved in the growth, development and function of a diverge range of tissues. The ER is a member of nuclear receptor (NR) superfamily, which shares a common structural organization including six independent but interacting functional domains. Two ER subtypes, although related, are separate genes and code for proteins of differing lengths. We are interested in the selection constraint acting on the functional domains of ER subtypes after gene duplication. Because the sequence differences of ER subtype in ligand binding domain (LBD) provide the molecular basis for the subtype physiological function, it is necessary to identify those functional related sites in order to investigate the functional divergence between two subtypes. In this study, we utilize the evolutionary rate analysis approach, combined with structural mapping, to predict the functionally important residues responsible for ER subtype divergence.

2 Materials and Methods

The ER, estrogen related receptor (ERR) and other steroid receptors (AR, PR, MR, and GR) amino acid sequences were collected from translated Genbank data bank by the keyword and BLAST search. The amino acid sequences of ER LBD were aligned by CLUSTAL X with the ambiguous positions manually adjusted based on the structural information. We calculate the functional divergence coefficient and predict the type I and II functional sites by site-specific rate shift method, using DIVERGE [1], and likelihood rate tests (LRTs) [2]. Type I sites represent the amino acid conservation in one subfamily, but highly variable in another, implying that these residues have experienced altered functional constraints. Type II sites represent both amino acid conservation in two subfamilies, but with different biochemical properties, implying that these residues may be responsible for functional specification in two subfamilies. This result was also compared with that obtained by another widely used Evolutionary Trace (ET) method [3].

3 Results

The coefficient of functional divergence θ between ER alpha and beta subfamily is 0.254400 ± 0.057323. It is significantly greater than 0, implying that altered functional constraints may take place at some amino acid residues. The posterior probability analysis was then conducted to predict important amino acid residues responsible for altered functional constraints between the two subfamilies. There were sixteen sites among total 246 sites surpassing the cut-off value of posterior probability 0.5.

The residues at these sites were identified as type I functional residues. Thirty-five sites were predicted by the LRTs methods with distinct functional divergence between ER alpha and beta. The designated type I functional sites in this analysis were largely consistent with those determined by DIVERGE. The rest were designated as type II or type I and II mixed functional sites. Five of the total thirty-seven residues at the predicted sites were also supported by other structure or mutation data (344G, 345L, 373H, 383L, 442G, see supplementary material). Then, all the residues at the predicted sites were mapped onto the 3D structure of human ER alpha (PDB file 3ERD, homodimer, Fig.1A). They were diversely distributed on the structure. Also, the functional residues predicted by ET were shown in Fig.1B. Compared with the mapping of ET prediction, there are fewer predicted residues predicted by the site-specific rate method.

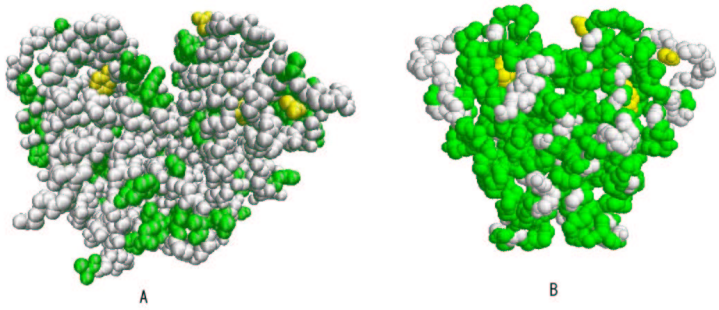

Figure 1: The predicted residues at functional sites were mapped onto the 3D structure of human ER alpha (3ERD), The backbone are in spacefill form with white color, ligands and coactivators with yellow color. The functional residues are shown in green color. A, evolutionary rate method. B. ET method.

4 Discussion

There are no evident clustering or special locations on the 3D structure for the predicted functional divergence residues using the site-specific rate analysis. This is different from the widely used Evolutionary Trace method, where the structure cluster on the structure is the criterion to identify the functional residues [3]. The difference lies in that ET utilize the extremely conserved amino acids in both subfamilies. Instead, the evolutionary rate shift method only covers the functional residues relating to functional divergence between two subfamilies. It seems reasonable to understand there is no evident cluster residues on the structure for the predicted residues related to functional divergence.

However, the dispersed functional residues on the structure indicated that, any position in a protein could be important for the overall function [4]. To some extent, the structure location of functional residues makes it difficult for prediction. The possibility of neutral substitutions at the amino acid level indeed exists, evenly with distinct rate shift between subfamilies. Because the structure folds are more conservative than sequence variation, therefore, at some positions, the residue substitutions can be tolerated without causing functional changes. Therefore, it is not enough to distinguish the adaptive mutation (function-related substitution) from the neutral one, when only considering the evolutionary rate shift at amino acid sites. Integration of more structure and sequence information is needed to perform the accurate function residue prediction.

References

[1] Gu, X. and Vander Velden K., DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family, *Bioinformatics*, 18(3):500–501, 2002.

[2] Knudsen, B., Miyamoto, M.M., Laipis, P.J., and Silverman, D.N., Using evolutionary rates to investigate protein functional divergence and conservation: a case study of the carbonic anhydrases, *Genetics*, 164(4): 1261–1269, 2003.

[3] Lichtarge, O., Bourne, H.R., and Cohen, F.E., An evolutionary trace method defines binding surfaces common to protein families, *J. Mol. Biol.*, 257(2):342–358, 1996.

[4] Philippe, H. and Lopez, P., On the conservation of protein sequences in evolution, *Trends Biochem. Sci.*, 26(7):414–416, 2001.