CHAPTER 4

STUDIES ON THE COEVOLUTION OF ORNITHINE DECARBOXYLASE AND ANTIZYME

Studies on the coevolution of ornithine decarboxylase and antizyme.

4.1 Introduction

Under this objective the degree to which the two proteins ODC and antizyme have coevolved is being investigated *in silico*. Ornithine decarboxylase (ODC) is a key enzyme in the biosynthesis of polyamines. Antizyme (Az) regulates the levels of polyamines in cells by targeting ODC for degradation. The interaction between ODC and antizyme is important for the regulation of the levels of polyamines in the cells. Hence, the question is these two proteins ODC and antizyme coevolved to what extent to maintain their interaction?

4.2 Hypothesis of the study

The interaction between ODC and Az is important for the regulation of the levels of polyamines in cells. Hence, these two proteins co-evolve.

We are assuming that

(i) The ODC and Az interaction interface is ‘same’ in all organisms,

(ii) Neither ODC nor Az uses their mutual interaction interface to interact with any other protein. If so, even mutations in the other interacting partner will influence the interface

(iii) ODCs and Azs that we are considering for analysis constitute homologous families.

4.3 Methodology

4.3.1 Collect the sequences of ODC and antizyme for several organisms

The fasta sequences of ODC and antizyme for several organisms (cover human to yeast) were collected from UniProt database. Only those sequences were collected which have citation or referred. Fasta sequences of ODCs and Azs that were considering for analysis correspond to homologous families.

4.3.2 Performed multiple sequence alignment

Multiple sequence alignment (MSA) was done for full length ODC and antizyme proteins in ClutalW by bootstrap method (replications 500). Human ODC and antizyme sequences were taken as references in MSA. MEGA6 software was used for editing.

4.3.3 Pairwise distance analysis

The pairwise distance was calculated for ODC and Az MSA results in MEGA6 software. The parameters were used as Bootstrap method (500 replications), variance method with Dayhoff substitution model.

4.3.4 Phylogenetic tree

The phylegenetic tree was generated in MEGA6 software by using the neighbor-joining method with 500 replications of bootstrap.

4.3.5 Correlation co-efficient

The Pearson Correlation coefficient was calculated for to correlate pairwise distance matrices of ODC and Az sequences.

Pearson Correlation Coefficient Formula:

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)}}$$

Where:

x_i: the distance value for ODC
y_i: the distance value for Az
n: total number of distance value of ODC or Az
r: correlation coefficient, $-1 <= r <= 1$
 1 represents strongly positively correlation,
 -1 represents strongly negatively correlation,
 0 represents no correlation.

Table 4.1 List of tools used in this study

Analysis	Software/Server/Database	References
Sequences of ODC and Az proteins	Universal Protein Resource (UniProt) database	(Apweiler et al., 2004; UniProtConsortium, 2014)
Multiple Sequence Alignment (MSA)	Clustal Omega/W	(Sievers et al., 2011)
Phylogenetic tree	The neighbor-joining method	(Saitou and Nei, 1987)
Compute pair wise distance matrices	MEGA 6	(Tamura et al., 2007)
Correlation between ODC and Az	http://www.endmemo.com/statistics/cc.php	The Pearson Correlation coefficient
Domain identification	Pfam28	(Finn et al., 2013)
Substitution mutation analysis	Blousm62	

4.4 Strategy to find co-evolution between ODC and antizyme proteins

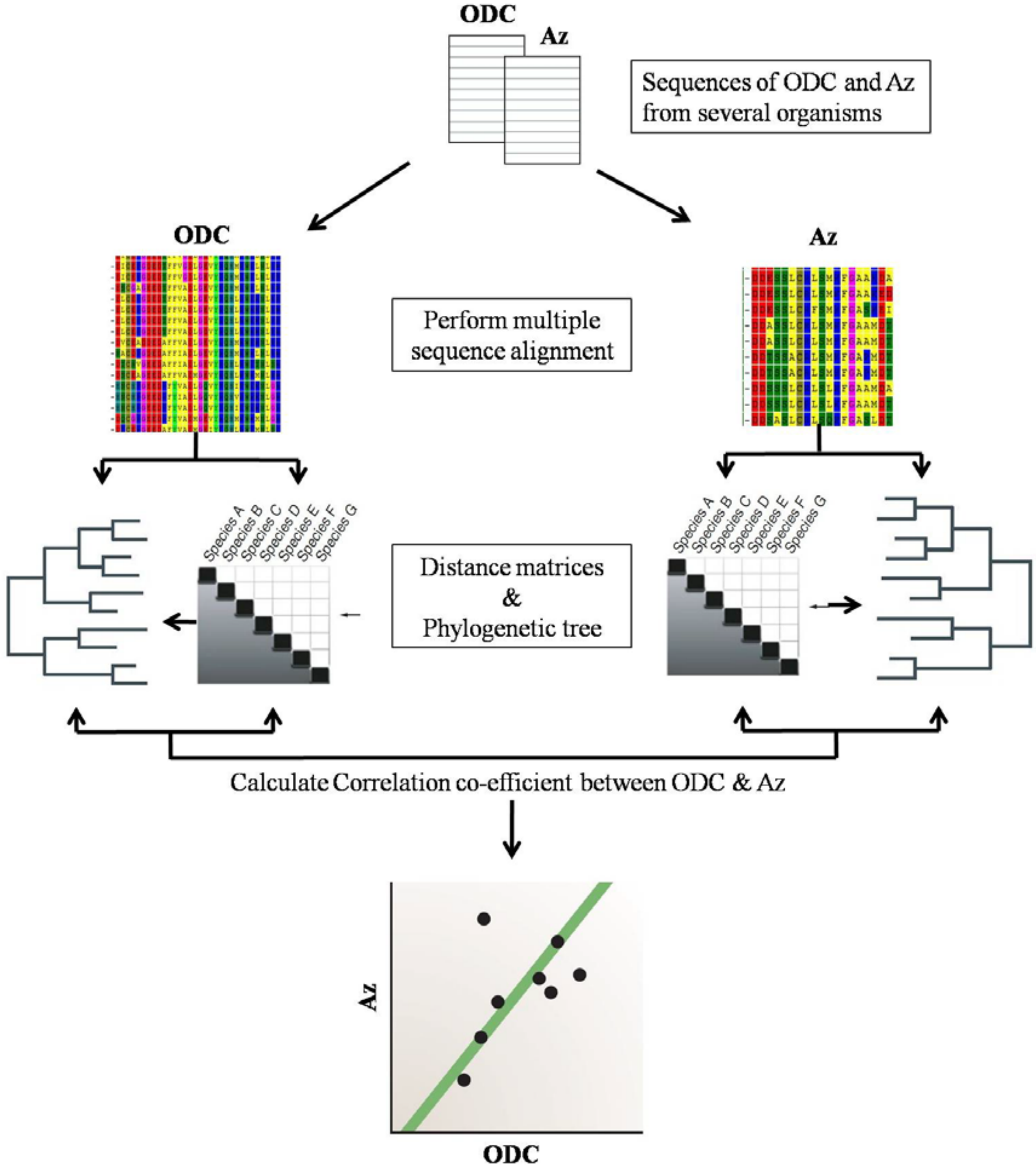


Figure 4.1 Schematic representation of strategy to find evolutionary correlation between ODC and Antizyme proteins

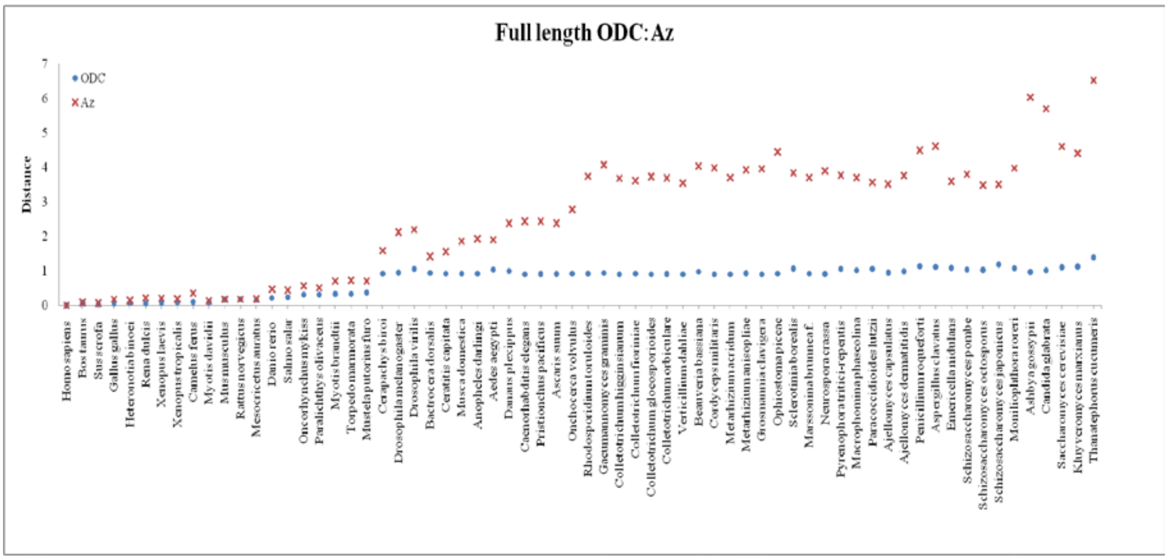
4.5 Results and discussion

4.5.1 Correlation between full length ODC and antizyme protein sequences

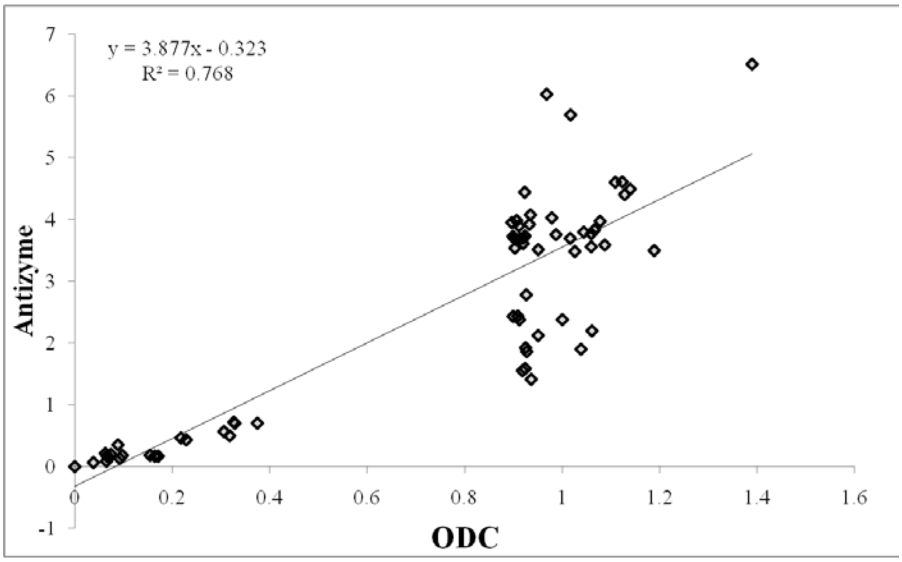
To check the correlation between ODC and antizyme proteins through evolution, Fasta sequences of ODC and antizyme for several organisms (cover human to yeast) were collected from UniProt database. Only those sequences were collected which have citation or referred. Fasta sequences of ODCs and Azs that were considering for analysis correspond to homologous families.

Multiple sequence alignment (MSA) was done for full length ODC and antizyme proteins in ClutalW (Appendix I). Human ODC and antizyme sequences were taken as references in MSA. MSA Output for ODC and antizyme were submitted to calculate pair wise distance. Bootstrap method (500 replications) variance method with Dayhoff substitution model was used to compute pair wise distance matrices.

Difference between selected organism sequence with respect to reference organism sequence is expressed as distance. Distance matrix is pairwise information about the evolutionary divergence of selected sequences. A value in matrix is equivalent to % of divergence of two homologous organisms’ protein pair.



(A)

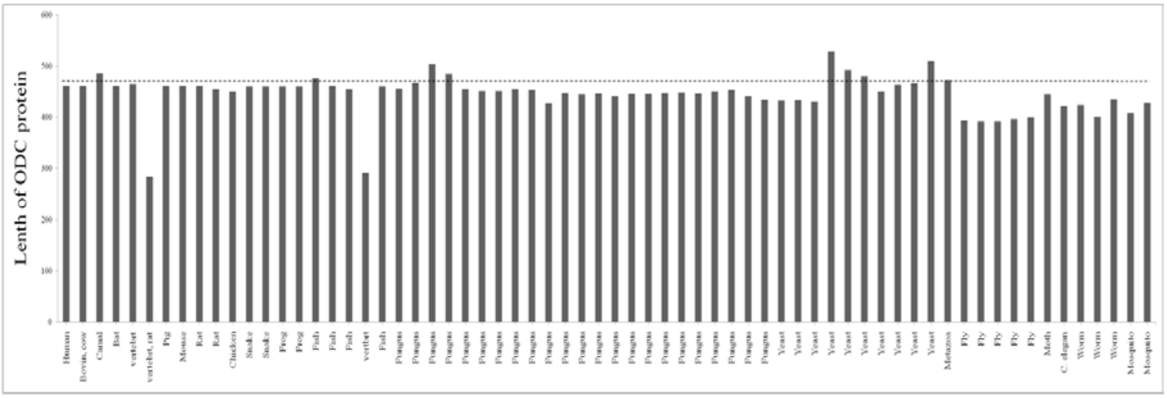


(B)

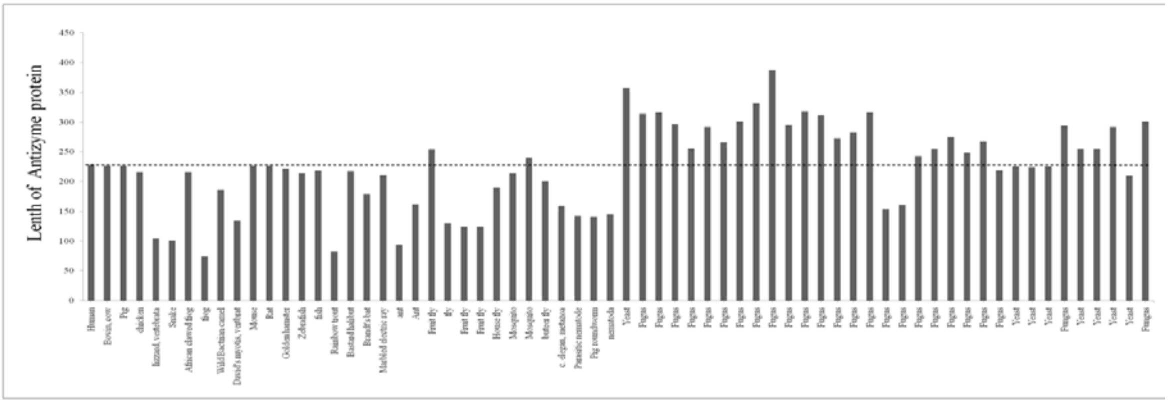
Figure 4.2 Distance calculation for full length ODC and Antizyme protein sequences of several organisms with respect to human.

(A) Comparison of distance for ODC and Antizyme protein for each organism
(B) XY plot for ODC vs Az distance value

Initially, distances were calculated for full length ODC and antizyme protein sequences of selected organisms. Human ODC and antizyme protein sequences were taken as references. Correlation coefficient for pairwise distance matrices of ODC and Az was **0.23.** There was poor correlation between selected ODC and Az sequences. All the other ODC:Az distance values correlate well with each other excluding fungus and yeast (Fig. 4.2). It could be due to the large number of insertions in fungus and yeast Az sequences compared to human Az sequence (Appendix II). There was no large difference in length of ODC sequences for all selected organism with respect to human. (Fig. 4.3A) However, in case of antizyme, there was variation in length (Fig. 4.3B).



(A)

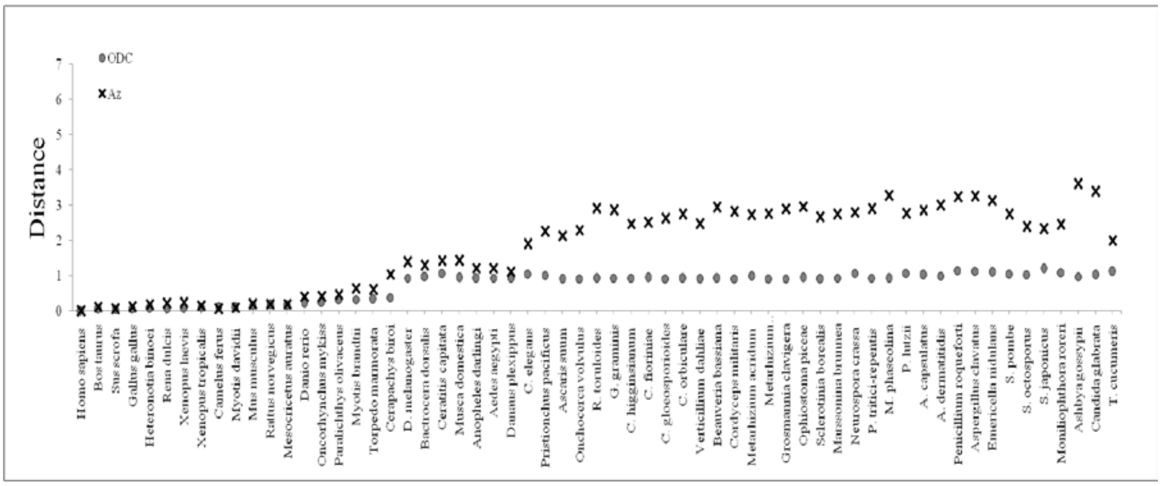


(B)

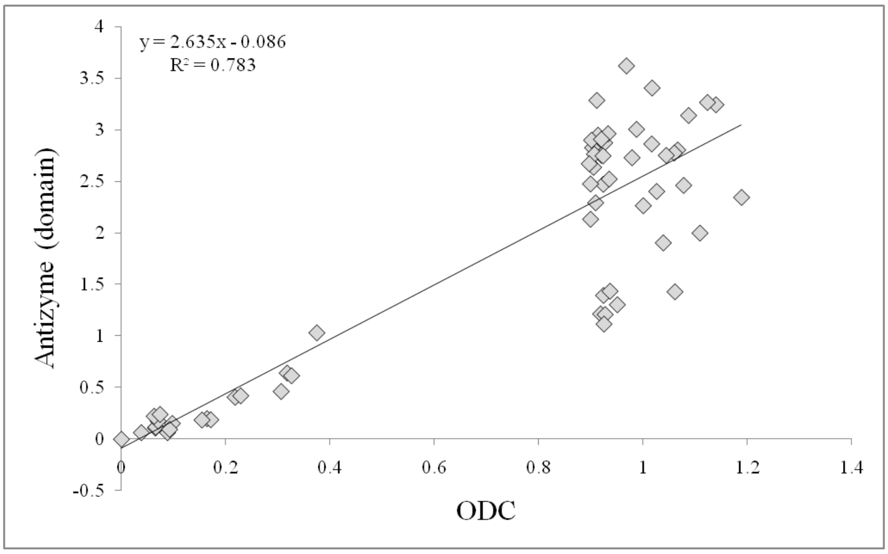
Figure 4.3 Compare length of ODC (A) and Antizyme (B) sequence for several organisms with respect to human

The next step was focused only on domain region of antizyme. Domain region of all Az sequences of selected organisms were identified by using Pfam domain database. Multiple sequence alignment was carried out for Az domain region sequences. Again same strategy was used to calculate correlation coefficient (Fig. 4.1).

Similar results were obtained. Correlation coefficient for pairwise distance matrices of ODC and Az(domain) was **0.36.** There was poor correlation between selected ODC and Az(domain) sequences. But compared to last observation (Fig. 4.2), better correlation was seen between ODC and Az(domain) (Fig. 4.4)



(A)

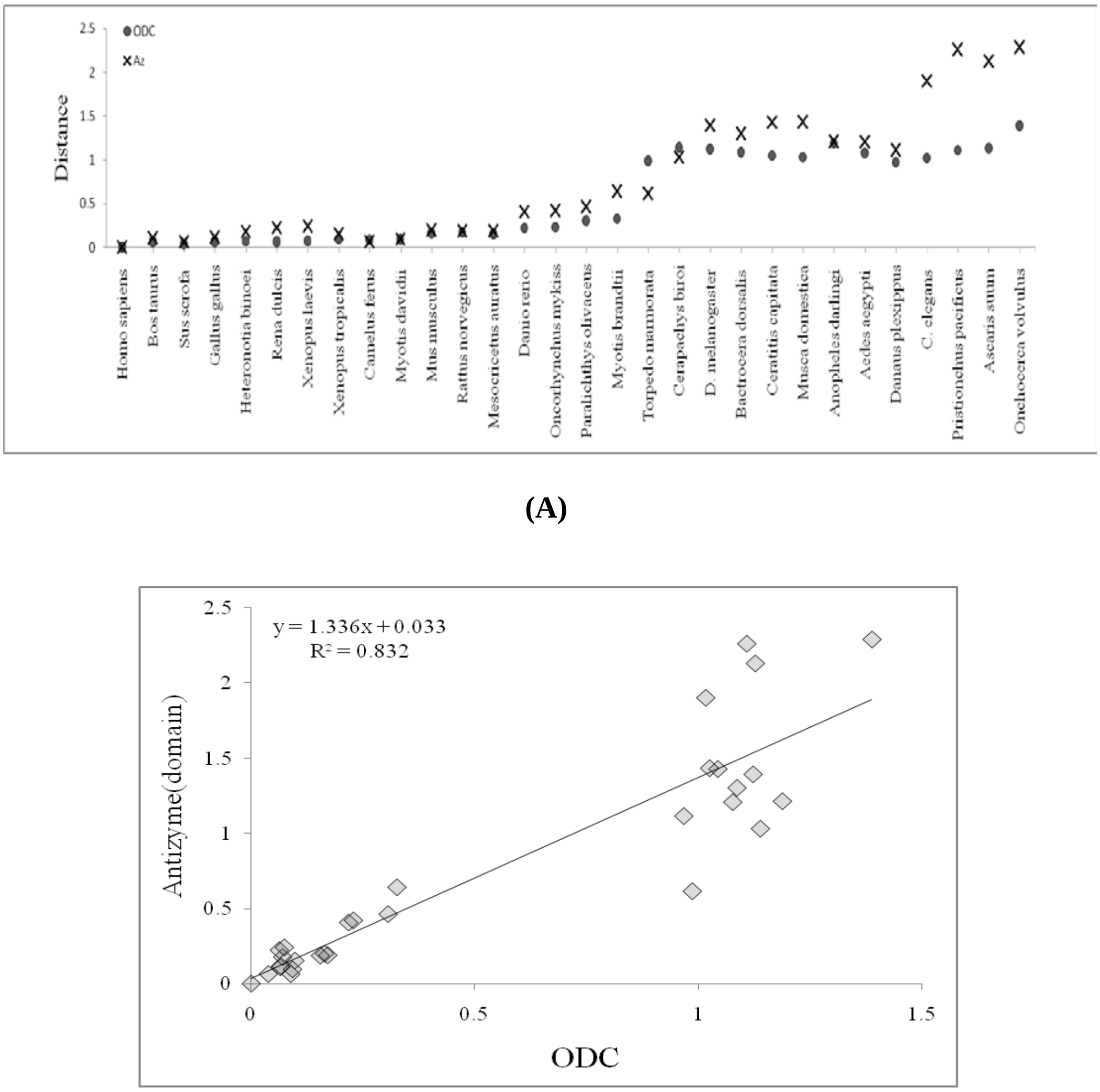


(B)

Figure 4.4 Distance calculation for ODC and Antizyme (domain) protein sequences of several organisms with respect to human.

(A) Comparison of distance for ODC and Antizyme (domain) protein for each organism
(B) XY plot for ODC vs Az(domain) distance value

Finally we have excluded fungus and yeast sequences of ODC and Az from our input data. The data was divided in two parts for analysis as (i) Higher eukaryotes (ii) Fungus and yeast.



(A)

(B)

Figure 4.5 Distance calculation for higher organisms ODC and Antizyme (domain) protein sequences with respect to human.

(A) Comparison of distance for ODC and Antizyme (domain) protein for each organism
(B) XY plot for ODC vs Az(domain) distance value

In higher eukaryotic distance values of ODC:Az correlate with each other (Fig. 4.5). Correlation coefficient with sequences of higher organisms for ODC and antizyme (domain) protein with respect to human was **0.91.**

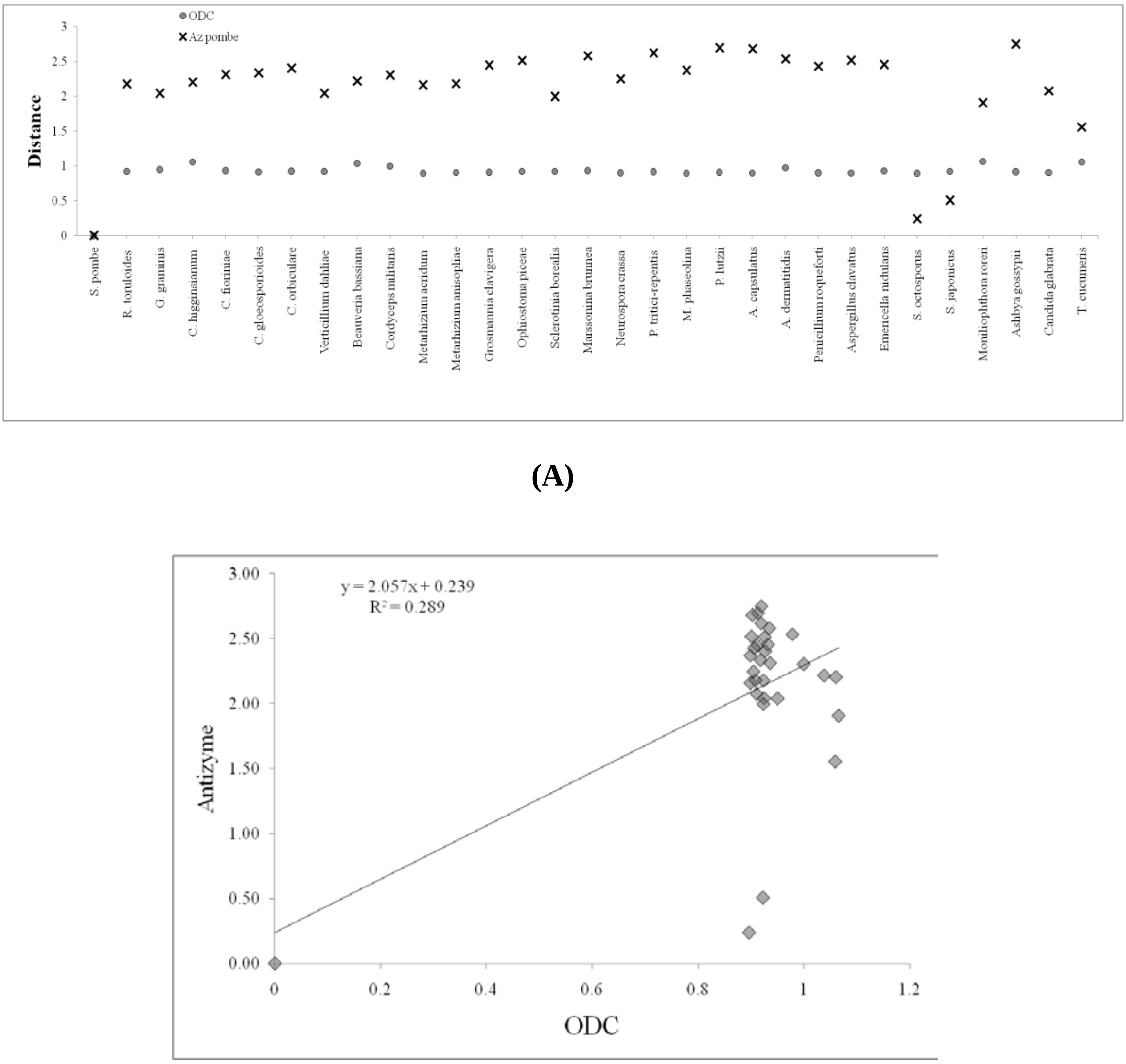


Figure 4.6 Distance calculation for fungus and yeast ODC and Antizyme (domain) protein sequences with respect to *S.pombe*.

(A) Comparison of distance for ODC and Antizyme (domain) protein for each organism
(B) XY plot for ODC vs Az(domain) distance value

In lower eukaryotic (fungus and yeast) distance values of ODC:Az correlate are shown in Fig. 4.6. Correlation coefficient of lower eukaryotes sequences for ODC and antizyme (domain) protein with respect to *S.pombe* was **(-)0.23.**

In higher eukaryotes ODC and antizyme correlate with each other. There was no correlation between ODC and Az in lower eukaryotes (fungus and yeast). It could be due to the fact that fungus and yeast diverged from the main line of evolution earlier than human.