Comparative Genomics Revealed a Novel DNA-Binding Regulatory Protein Involved in Homologous Recombination in Bacteria

Yang Gao
Computer Network Information Center,
Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College,
Beijing, 100005 China

Yan Zhang*
Key Laboratory of Systems Biology,
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences,
Shanghai, 200031 China
* Corresponding author, Tel: 86-21-54921124; Fax: 86-21-54920061; E-mail: yanzhang01@sibs.ac.cn

Abstract—**Homologous recombination is a fundamental cellular process that is most widely used by cells to rearrange genes and accurately repair DNA double-strand breaks. It may result in the formation of a critical intermediate named Holliday junction, which is a four-way DNA junction and needs to be resolved to allow chromosome segregation. Different Holliday junction resolution systems and enzymes have been characterized from all three domains of life. In bacteria, the RuvABC complex is the most important resolution system. In this study, we conducted comparative genomics studies to identify a novel DNA-binding protein, YebC, which may serve as a key regulator of RuvABC resolvasome. On the other hand, the presence of YebC orthologs in some organisms lacking RuvC implied that it might participate in other biological processes. Further phylogenetic analysis of YebC protein sequences revealed two functionally different subtypes of this family: YebC_I and YebC_II. Only YebC_I subgroup may play an important role in regulating RuvABC gene expression in bacteria. Investigation of YebC-like proteins in eukaryotes suggested that they may have originated from YebC_II proteins and evolved a new function as a specific translational activator in mitochondria. Finally, additional phylum-specific genes associated with Holliday junction resolution were predicted. Overall, this study provides new insight into the basic mechanism of Holliday junction resolution and homologous recombination in bacteria.**

Keywords- comparative genomics; homologous recombination; Holliday junction; bacteria

I. INTRODUCTION

Homologous recombination is a fundamental mechanism in biology that rearranges genes within and between chromosomes, promotes DNA repair, and guides segregation of chromosomes at division. This process is common to all forms of life and involves the exchange (i.e., breakage and reunion) of DNA sequences between two chromosomes or DNA molecules (1-4). Such exchanges contribute to the generation of genetic diversity and the conservation of genetic identity.

Although homologous recombination varies widely among different organisms and cell types, most forms of it involve the same basic steps: (i) after a DNA break occurs, sections of DNA around the break on the 5' end of the damaged chromosome are removed in a process called resection; (ii) in the strand invasion step that follows, an overhanging 3' end of the damaged chromosome then "invades" an undamaged homologous chromosome; (iii) after strand invasion, one or two cross-shaped structures (called Holliday junctions) are formed to connect the two chromosomes. Holliday junction (or four-way junction) has been generally assumed as a key intermediate in genetic recombination and DNA repair since its discovery in 1964 (5). They are highly conserved structures from prokaryotes to mammals, which adjoin two DNA duplexes, forming a branch point where four helices are interconnected by strand exchange (6).

Because Holliday junctions provide a covalent linkage between chromosomes, their efficient resolution is essential for proper chromosome segregation. Enzymes that resolve Holliday junctions by endonucleolytic cleavage have been isolated from bacteriophages, bacteria, archaea and certain eukaryotes (7-10). In *Escherichia coli*, the enzymes that are involved in resolution of Holliday junction include RuvABC, RecG, and RusA (11-13). The RuvABC proteins (or RuvABC resolvasome) constitute a simple and the most widely used system for the processing of Holliday junctions. RuvAB proteins catalyze the branch migration whereas RuvC endonuclease resolves the Holliday junction into duplex products (13). The RecG protein is a DNA helicase and may promote branch migration of a variety of branched DNAs including Holliday junctions (14, 15). The RusA protein is a homodimeric Holliday junction-specific endonuclease and can bind a variety of branched DNA structures (16, 17). RecG may be required by RusA to branch migrate Holliday junctions to cleavable sequences (7). However, the precise mechanism regulating the activities of these enzymes is unknown and the factors involved remain unidentified.

Homologs of RuvABC, RecG, and RusA are absent from almost all sequenced archaea and eukaryotes. In archaea, the Hjc protein, a distantly related member of the type II

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

restriction endonuclease family, has been characterized to serve as a Holliday junction resolving enzyme (18, 19). Little is known about the mechanism of eukaryotic Holliday junction resolution and the enzymes involved. It was reported that Saccharomyces cerevisiae contains a Holliday junction resolvase Cce1 (20, 21), an equivalent enzyme from Schizosaccharomyces pombe (named Ydc2) has also been found (22). However, these enzymes are targeted to the mitochondria, suggesting that they can only cleave junctions formed during recombination of mitochondrial DNA. Very recently, a nuclear Holliday junction resolvase was first identified from both humans and yeast (23). These resolvases (GEN1 in human and its yeast ortholog Yen1) represent a new subclass of the the Rad2/XPG family of nucleases, and promote Holliday junction resolution in a manner similar to that shown by the *E. coli* RuvC (24, 25).

In this study, we carried out comparative genomics approaches to investigate the mechanisms of Holliday junction resolution in prokaryotes. Occurrence of known components (e.g., RuvABC) could be easily identified by comparative genomics. Our analysis also generated evidence for a novel DNA-binding regulatory protein family involved in Holliday junction resolution in bacteria. Homologs of this family were detected in a variety of eukaryotes and are predicted to be localized in mitochondria. Overall, these data provide new insight for better understanding the basic mechanism of homologous recombination in nature.

II. Materials and Methods

A. Genomes, sequences and resources

Fully sequenced prokaryotic and eukaryotic genomes from the Entrez Genome Database were used in this study (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). Because of the large number of strains for some bacterial species, only one strain was selected for each species. A total of 786 bacteria, 51 archaea and 170 eukaryotes were analyzed.

We used *E. coli* RuvA (COG0632, Holliday junction resolvasome DNA-binding subunit), RuvB (COG2255, Holliday junction resolvasome helicase subunit) and RuvC (COG0817, Holliday junction resolvasome endonuclease subunit) sequences as queries to search for RuvABC-dependent Holliday junction resolution trait. For each of these proteins, TBLASTN (26) was initially used to identify genes coding for homologs with a cutoff of E-value ≤0.1. Orthologous proteins were then defined as bidirectional best hits (27). When necessary, orthologs were also confirmed by genomic location analysis or building phylogenetic trees for the corresponding protein families.

The STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database and programs (http://string-db.org/) were used to identify gene candidates that may be functionally related to RuvABC resolvasome. Different parameters were used for better performance.

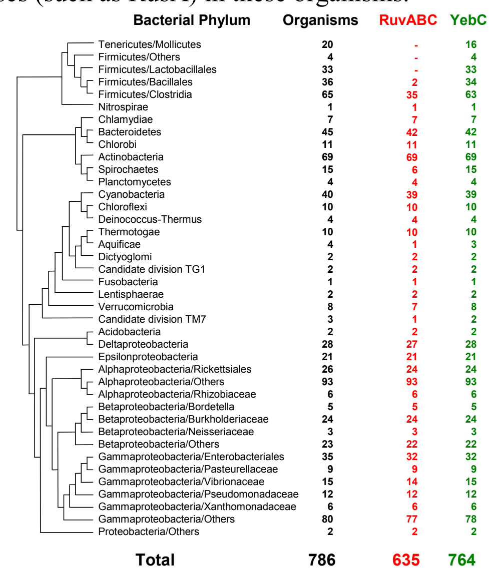
B. Multiple sequence alignment and phylogenetic analysis

Sequence alignments were performed with CLUSTALW (28) using default parameters. Ambiguous alignments in highly variable (gap-rich) regions were excluded. The resulting multiple alignments were then checked for conservation of residues and manually edited. Phylogenetic analyses were performed using PHYLIP programs (29). Pairwise distance matrices were calculated by PROTDIST to estimate the expected amino acid replacements per position. Neighbor-joining (NJ) trees were obtained with NEIGHBOR and the most parsimonious trees were determined with PROTPARS.

III. Results and Discussion

A. Distribution of RuvABC-dependent Holliday junction resolution trait in bacteria

Except a very small number of organisms (less than 10) with small and condensed genomes (mostly parasites), all sequenced bacteria contain RuvA and RuvB genes. As RuvAB complex may catalyze both Holliday junction branch migration and replication fork reversal (30, 31), the occurrence of their genes may not precisely reflect the Holliday junction resolution trait. Thus, we used the co-occurrence of RuvABC as a gene signature for the presence of RuvABC-dependent Holliday junction resolution trait. Sequence analysis of bacterial genomes revealed a wide distribution of RuvABC resolvasome. We identified 635 (80.8% of all sequenced genomes) organisms that contain this system. All RuvC-containing organisms have RuvA and RuvB, many of which have RuvABC genes in the same operon. Fig. 1 shows the distribution of RuvABC system in different bacterial taxa based on a highly resolved phylogenetic tree of life developed by Ciccarelli and coworkers (32). Although RuvC is the most widely used resolvase in bacteria, the absence of this gene in ~20% bacterial genomes suggested the presence of alternative resolvases (such as RusA) in these organisms.



978-1-4577-1666-9/11/$26.00 ©2011 IEEE

Figure 1. Distribution of RuvABC system and YebC family in different bacterial taxa.

B. Identification of a new family involved in Holliday junction resolution in bacteria

Since the RuvABC complex has been shown to be the most important system for the resolution of Holliday junctions in bacteria, identification of functional linkages involving RuvABC (especially RuvC which is specific for Holliday junction resolution) may help understand the details of this process. We used STRING web server to examine such functional linkages based on neighborhood, gene fusion and co-occurrence analyses. The top candidates for RuvA, RuvB or RuvC are shown in Table 1.

Table 1. STRING analysis of genes functionally associated with RuvABC resolvasome.

Rank	RuvA	RuvB	RuvC
1	RuvB	RuvA	RuvB
2	RuvC	RuvC	RuvA
3	YebC	YebC	YebC
4	YbgC	QueA	YbgC
5	TolB	YbgC	YeeN
6	QueA	TolB	TolB
7	FolC	PanB	CysS
8	MaeB	TolQ	TolQ
9	YeeN	TolR	QueA
10	TolQ	YjeS	PurH

Except the components of RuvABC, the protein with the best score was YebC, a putative cytoplasmic protein of unknown function (COG0217, uncharacterized conserved protein; pfam01709, domain of unknown function DUF28). This gene was located very close to or even in the same operon with RuvC in many bacteria. Moreover, YebC and RuvABC showed similar patterns of occurrence in most bacterial phyla based on the STRING output. The next predicted RuvABC link was YbgC, a bacterial 4-hydroxybenzoyl-CoA thioesterase involved in phospholipid metabolism and is also associated with the Tol-Pal system (33, 34). Most of other candidates predicted by STRING are also involved in Tol-Pal system. It has been known that this system is important for cell envelope integrity and is part of the cell division machinery. In *E. coli*, the Tol-Pal system is composed of the YbgC, TolQ, TolA, TolR, TolB, Pal and YbgF proteins. It is unclear whether some of these proteins are involved in DNA repair and recombination. Here, we only focus on the YebC family.

Since YebC might be functionally associated with RuvABC, we further analyzed the distribution of this family in all sequenced prokaryotes. Homologs of YebC were not detected in archaea, implying that YebC may either have evolved in bacteria or lost in the ancestors of archaea (more likely). In bacteria, the distribution of YebC appeared to be wider than RuvC (Fig. 1), suggesting that YebC may be also involved in other processes independent of RuvC. However, the facts that all RuvC-containing organisms have YebC, and that YebC is always a neighboring gene of RuvC in approximately half of the RuvC-containing organisms, indicate a strong relationship between the two genes. These results are consistent with a previous analysis of some "hypothetical" genes expressed in *Haemophilus influenzae*, which also suggested a potential association of YebC with RuvABC in this organism (35).

The majority of YebC-containing bacteria (93.5%) have single copy of this gene. Intriguingly, in the organisms that have more than one YebC homologs, there is always one protein whose gene is localized very close to either RuvC or RuvAB (when RuvC is absent) genes, implying that YebC proteins may have different subgroups. Phylogenetic analysis of YebC proteins from sequenced bacteria revealed that the YebC family could be divided into two subtypes: YebC_I and YebC_II (Fig. 2).

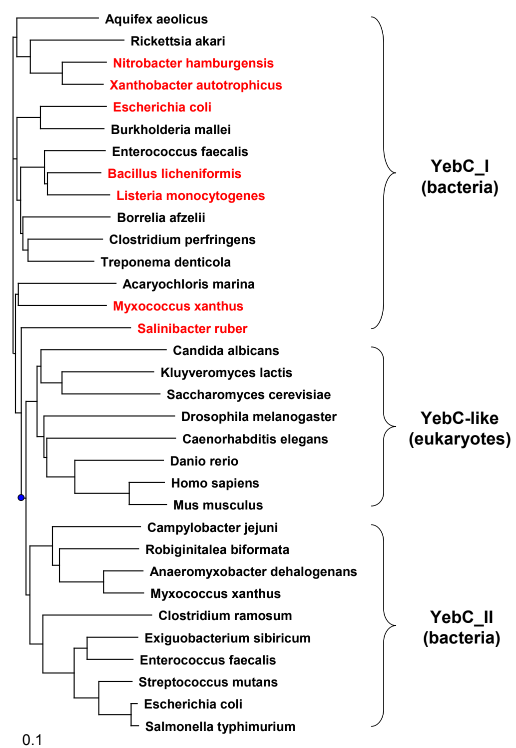


Figure 2. Phylogenetic analysis of YebC family. Organisms where YebC genes are localized very close to RuvABC genes are shown in red. The root node of the tree is shown as a blue dot. Separate branches for the two subtypes of YebC in bacteria and the eukaryotic YebC-like proteins are also shown.

Only members of YebC_I subgroup were found to be localized very close to the genes encoding RuvC. In contrast, there is no evidence that YebC_II subgroup could be involved in Holliday junction resolution, even though only YebC_II members were observed in most Bacteroidetes that have RuvABC resolvasome. Thus, it appeared that YebC_I proteins are functionally associated with RuvABC system. YebC_II might have evolved from YebC_I proteins with novel function. Multiple alignment of YebC_I and YebC_II sequences suggested several specific residues which are only

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

present in each subfamily (Fig. 3). An attractive hypothesis is that YebC_I is functionally different from YebC_II, perhaps distinguished by some of these conserved residues.

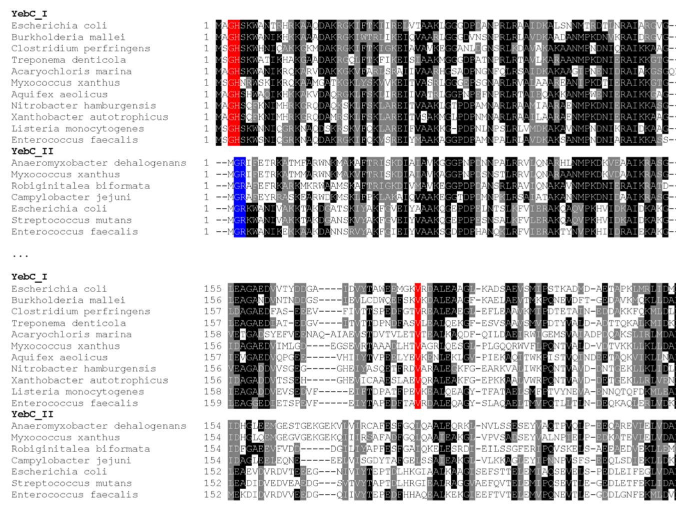


Figure 3. Multiple sequence alignment of YebC proteins in bacteria. Representative sequences were divided into YebC_I and YebC_II subgroups. Residues which are strictly conserved in the YebC_I subgroup are shown in red background. Residues which are strictly conserved in the YebC_II subgroup are shown in blue. Other residues shown in white on black or grey are conserved in homologs.

Although YebC is a large family of widespread conserved proteins whose function is unknown, this group of proteins has been extensively characterized from the structural perspective. To date, the crystal structures of YebC proteins from *Aquifex aeolicus* (YebC_I group), *E. coli* (YebC_I group), and *Helicobacter pylori* (YebC_II group) have been solved (PDB ID codes 1LFP, 1KON, and 1MW7, respectively). A previous structural analysis of *A. aeolicus* YebC revealed a large cavity with a predominance of negatively charged residues on the surface of this protein (36). Interestingly, all three structure-solved proteins have a putative DNA binding function. A recent study claimed that the YebC protein in *Pseudomonas aeruginosa* (PA0964, YebC_I subgroup) may serve as a potential transcription regulator, which is involved in negatively regulating the quorum-sensing response regulator pqsR of the PQS system by binding at its promoter region (37). This result implied the complexity of the function of YebC proteins in nature.

Although the functions of YebC proteins and the biological pathways they are involved in are unclear, our studies provide some useful information about this family: (i) both YebC_I and YebC_II subgroups may bind DNA; (ii) YebC_I proteins may serve as a multi-functional transcription regulator mainly involved in regulating the expression of RuvABC as well as other genes such as pqsR; (iii) YebC_II might have evolved from YebC_I by gene duplication and have novel function independent of Holliday junction resolution or even DNA recombination. A future challenge would be to understand the DNA binding patterns of YebC_I proteins as well as additional processes they may regulate.

C. Investigation of YebC-like proteins in eukaryotes

Significant YebC homologs were also detected in a variety of eukaryotes, including fungi, plants and animals. Very recently, it was reported that a mutation in the human gene encoding a YebC homolog (named CCDC44, localized to the mitochondria) led to a specific defect in the synthesis of the mitochondrial DNA-encoded cytochrome c oxidase subunit I (COX I) (38). Thus, the human CCDC44 protein was renamed as TACO1, which may serve as a mammalian mitochondrial translational activator of COX I. Possible mechanisms of TACO1 action to ensure translation of COX 1 were also considered: (i) securing an accurate start of translation; (ii) stabilizing the elongating polypeptide; and (iii) interacting with the peptide release factor (38, 39).

We analyzed the sequences of all eukaryotic YebC-like proteins and the relationship with their bacterial counterparts. All detected YebC-like proteins in eukaryotes have mitochondrial signal sequences, suggesting that they are mitochondria-targeted proteins. Phylogenetic analysis of both bacterial and eukaryotic YebC proteins showed that the eukaryotic YebC-like proteins were clustered with YebC_II subfamilies (Fig. 2), implying that eukaryotic YebC-like proteins (including human TACO1) might have evolved from ancient YebC_II proteins. The mitochondrial signal sequences were then added to target them into the mitochondria as a specific translational activator, at least in metazoan mitochondrial genome. As eukaryotes lack the RuvABC resolvasome, it is unclear whether these YebC-like proteins are involved in homologous recombination in mitochondria, or whether they still have the capacity to bind mitochondrial DNA. Further studies are required to determine the substrates and function of YebC-like proteins in other organisms as well as their relationship with DNA repair and recombination in mitochondria.

D. Identification of additional phylum-specific genes that might be associated with Holliday junction resolution

Comparative genomics studies also suggested additional candidate genes involved in Holliday junction resolution in certain bacterial phyla. For example, in Firmicutes/Clostridia, most organisms possess a conserved hypothetical protein (CTC02214 in *Clostridium tetani*) whose gene is always located next to either YebC or RuvC gene, implying a potential functional link with them. However, orthologs of this protein family were exclusively detected in Clostridia, suggesting that this protein might be newly evolved in this phylum. Similarly, another conserved hypothetical protein (DUF208 super family; COG1636, uncharacterized protein conserved in bacteria) was also identified in a variety of distantly related organisms where its gene is often located close to either YebC or RuvABC genes. Further studies, however, are needed to verify their function and the relationship between these genes and genetic recombination.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

IV. CONCLUSION

In this study, we carried out comparative genomics to identify a novel DNA-binding protein family, YebC, which was strongly linked to Holliday junction resolution in bacteria. Phylogenetic analysis revealed that YebC might be divided into two functional subgroups: YebC_I and YebC_II. The YebC_I subfamily might be a multi-functional transcriptional regulator that mainly regulates the expression of RuvABC resolvasome in bacteria. It cannot be excluded that YebC_II subfamily is involved in homologous recombination, but current evidence does not provide strong support for this possibility. Further studies on eukaryotic YebC-like proteins suggested that they may have evolved from YebC_II subgroup and have different functions to serve as a specific translational activator in mitochondria.

REFERENCES

[1] A. Shinohara and T. Ogawa, “Homologous recombination and the roles of double-strand breaks,” Trends Biochem. Sci. vol. 20, pp. 387-391, October 1995.

[2] R. D. Camerini-Otero and P. Hsieh, “Homologous recombination proteins in prokaryotes and eukaryotes,” Annu. Rev. Genet. vol. 29, pp. 509-552, 1995.

[3] A. Kuzminov, “DNA replication meets genetic exchange: chromosomal damage and its repair by homologous recombination,” Proc. Natl. Acad. Sci. U.S.A. vol. 98, pp. 8461-8468, July 2001.

[4] J. Yanowitz, “Meiosis: making a break for it,” Curr. Opin. Cell. Biol. vol. 22, pp. 744-751, December 2010.

[5] R. Holliday, “A mechanism for gene conversion in fungi,” Genet. Res. vol. 5, pp. 282–304, 1964.

[6] A. Constantinou, A. A. Davies, and S. C. West, “Branch migration and Holliday junction resolution catalyzed by activities from mammalian cells,” Cell. vol. 104, pp. 259-268, January 2001.

[7] G. J. Sharples, S. M. Ingleston, and R. G. Lloyd, “Holliday junction processing in bacteria: insights from the evolutionary conservation of RuvABC, RecG, and RusA,” J. Bacteriol. vol. 181, pp. 5543-5550, September 1999.

[8] M. F. White, “Homologous recombination in the archaea: the means justify the ends,” Biochem. Soc. Trans. vol. 39, pp. 15-19, January 2011.

[9] G. J. Sharples, “The X philes: structure-specific endonucleases that resolve Holliday junctions,” Mol. Microbiol. vol. 39, pp. 823-834, February 2001.

[10] A. C. Déclais and D. M. Lilley, “New insight into the recognition of branched DNA structure by junction-resolving enzymes,” Curr. Opin. Struct. Biol. vol. 18, pp. 86-95, February 2008.

[11] S. C. Kowalczykowski, D. A. Dixon, A. K. Eggleston, S. D. Lauder, and W. M. Rehrauer, “Biochemistry of homologous recombination in Escherichia coli,” Microbiol. Rev. vol. 58, pp. 401-465, September 1994.

[12] S. C. West, “Enzymes and molecular mechanisms of genetic recombination,” Annu. Rev. Biochem. vol. 61, pp. 603-640, 1992.

[13] S. C. West, “The RuvABC proteins and Holliday junction processing in Escherichia coli,” J. Bacteriol. vol. 178, pp. 1237-1241, March 1996.

[14] R. G. Lloyd and G. J. Sharples, “Dissociation of synthetic Holliday junctions by E. coli RecG protein,” EMBO J. vol. 12, pp. 17-22, January 1993.

[15] C. J. Rudolph, A. L. Upton, G. S. Briggs, and R. G. Lloyd, “Is RecG a general guardian of the bacterial genome?” DNA Repair (Amst). vol. 9, pp. 210-223, March 2010.

[16] G. J. Sharples, S. N. Chan, A. A. Mahdi, M. C. Whitby, and R. G. Lloyd, “Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions,” EMBO J. vol. 13, pp. 6133-6142, December 1994.

[17] S. N. Chan, S. D. Vincent, and R. G. Lloyd, “Recognition and manipulation of branched DNA by the RusA Holliday junction resolvase of Escherichia coli,” Nucleic Acids Res. vol. 26, pp. 1560-1566, April 1998.

[18] K. Komori, S. Sakae, H. Shinagawa, K. Morikawa, and Y. Ishino, “A Holliday junction resolvase from Pyrococcus furiosus: functional similarity to Escherichia coli RuvC provides evidence for conserved mechanism of homologous recombination in Bacteria, Eukarya, and Archaea,” Proc. Natl. Acad. Sci. U.S.A. vol. 96, pp. 8873-8878, August 1999.

[19] H. Daiyasu, K. Komori, S. Sakae, Y. Ishino, and H. Toh, “Hjc resolvase is a distantly related member of the type II restriction endonuclease family,” Nucleic Acids Res. vol. 28, pp. 4540-4543, November 2000.

[20] S. Kleff, B. Kemper, and R. Sternglanz, “Identification and characterization of yeast mutants and the gene for a cruciform cutting endonuclease,” EMBO J. vol. 11, pp. 699-704, February 1992.

[21] M. F. White and D. M. Lilley, “The resolving enzyme CCE1 of yeast opens the structure of the four-way DNA junction,” J. Mol. Biol. vol. 266, pp. 122-134, February 1997.

[22] M. F. White and D. M. Lilley, “Characterization of a Holliday junction-resolving enzyme from Schizosaccharomyces pombe,” Mol. Cell. Biol. vol. 17, pp. 6465-6471, November 1997.

[23] S. C. Ip, U. Rass, M. G. Blanco, H. R. Flynn, J. M. Skehel, and S. C. West, “Identification of Holliday junction resolvases from humans and yeast,” Nature. vol. 456, pp. 357-361, November 2008.

[24] J. M. Svendsen and J. W. Harper, ”GEN1/Yen1 and the SLX4 complex: Solutions to the problem of Holliday junction resolution,” Genes Dev. vol. 24, pp. 521-536, March 2010.

[25] U. Rass, S. A. Compton, J. Matos, M. R. Singleton, S. C. Ip, M. G. Blanco, et al., “Mechanism of Holliday junction resolution by the human GEN1 protein,” Genes Dev. vol. 24, pp. 1559-1569, July 2010.

[26] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, “Basic local alignment search tool,” J. Mol. Biol. vol. 215 pp. 403-410, October 1990.

[27] R. L. Tatusov, M. Y. Galperin, D. A. Natale, and E. V. Koonin, “The COG database: a tool for genome-scale analysis of protein functions and evolution,” Nucleic Acids Res. vol. 28, pp. 33-36, January 2000.

[28] D. Higgins, J. Thompson, T. Gibson, J. D. Thompson, D. G. Higgins, and T. J. Gibson, “CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice,” Nucleic Acids Res. vol. 22, pp. 4673-4680, November 1994.

[29] J. Felsenstein, “PHYLIP: Phylogeny Inference Package (Version 3.2),” Cladistics. vol. 5, pp. 164-166, 1989.

[30] Z. Baharoglu, M. Petranovic, M. J. Flores, and B. Michel, “RuvAB is essential for replication forks reversal in certain replication mutants,” EMBO J. vol. 25, pp. 596-604, February 2006.

[31] A. S. Bradley, Z. Baharoglu, A. Niewiarowski, B. Michel, and I. R. Tsaneva, “Formation of a stable RuvA double tetramer is required for efficient branch migration in vitro and for replication fork reversal in vivo,” J. Biol. Chem. in press.

[32] F. D. Ciccarelli, T. Doerks, C. von Mering, C. J. Creevey, B. Snel, and P Bork, “Toward automatic reconstruction of a highly resolved tree of life,” Science. vol. 311, pp. 1283-1287, March 2006.

[33] Z. Zhuang, F. Song, B. M. Martin, and D. Dunaway-Mariano, “The YbgC protein encoded by the ybgC gene of the tol-pal gene cluster of Haemophilus influenzae catalyzes acyl-coenzyme A thioester hydrolysis,” FEBS Lett. vol. 516, pp. 161-163, April 2002.

[34] D. Gully and E. Bouveret, “A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in Escherichia coli,” Proteomics. vol. 6, pp. 282-293, January 2006.

[35] E. Kolker, K. S. Makarova, S. Shabalina, A. F. Picone, S. Purvine, T. Holzman T, et al., “Identification and functional analysis of

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

'hypothetical' genes expressed in Haemophilus influenzae,” Nucleic Acids Res. vol. 32, pp. 2353-2361, April 2004.

[36] D. H. Shin, H. Yokota, R. Kim, and S. H. Kim, “Crystal structure of conserved hypothetical protein Aq1575 from Aquifex aeolicus,” Proc. Natl. Acad. Sci. U.S.A. vol. 99, pp. 7980-7985, July 2002.

[37] H. Liang, L. Li, Z. Dong, M. G. Surette, and K. Duan, “The YebC family protein PA0964 negatively regulates the Pseudomonas aeruginosa quinolone signal system and pyocyanin production,” J. Bacteriol. vol. 190, pp. 6217-6227, September 2008.

[38] W. Weraarpachai, H. Antonicka, F. Sasarman, J. Seeger, B. Schrank, J. E. Kolesar, et al., “Mutation in TACO1, encoding a translational activator of COX I, results in cytochrome c oxidase deficiency and late-onset Leigh syndrome,” Nat. Genet. vol. 41, pp. 833-837, July 2009.

[39] M. Y. Galperin and E. V. Koonin, ” From complete genome sequence to ‘complete’ understanding?” Trends Biotechnol. vol. 28, pp. 398-406, August 2010.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE