

Genome-wide comprehensive analysis of human helicases

Pavan Umate, Narendra Tuteja* and Renu Tuteja*

International Center for Genetic Engineering and Biotechnology; Aruna Asaf Ali Marg; New Delhi, India

Dedicated to the memory of Professor Arturo Falaschi

Key words: DEAD-box, DHX and DDX helicases, human helicases, MCM proteins, RecQ helicases, RNA helicases

Abbreviations: ATP, adenosine triphosphate; BLM, Bloom; MCM, minichromosome maintenance; RTS, Rothmund-Thomson syndrome; SKIV2L2, superkiller viralicidic activity 2-like 2; WRN, Werner

Helicases are motor proteins that catalyze the unwinding of duplex nucleic acids in an ATP-dependent manner. They are involved in almost all the nucleic acid transactions. In the present study, we report a comprehensive analysis of helicase gene family in human and its comparison with homologs in model organisms. The human genome encodes for 95 non-redundant helicase proteins, of which 64 are RNA helicases and 31 are DNA helicases. 57 RNA helicases are validated based on annotations and occurrence of conserved helicase signature motifs. These include 14 DExH and 37 DExD subfamily members, six other members such as U5.snRNP, ATR-X, Suv3, FANCI, and two of superkiller viralicidic activity 2-like helicases. 31 DNA helicases are also identified, which include RecQ, MCM and RuvB-like helicases. Finding a set of helicases in human and almost similar sequences in model organisms suggests that the "core" members of helicase gene family are highly conserved throughout evolution. The present study gives an overview of members of RNA and DNA helicases encoded by the human genome along with their conserved motifs, phylogeny and homologs in model organisms. The study on comparing these homologs will spread light on the organization and complexity of helicase gene family in model organisms. The comprehensive analysis of human helicases presented in this study will further provide an invaluable resource for elaborate biological research on these helicases.

Introduction

The separation of complementary strands of a DNA and/or RNA duplex is fundamental requirement for many cellular metabolic processes. This is mediated by DNA and/or RNA helicases that unwind nucleic-acid duplexes in an ATP-dependent manner. The energy of ATP hydrolysis is required for an active unwinding process. Generally, helicases exhibit specific polarity depending upon the strand on which they move either in 3'-5' or 5'-3' direction.

It is interesting to note that the unwinding reaction is independent of sequence specificity. The unwinding property of helicases depends on the type and structure of nucleic acid substrate. Thus all helicases are classified as those that preferentially unwind DNA or RNA duplexes. The unwinding of duplexes provides access to the template for proteins of the transcription, recombination, replication and repair machineries.¹ The proteins with which the DNA and/or RNA helicases interact, and their substrates, might be responsible for difference in functioning of DNA and/or RNA helicases.

Each helicase *in vivo* requires a specific structure of the substrate, interaction with specific proteins in order to participate in a very specific pathway of nucleic acid metabolism. Thus, each helicase must be meticulous in selecting those proteins and substrates with which it will interact. Helicases act in coordination with other proteins/factors and provide molecular motor function. Helicases are ubiquitous and several DNA and RNA helicases have been identified, isolated and characterized till date from a variety of organisms.²⁻⁵ The sequence analysis of these enzymes revealed that these proteins contain some well conserved helicase motifs that are involved in ATPase and helicase activities.^{3,6,7} Due to presence of the sequence DExD in motif II, these proteins are commonly known as DExD box proteins.^{3,7} Likewise, the RNA helicase families are named as DEAD (DDX), DEAH or DExH (DHX) according to the conserved motif II. These conserved helicase signature motifs of the DExD/H helicases are concentrated in a central core region that spans ~350–400 amino acids.^{8,9} But the N- and C-terminal regions are highly variable in size and composition⁹ and are most probably required for cellular localization and/or interactions with RNAs or other proteins.¹⁰ The DExD/H-box proteins play vital role in almost all cellular events that involve DNA and/or RNA. Biochemical characterization by studying the enzymatic activity of DExD/H-box proteins is often used to study the function of these proteins. The DExD/H helicase family constitutes DEAD-box and related DEAH, DExH and DExD families which belong to SF2 family.^{8,9,11} The variation within their conserved motifs distinguishes the closely related helicase families. The largest of all the families is the DEAD-box family which is characterized by the presence of nine (Q, I, Ia, Ib and from II to VI) conserved helicase motifs.^{3,6,7}

*Correspondence to: Narendra Tuteja and Renu Tuteja; Email: narendra@icgeb.res.in and renu@icgeb.res.in
Submitted: 10/01/10; Accepted: 10/03/10
DOI: 10.4161/cib.4.1.13844

The DExD/H-box proteins can use ATP hydrolysis to promote directly or indirectly the RNA-RNA unwinding,^{12,15} RNA-protein dissociation,¹⁴ and protein-protein interaction.¹⁵ The unwinding of dsRNA in vitro was demonstrated by a subset of DExD/H-box proteins.¹⁶⁻¹⁹ The cellular functions of DExD/H-box protein are also fulfilled upon displacement of protein factors by DExD/H-box proteins from the single-stranded RNA.¹⁴

The sequence homologies among the members of DEAD-box helicases are shared beyond and between the nine conserved helicase motifs. The role of these motifs in ATP-binding, ATP-hydrolysis (ATPase), RNA binding and dsRNA unwinding/helicase activities was shown based on structural, mutational and biochemical analyses.²⁰⁻²² The motifs I (AxxGxGKT) and II (VLDEAD) are characteristic of nucleotide triphosphate (NTP) hydrolyzing proteins, the motif III (SAT) is involved in coupling of ATP hydrolysis and unwinding. The motifs IV (FVNT) and V (RGxD) are involved in RNA binding and the motif VI (HRIGRxxR) is required for the nucleic acid-dependent NTP hydrolysis.^{7,23} Using the energy derived from NTP hydrolysis RNA helicases function as molecular motors that rearrange RNA secondary structure thus playing vital role in many cellular processes.²⁴ All cellular organisms and many viral genomes contain RNA helicases. In yeast *Saccharomyces cerevisiae*, the study of DEAD-box and related proteins has provided the best overview of cellular processes involving RNA helicases.²⁵ Extensive genetic analyses in *S. cerevisiae* revealed the role of RNA helicases in RNA metabolism and all RNA-related processes. This genetic analysis has shown the role of individual RNA helicases in specific cellular processes, while biochemical experiments have shown that these proteins in the presence of ATP and Mg²⁺ can dissociate short RNA duplexes. Apart from disrupting secondary structures or short RNA-RNA interactions of a few base pairs, RNA helicases are also involved in unwinding long stretches of double stranded RNA.²⁶ The dynamic RNA-RNA interactions and denaturation of short and/or long RNA duplexes seems to be necessary for all cellular processes that involve RNA helicases.

The DEAD-box RNA helicases which are found in all eukaryotes and most prokaryotes constitute a large family of proteins.^{4,5,20,24,27} In the genomes of *Caenorhabditis elegans* and *Drosophila melanogaster*, nearly 30 genes that encode DEAD-box RNA helicases were identified.²⁸ The *Arabidopsis thaliana* and rice (*Oryza sativa* L.) genomes encode ~32 and ~21 members of DEAD-box RNA helicases respectively.^{5,27,28} *Plasmodium falciparum* genome also encodes a number of DEAD/DEAH box and other helicases.⁴ As reported earlier, there are at least 36 and 14 members of the DEAD and DEAH families of putative RNA helicases in humans.²⁹

In the present study, we have compiled complete set of 95 DNA and RNA helicases encoded by the human genome. We report identification of 58 DExD/H RNA helicases which includes 42 and 16 members of the DEAD (DDX) and DEAH (DHX) families of putative RNA helicases respectively. In addition, we have also identified six RNA helicases which are independent from DDX and DHX families. These are the U5.snRNP 200 kDa helicase, transcriptional regulator ATR-X helicase, RNA helicase SUPV3L1 (mitochondria), superkiller viralicidic activity 2-like 2

(SKIV2L2), helicase SKI2W (also called as DDX13) and helicase FANCJ. A total of 31 DNA helicases (like RecQ members, KU70, MCM proteins, XPB and XPD repair helicases, RuvB-like helicases, Pif1, Twinkle, BACH1, RecQ alpha, beta and gamma and RTEL1) are also identified. Further, these helicases are validated based on protein alignment studies for conserved helicase signature motifs. A phylogenetic relationship is drawn among different members of RNA and DNA helicases. A comparative bioinformatics approach is adopted to identify homologs of human RNA and DNA helicases in different model organisms like *S. cerevisiae*, *A. thaliana*, *O. sativa* and *P. falciparum*. The data described in this manuscript clearly indicate that a multitude of helicases are encoded by the human genome. The present studies ultimately lead toward a holistic view of the structure, composition and phylogeny of helicase gene family in humans. Insights derived from this study combined with the comparative in silico approach toward identification of homologs in other model organisms will further our understanding of these ubiquitous motor proteins.

Results

RNA helicases. We have identified 64 RNA helicases from human genome. A list of these helicases with protein Id, protein name, synonym and gene name is given in Table 1. These 64 RNA helicases are categorized into DHX, DDX and independent families. The human genome contains 16 of DHX, 42 of DDX and six independent helicases.

DExH (DHX) helicases. The identified 16 DHX members of RNA helicases are as follows: DHX8, DHX9, DHX15, DHX16, DHX29, DHX30, DHX32, DHX33, DHX34, DHX35, DHX36, DHX37, DHX38, DHX40, DHX57 and DHX58 (Table 1). The protein alignment studies indicated the presence of 8 conserved helicase motifs (except for DHX32 and DHX58) in all 14 of the DHX members (Fig. 1A). These conserved helicase motifs are as follows, GxxGxGKT, TQPRRV, TDGML, DExH, SAT, FLTG, TNIAET and QRxGRAGR (Fig. 1A). The protein residues for helicase motifs in DHX32 and DHX58 showed major alterations from the canonical sequences that are typical for helicase motifs of DHX members of RNA helicases (Fig. 1B). DHX32 showed absence of typical helicase motifs like SAT, FLTG, TNIAET and QRxGRAGR while changes in the sequence for helicase motifs like GxxKxGKS, TQVHKQ, TDDML and DDIH are observed (Fig. 1B). Similarly, DHX58 has motif DECH but showed change in motifs, LxxGxGKT, TAS and QARGRAR (Fig. 1B).

Supplemental Figure 1 shows phylogenetic analysis performed on complete protein sequences for the 14 DHX members of RNA helicases. Most of these members showed high degree of sequence conservation as indicated from higher bootstrap values (Sup. Fig. 1). The members DHX36 and DHX57 emerged as a single clade with striking sequence identity (Sup. Fig. 1). A low bootstrap value is obtained when phylogenetic analysis was made on complete protein sequence for members DHX16 and DHX35 (Sup. Fig. 1). However, the regions corresponding to helicase motifs in DHX16 and DHX35 are well conserved (see Fig. 1A).

Table 1. Human RNA helicases (continued)

S. no.	Swiss-Prot Id	Protein name	Synonym	Gene name
1	Q92499	RNA helicase DDX1	DEAD-box protein 1	<i>DDX1</i>
			DEAD-box protein retinoblastoma	
			DBP-RB	
2	P60842	RNA helicase eIF4A-1	Eukaryotic initiation factor eIF4A-I	<i>eIF4A-I</i>
			DDX2A	<i>DDX2A</i>
3	Q14240	RNA helicase eIF4A-2	Eukaryotic initiation factor 4A-II	<i>eIF4A-II</i>
			DDX2B	<i>DDX2B</i>
4	O00571	RNA helicase DDX3X	DEAD-box protein 3, X-chromosomal	<i>DDX3X</i>
			Helicase-like protein 2	<i>DBX, DDX3</i>
			HLP2	
			DEAD-box, X isoform	
5	O15523	RNA helicase DDX3Y	DEAD-box Y (DBY), Y-chromosomal	<i>DDX3Y</i>
6	Q9NQ10	RNA helicase DDX4	DEAD-box protein 4	<i>DDX4</i>
			Vasa homolog	
7	P17844	RNA helicase DDX5	DEAD-box protein 5	<i>DDX5</i>
			RNA helicase p68	
8	P26196	RNA helicase DDX6	DEAD-box protein 6	<i>DDX6</i>
			RNA helicase p54	<i>HLR2</i>
			Oncogene RCK	<i>RCK</i>
9	Q14562	RNA helicase DHX8	DEAH-box protein 8	<i>DHX8</i>
			RNA helicase HRH1	
10	Q08211	RNA helicase A	Nuclear DNA helicase II	<i>DHX9</i>
			NDH II	
			DEAH-box protein 9	
11	Q13206	RNA helicase DDX10	DEAD-box protein 10	<i>DDX10</i>
12	Q96FC9	RNA helicase DDX11	DEAD/H-box protein 11	<i>DDX11</i>
			CHL1-related protein 1 (hCHLR1)	
			Keratinocyte growth factor-regulated gene	
			KRG-2	
13	Q92771	RNA helicase DDX12	DEAD/H-box protein 12	<i>DDX12</i>
			CHL1-related protein 2 (hCHLR2)	
14	Q15477	Helicase SKI2W	Helicase-like protein	<i>SKIV2L</i>
			HLP	<i>DDX13</i>
15	O43143	RNA helicase DHX15	DEAH-box protein 15	<i>DHX15</i>
			ATP-dependent RNA helicase #46	
16	O60231	RNA helicase DHX16	DEAH-box protein 16	<i>DHX16</i>
			ATP-dependent RNA helicase #3	
17	Q92841	RNA helicase DDX17	DEAD-box protein 17	<i>DDX17</i>
			RNA-dependent helicase p72	
			DEAD-box protein p72	
18	Q9NVP1	RNA helicase DDX18	DEAD box protein 18	<i>DDX18</i>
			Myc-regulated DEAD-box protein	
			MrDb	
19	Q9NUU7	RNA helicase DDX19A	DEAD-box protein 19A	<i>DDX19A</i>
			DDX19-like protein	
20	Q9UMR2	RNA helicase DDX19B	DEAD box protein 19B	<i>DDX19B</i>
			DEAD box RNA helicase DEAD5	

Table 1. Human RNA helicases (continued)

21	Q9UHI6	RNA helicase DDX20	DEAD-box protein 20	<i>DDX20</i>
			DEAD-box protein DP 103	
			Component of gems 3	
			Gemin-3	
22	Q9NR30	Nucleolar RNA helicase 2	Nucleolar RNA helicase II	<i>DDX21</i>
			Nucleolar RNA helicase Gu	
			RH II/Gu	
			Gu-alpha	
			DEAD-box protein 21	
23	Q9BUQ8	RNA helicase DDX23	DEAD-box protein 23	<i>DDX23</i>
			100 kDa U5 snRNP-specific protein (U5-100 kD)	
			PRP28 homolog	
24	Q9GZR7	RNA helicase DDX24	DEAD-box protein 24	<i>DDX24</i>
25	Q9UHL0	RNA helicase DDX25	DEAD-box protein 25	<i>DDX25</i>
			Gonadotropin-regulated testicular RNA helicase	
26	Q96GQ7	RNA helicase DDX27	DEAD-box protein 27	<i>DDX27</i>
27	Q9NUL7	RNA helicase DDX28	Mitochondrial DEAD-box protein 28	<i>DDX28</i>
28	Q7Z478	RNA helicase DHX29	DEAH-box protein 29	<i>DHX29</i>
			Nucleic acid helicase DDXx	
29	Q7L2E3	RNA helicase DHX30	DEAH-box protein 30	<i>DHX30</i>
30	Q9H8H2	RNA helicase DDX31	DEAD-box protein 31	<i>DDX31</i>
31	Q7L7V1	RNA helicase DHX32	DEAH-box protein 32	<i>DHX32</i>
			DEAD/H helicase-like protein 1 (DHLP1)	
			DEAD/H box 32	
			HuDDX32	
32	Q9H6R0	RNA helicase DHX33	DEAH-box protein 33	<i>DHX33</i>
33	Q14147	RNA helicase DHX34	DEAH-box protein 34	<i>DHX34</i>
34	Q9H5Z1	RNA helicase DHX35	DEAH-box protein 35	<i>DHX35</i>
35	Q9H2U1	RNA helicase DHX36	DEAH-box protein 36	<i>DHX36</i>
			MLE-like protein 1	
			RNA helicase associated with AU-rich element ARE	
36	Q8IY37	RNA helicase DHX37	DEAH-box protein 37	<i>DHX37</i>
37	Q92620	RNA helicase PRP16	RNA helicase DHX38	<i>DHX38</i>
			DEAH-box protein 38	
38	O00148	RNA helicase DDX39	DEAD-box protein 39	<i>DDX39</i>
			Nuclear RNA helicase URH49	
			DEAD-box protein UAP56	
			Spliceosome RNA helicase BAT1	
			RNA helicase p47	
39	Q8IX18	RNA helicase DHX40	DEAH-box protein 40	<i>DHX40</i>
			Protein PAD	
40	Q9UJV9	RNA helicase DDX41	DEAD-box protein 41	<i>DDX41</i>
41	Q86XP3	RNA helicase DDX42	DEAD-box protein 42	<i>DDX42</i>
			Splicing factor 3B-associated 125 kDa protein	
			SF3b DEAD-box protein	

Table 1. Human RNA helicases (continued)

42	Q9NXZ2	RNA helicase DDX43	DEAD-box protein 43	<i>DDX43</i>
			DEAD box protein HAGE	
			Helical antigen	
			Cancer/testis antigen 13	
			CT13	
43	Q7L014	RNA helicase DDX46	DEAD-box protein 46	<i>DDX46</i>
			PRP5 homolog	
44	Q9H0S4	RNA helicase DDX47	DEAD-box protein 47	<i>DDX47</i>
45	P38919	RNA helicase eIF4A-3	Eukaryotic initiation factor 4A-III	<i>eIF4A-III</i>
			RNA helicase DDX48	<i>DDX48</i>
			DEAD box protein 48	
			Nuclear matrix protein 265 (hNMP 265)	
46	Q9Y6V7	RNA helicase DDX49	DEAD-box protein 49	<i>DDX49</i>
47	Q9BQ39	RNA helicase DDX50	DEAD-box protein 50	<i>DDX50</i>
			Nucleolar protein Gu2	
			Gu-beta	
48	Q8N8A6	RNA helicase DDX51	DEAD-box protein 51	<i>DDX51</i>
49	Q9Y2R4	RNA helicase DDX52	DEAD-box protein 52	<i>DDX52</i>
			RNA helicase ROK1-like	
50	Q86TM3	RNA helicase DDX53	DEAD box protein 53	<i>DDX53</i>
			DEAD box protein CAGE	
			Cancer-associated gene protein	
			Cancer/testis antigen 26/CT26	
51	Q8TDD1	RNA helicase DDX54	DEAD-box protein 54	<i>DDX54</i>
			RNA helicase DP97	
			DEAD-box RNA helicase 97 kDa	
52	Q8NHQ9	RNA helicase DDX55	DEAD box protein 55	<i>DDX55</i>
53	Q9NY93	RNA helicase DDX56	DEAD-box protein 56	<i>DDX56</i>
			ATP-dependent 61 kDa nucleolar RNA helicase	
			DEAD-box protein 21	
54	Q6P158	RNA helicase DHX57	DEAH-box protein 57	<i>DHX57</i>
55	O95786	RNA helicase DDX58	DEAD-box protein 58	<i>DDX58</i>
			Retinoic acid-inducible gene 1 protein	
			RIG-1	
56	Q96C10	RNA helicase DHX58	DEAH-box protein 58	<i>DHX58</i>
			helicase LGP2	
			Protein D11Lgp2 homolog	
			RIG-I-like receptor LGP2	
57	Q5T1V6	RNA helicase DDX59	DEAD-box protein 59	<i>DDX59</i>
			Zinc finger HIT domain-containing protein 5	
58	Q8IY21	RNA helicase DDX60	DEAD-box protein 60	<i>DDX60</i>
59	Q13838	Spliceosome RNA helicase BAT1	DEAD box protein UAP56	<i>BAT1</i>
			56 kDa U2AF65-associated protein	<i>UAP56</i>
			ATP-dependent RNA helicase p47	
			HLA-B-associated transcript 1 protein	
60	O75643	U5.snRNP 200 kDa helicase	U5 snRNP-specific 200 kDa protein	<i>ASCC3L1</i>
			U5-200KD	

Table 1. Human RNA helicases (continued)

			Activating signal cointegrator 1 complex subunit 3-like 1	
			BRR2 homolog	
61	P46100	Transcriptional regulator ATRX helicase	ATP-dependent helicase ATRX	<i>ATRX</i>
			X-linked helicase II	
			X-linked nuclear protein (XNP)	
			Znf-HX	
62	Q8IYB8	RNA helicase SUPV3L1, mitochondrial	SUV3-like protein 1	<i>SUPV3L1</i>
			Suppressor of var1 3-like protein 1	<i>SUV3</i>
63	P42285	Superkiller viralicidic activity 2-like 2	ATP-dependent helicase SKIV2L2	<i>SKIV2L2</i>
			KIAA0052	
64	Q9BX63	Fanconi anemia group J protein	RNA helicase BRIP1	<i>FANCI</i>
		Protein FACJ	BRCA1-interacting protein 1	

DExD (DDX) helicases. The present analysis has identified 42 DDX members of RNA helicases (Table 1), of which 37 DDX members are validated based on conserved helicase motif signatures (Fig. 2A). These 37 members are, DDX2A, DDX2B, DDX3X, DDX3Y, DDX4, DDX5, DDX6, DDX10, DDX17, DDX18, DDX19A, DDX19B, DDX20, DDX21, DDX23, DDX24, DDX25, DDX27, DDX28, DDX31, DDX39, DDX41, DDX42, DDX43, DDX46, DDX47, DDX48, DDX49, DDX50, DDX51, DDX52, DDX53, DDX54, DDX55, DDX56, DDX59 and BAT1 (Table 1). The protein alignment for the above 37 DDX members of RNA helicases is shown in Figure 2A. In these members, the 9 conserved helicase motifs which are identified are as follows, GFxxPxxIQ, AxxGxGKT, PTRELA, TPGR, DExD, SAT, FVxT, RGxD and HRxGRxxR (Fig. 2A).

Some of these helicase motifs are either completely absent or showed changes in amino acid composition in 5 DDX members, DDX1, DDX11, DDX12, DDX58 and DDX60. In case of DDX1, the motifs GFxxPxxIQ, AxxGxGKT are absent while, motifs PSRELA (instead of PTRELA), TPGR, DEAD, SAT, FCRT (instead of FVNT), RGID and HRIGRVGR are all present (Fig. 2B). The diagrammatic representation of helicase motifs found in DDX11, DDX12, DDX58 and DDX60 is shown in Figure 1B. These helicases, although classified as DDX members, contain a DExH motif (Fig. 1B).

The protein sequence based phylogenetic analysis revealed high level identity features among various DDX family members of RNA helicases (Sup. Fig. 2). For instance, close relationships are established between members like DDX2A-DDX2B, DDX19A-DDX19B, DDX39-BAT1, DDX21-DDX50, DDX3X-DDX3Y, DDX5-DDX17, DDX43-DDX53, DDX10-DDX18

and DDX47-DDX49 which emerged as independent clades with high bootstrap values (Sup. Fig. 2). For the remaining DDX members, a moderately high to low bootstrap support is obtained (Sup. Fig. 2) although these members show a high degree of conservation in the regions composed of helicase motifs (see Fig. 2A).

Independent category of RNA helicases. This category includes helicases which neither belong to DHX nor to DDX members of RNA helicases. We classified six members under this category like the U5.snRNP 200 kDa helicase, transcriptional regulator ATR-X [X-linked alpha-thalassemia/mental retardation syndrome (ATR-X syndrome)], RNA helicase SUPV3L1 (mitochondria), superkiller viralicidic activity 2-like 2 (SKIV2L2), helicase SKI2W and helicase FANCI (Table 1). In humans, the superkiller helicase family contains only 2 members, SKIV2L2 and SKI2W. The protein alignment of these superkiller helicases showed 8 conserved helicase motifs (Fig. 3). The motif I (AHTSAGKT), motif Ia (TSPIKALSNQ), motif Ib (MTTEIL), motif II (DExH), motif III (SAT), motif IV (xxFTFSx), motif V (TETFAMGxNMPA) and motif VI (QMxGRAGR) are identified in SKIV2L2 and SKI2W helicases (Fig. 3).

DNA helicases. A list of identified DNA helicases from human genome is given in Table 2. Altogether, 31 DNA helicases have been identified which includes 5 RecQ members, KU70, 9 MCM (minichromosome maintenance) proteins, nucleolin, 2 chromodomain helicases, 2 DNA repair helicases, lymphoid-specific helicases, DNA helicase INO80 complex homolog 1 and 2 RuvB-like helicases, Pif1, Twinkle, BACH1 and three RecQ5 isoforms alpha, beta and gamma, and regulator of telomere elongation helicase (RTEL1) (Table 2).

Figure 1 (See next page). Protein alignment of 14 DHX family members of RNA helicases. (A) Entire protein sequence is aligned with ClustalX software. The eight conserved helicase signature motifs are identified on top of the alignment. The names of DHX helicases are identified on the left and sequence positions are given on right. Asterisks and dots drawn on top of sequence indicate identical residues and conservative amino acid changes, respectively. Gaps in the amino acid sequences are introduced to improve the alignment. Only a part of protein alignment with conserved helicase motifs is shown. (B) Schematic representations for DHX32 and DHX58 helicase motifs are shown. Note the presence of DExH motif in DDX members like DDX11, DDX12, DDX58 and DDX60. The typical motifs for DHX-type family members are represented on top.

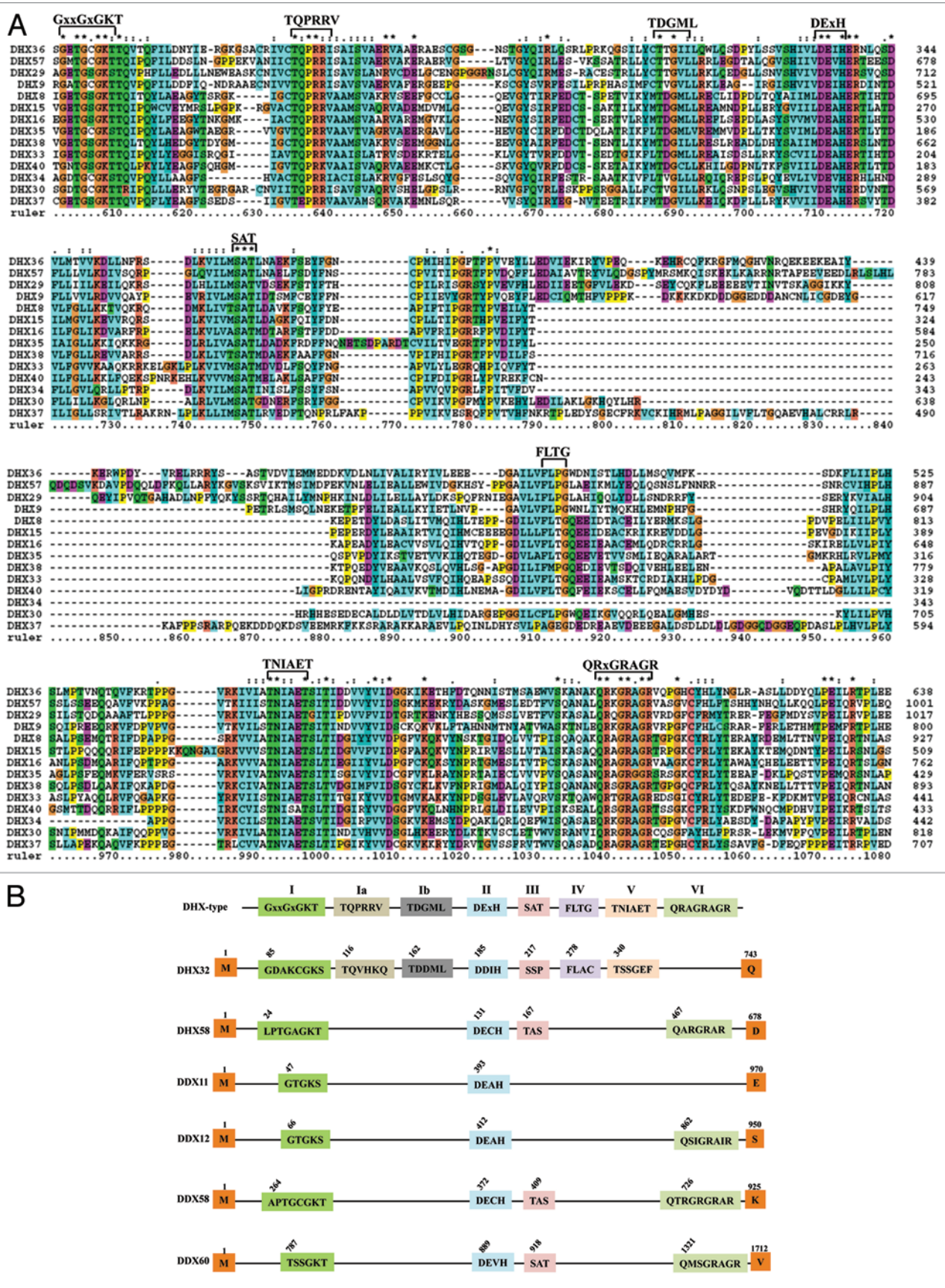


Figure 1. For figure legend, see page 123.

The 5 RecQ members which are identified are, RecQ1 (DNA helicase Q1), RecQ2 [Bloom (BLM) syndrome], RecQ3 [Werner (WRN) syndrome], RecQ4 [DNA helicase Q4 or Rothmund-Thomson syndrome (RTS)] and RecQ5 (DNA helicase Q5) (Table 2). Of the 5 RecQ members that are identified, two are smaller (RecQ1 and RecQ5 of 649 and 991 amino acids, respectively) and three larger (BLM, WRN and RecQ4 of 1,417, 1,432 and 1,208 amino acids, respectively). A common structural feature in this family of enzymes is the presence of central domain with ~400–450 amino acids containing the 7 helicase signature motifs (Fig. 4). These motifs are MPTGGGKSL, VISPLRSL, DExH, TATA, SGIIYC, ATIAFGMGID and QESGRAGR (Fig. 4). It is interesting to observe a high degree of sequence conservation within these 7 helicase motifs among all the members of the RecQ family (Fig. 4). The phylogenetic analysis of the 5 RecQ family members is shown in Supplemental Figure 3. The RecQ1-RecQ2 and RecQ3-RecQ4 form separate clades with 97 and 58% of bootstrap support respectively (Sup. Fig. 3).

The MCM complex in humans is composed of subunits MCM2-7 (Table 2). Other MCMs like MCM8-10 are also identified (Table 2). The MCM2-7, MCM8 and MCM9 are characterized by the presence of the MCM-box of ~200 residues (Fig. 5). The MCM-box showed two ATPase consensus motifs, the Walker A and Walker B motifs (Fig. 5). An invariant lysine (K) residue which is found in all ATP-binding proteins also occurs in the Walker A motif (Fig. 5). The MCM-specific consensus, GDPxx(A/S)KSQ, is present in Walker A consensus motif wherein a glycine (G) occurs in place of alanine (A) or serine (S) in MCM8 and MCM9 along with several additional conserved residues (Fig. 5). The Walker B element consensus, IDEFDKM, is present in all MCMs (MCM2-7 and MCM8) except MCM9 wherein the amino acids residues in this motif are only partially conserved (IDEFNLS) (Fig. 5). All MCM proteins contain a short motif SRFD known as the “arginine finger” approximately 60 residues after the Walker B element (Fig. 5). MCM10 has no obvious sequence similarity with MCM2-7. The protein sequences of MCM2-7, MCM8 and MCM9 are analyzed for their phylogenetic relationship (Sup. Fig. 4). The study revealed a close relationship between MCM3 and 7, MCM4 and 5 and MCM8 and 9 (Sup. Fig. 4).

The DNA repair helicases, XPB and XPD, have also been identified in human genome (Table 2). Each of these helicases shows the presence of 7 consensus motifs indicated as I, Ia, II, III, IV, V and VI (Fig. 6A and B). However, the amino acid residues of the 7 consensus motifs differ between these two helicases (Fig. 6A and B). In the XPB helicase, the amino acid composition of these motifs is, CGAGKS (motif I), VEQWK (motif Ia), LDEVH (motif II), LGLTATLxRED (motif III), PTSQGER (motif IV), FDLPEANV (motif V) and QRLGRVLR (motif VI) (Fig. 6A). These 7 consensus motifs in XPD helicase are, SGTGKT (motif I), YCSRTV (motif Ia), FDEAH (motif II), TSGTSLPL (motif III), FTSYQYME (motif IV), SVARGKVS (motif V) and QCVGRAIR (motif VI) (Fig. 6B). The protein alignment of DNA helicases RuvB-like 1 and RuvB-like 2 is shown in Figure 7. The protein length is 456 and 463 amino acids in RuvB-like 1 and RuvB-like 2 respectively. The protein

sequence of these helicases is relatively well conserved (Fig. 7). The characteristic Walker A and Walker B motifs are present (Fig. 7). The Walker A consensus, G(x)₄GKT and Walker B consensus DExH are well conserved in these helicases (Fig. 7).

Homology search in model organisms. We identified the homologs of human RNA and DNA helicases in model organisms like *S. cerevisiae*, *A. thaliana*, *O. sativa* and *P. falciparum*. The Table 3 shows homologs of human RNA helicases in these organisms. Several human homologs for DDX and DHX members in rice have been annotated as DEAD-box, or DEAD/DEAH-box proteins respectively (Table 3). The homologs of human DNA helicases in *S. cerevisiae*, *A. thaliana*, *O. sativa* and *P. falciparum* are given in Table 4. We have also identified human homologs of RNA and DNA helicases in *P. falciparum* and reconfirmed their names/annotations based on the BLAST searches (Sup. Tables 1 and 2).

Discussion

Helicases are DNA and RNA unwinding enzymes found in various organisms ranging from prokaryotes to mammals, cellular pathogens and many viruses. With the availability of genome sequences several new helicases have been identified in various organisms. Many of these helicases are ‘putative,’ which are partially or completely uncharacterized in the databases. Nevertheless, defining the complete set of genome-encoded helicases is a prerequisite for detailed characterization of each helicase type. There has been tremendous interest in studying helicases as crucial regulators of several key aspects of cellular metabolism including DNA replication, DNA repair, RNA transcription, translation, and also as potential drug targets.⁴ In addition, the importance of helicases can be reemphasized from our present findings that a human cell possesses a large number of helicases, up to 31 DNA and 64 RNA helicases, respectively. The present study has identified and validated sets of both RNA and DNA helicases from human. We have unequivocally validated 14 DHX and 37 DDX family members of RNA helicase based on protein alignment studies followed by identification for the presence or absence of helicase consensus motifs. It is worth mentioning here that some of the putative RNA helicase members like DHX32, DHX58, DDX1, DDX11, DDX12, DDX58 and DDX60 wherein the helicase motifs are partially conserved, or absent, need to be characterized biochemically before classifying them as bona fide helicases.

Based on in vitro experiments the contribution of several DExD/H box proteins in the accomplishment of crucial cellular functions has been revealed. DDX and DHX members of RNA helicases have been the focus of intensive research with several important findings associated with these multifunctional enzymes. It is well established that the members of DExD/H-box protein family also act as splicing factors that are involved in many steps of spliceosome assembly and require ATP hydrolysis.³⁰ For instance, it has been shown that DHX16 is required for human pre-mRNA splicing after formation of a pre-catalytic spliceosome.³¹ The cells expressing a mutant DHX16 with mutation in the helicase domain accumulated unspliced, intron-containing

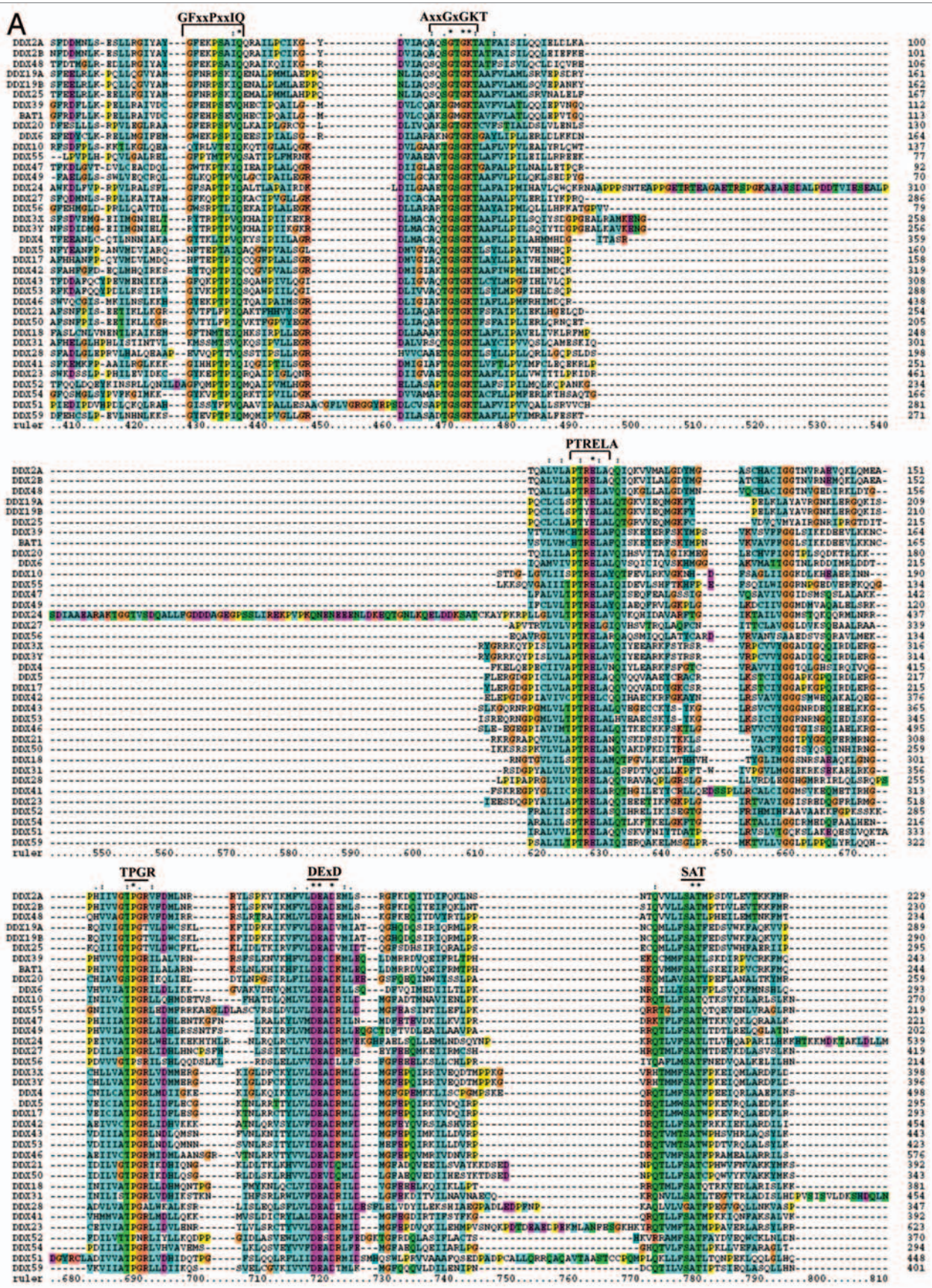


Figure 2A. For figure legend, see page 127.

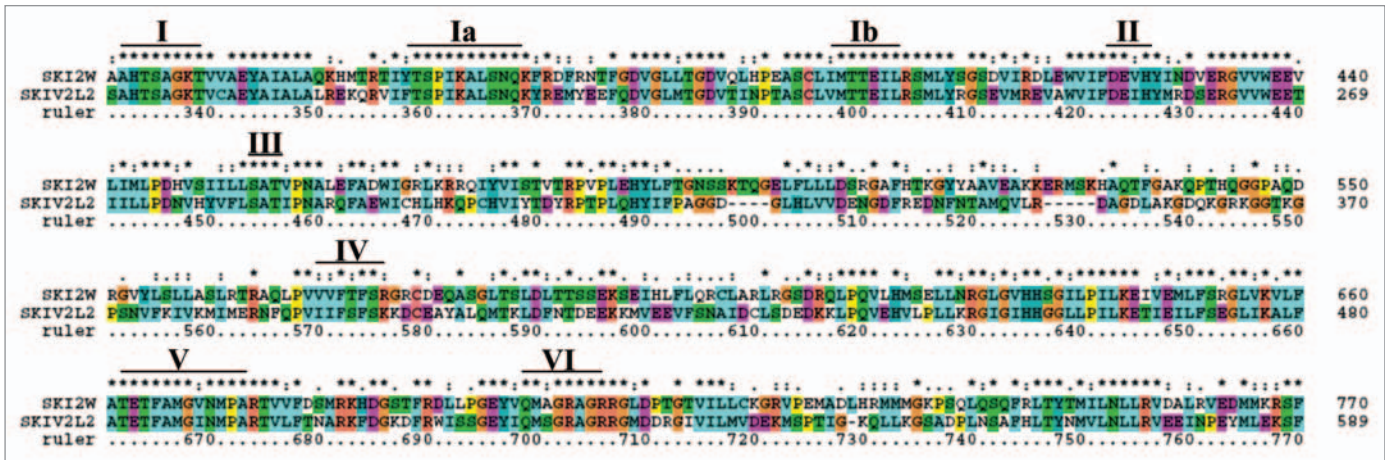


Figure 3. Protein alignment of superkiller helicase group members, SKI2W and SKIV2L2. The consensus for eight conserved motifs, motif I (AHTSAG-KT), motif Ia (TSPIKALSNQ), motif Ib (MTTEIL), motif II (DExH), motif III (SAT), motif IV (xxFTFSx), motif V (TETFAMGxNMPA) and motif VI (QMxGRAGR) are identified on top of the alignment made with ClustalX alignment program. The names of superkiller helicases are identified on the left and sequence positions are given on right. Asterisks and dots drawn on top of sequence indicate identical residues and conservative amino acid changes, respectively. Gaps in the amino acid sequences are introduced to improve the alignment. Only a part of protein alignment with conserved helicase motifs is shown.

of heat shock protein 60 (Hsp60). The expression of DHX36 mRNA in human organs is highest in testis and this protein is being utilized in cells with high rates of proliferation.³⁷ DHX36 has been reported to be responsible for major G4 DNA resolvase activity because proliferating cells contain G4 DNA structures.³⁸ DDX47 has been identified as a binding partner of GABA(A) receptor-associated protein (GABARAP) in cortical neurons and the co-transfection of GABARAP and DDX47 cDNA into a tumor cell line induced apoptosis.³⁹ Human DEAD box protein DDX42 possesses RNA helicase, protein displacement and RNA annealing activities.⁴⁰ The importance of DDX3 in anti-viral immunity, and that viruses target DDX3 for immune evasion was proposed.⁴¹ Further, a method to study the role of DDX3 in HIV-1 replication was also studied.⁴² It has been shown that DDX3 has a conserved general function mediated via its interaction with eIF3 in promoting translation.⁴³

The X- and Y-chromosome encoded DEAD-box helicases are DDX3X and DDX3Y respectively. The DEAD-box helicases are mainly RNA helicases but DDX3X can also bind to DNA.⁴⁴ The DDX3X has been reported to directly interact with transcription factors and promote transcription.⁴⁵ The orthologs of DDX3 have previously been shown to play role in diverse cellular processes like nuclear export of RNA, splicing and translation regulation.⁴⁶ This helicase has also been implicated in cancer biogenesis.⁴⁷ It has been reported that DDX3 undergoes nucleocytoplasmic shuttling and interestingly this property of DDX3 is exploited by the HIV for exporting its mRNA out of the nucleus.⁴⁸ DDX3Y is also called as DEAD-box RNA helicase Y (DBY) and DBY plays an important role in human spermatogenesis.⁴⁹ In a recent study it has been reported that the human DBY RNA helicase gene is the major azoospermia factor a (AZFa) gene.⁵⁰ Men with its deletion display no somatic pathologies, but suffer from complete absence of germ cells.⁵⁰ Accordingly, DDX3Y protein is expressed only

in the germline in spermatogonia, although the transcripts were found in many tissues.⁵⁰

Suv3 is a nuclear-gene encoded mitochondrial RNA helicase and its gene product localizes to mitochondria.^{51,52} The yeast Suv3 shows NTP-dependent RNA helicase activity in vitro.⁵¹ The multifunctional complex formed by the Suv3 together with a 3'-5' exoribonuclease (Dss1) is involved in mitochondrial RNA degradation.^{52,53} In yeast, genetic ablation of Suv3 causes respiratory negative phenotype, while it also shows an effect on stability and processing of mitochondrial RNAs, a total block of mitochondrial translation and accumulation of RNA by-products.⁵⁴ The human Suv3 shows an ATP-dependent unwinding activity in vitro and localizes to the mitochondrial matrix.⁵⁵ The Suv3 knockout mice die in utero.⁵⁶ Depletion of Suv3 leads to the accumulation of shortened mitochondrial RNA species, which is likely due to impaired mitochondrial RNA degradation.⁵⁴ A similar accumulation of shortened mitochondrial RNA like omega RNA and truncated mitochondrial RNA that cannot be degraded was observed in yeast Suv3 null cells.⁵⁷

The present in silico studies have identified 31 DNA helicases which includes several prominent members like RecQ helicases, MCM proteins and DNA repair helicases, XPB and XPD. The RecQ members are DNA helicases named after the RecQ protein in *Escherichia coli*, which was identified during a search for mutants sensitive to thymine starvation.⁵⁸ The RecQ family helicases found in prokaryotes to eukaryotes are involved in cellular processes such as recombination and DNA replication.⁵⁸ All RecQ enzymes possess 3' to 5' helicase activity.⁵⁸ WRN also encodes a 3'-5' DNA exonuclease activity.⁵⁹ Within RecQ family, there are members with short sequences like RecQ1 or RecQL,⁶⁰ and members with extremely long sequences like FFA from *Xenopus laevis*⁶¹ and QDE3 from *Neurospora crassa*,⁶² which show considerable variation in length and sequence motifs beyond the conserved helicase motifs.⁶⁰ Only 1 RecQ homolog is present in

Table 2. Human DNA helicases

S. no.	Swiss-Prot Id	Protein name	Synonym	Gene name
1	P46063	DNA helicase Q1	RECQ1 DNA-dependent ATPase Q1	<i>RECQ1</i>
2	P54132	Bloom syndrome protein	RECQ2; RECQL2; RECQL3 RecQ-like type 2 RecQ protein-like 3	<i>BLM</i>
3	Q14191	Werner syndrome protein	RECQ3; RECQL2; RECQL3	<i>WRN</i>
4	O94761	DNA helicase Q4	RecQ protein-like 4 RECQ4; RECQL4	<i>RECQ4</i>
5	O94762	DNA helicase Q5	RecQ protein-like 5 RECQ5; RECQL5	<i>RECQ5</i>
6	P12956	DNA helicase 2 subunit 1	DNA helicase II 70 kDa subunit Lupus Ku autoantigen protein p70 (ku70) 70 kDa subunit of Ku antigen Thyroid-lupus autoantigen (TLAA) CTC box-binding factor 75 kDa subunit (CTCBF) DNA repair protein XRCC6	<i>KU70</i>
7	P49736	MCM2	DNA replication licensing factor MCM2 Nuclear protein BM28	<i>MCM2</i>
8	P25205	MCM3	DNA replication licensing factor MCM3 RLF subunit beta P1-MCM3	<i>MCM3</i>
9	P33991	MCM4	DNA replication licensing factor MCM4 CDC21 homolog P1-CDC21	<i>MCM4</i>
10	P33992	MCM5	DNA replication licensing factor MCM5 CDC46 homolog P1-CDC46	<i>MCM5</i>
11	Q14566	MCM6	DNA replication licensing factor MCM6 p105MCM	<i>MCM6</i>
12	P33993	MCM7	DNA replication licensing factor MCM7 CDC47 homolog P1.1-MCM3	<i>MCM7</i>
13	Q9UJA3	MCM8	DNA replication licensing factor MCM8 Minichromosome maintenance 8	<i>MCM8</i>
14	Q9NXL9	MCM9	DNA replication licensing factor MCM9	<i>MCM9</i>
15	Q7L590	MCM10	Protein MCM10 homolog	<i>MCM10</i>
16	P19338	Nucleolin	Protein C23	<i>NCL</i>
17	O14647	CHD2	Chromodomain-helicase-DNA-binding protein 2	<i>CHD2</i>
18	Q9P2D1	CHD7	ATP-dependent helicase CHD7 CHD-7	<i>CHD7</i>
19	P19447	XPB	TFIIH basal transcription factor complex 89 kDa subunit	<i>ERCC3</i>
20	P18074	XPD	TFIIH basal transcription factor complex 80 kDa subunit	<i>ERCC2</i>
21	Q9NRZ9	Lymphoid-specific helicase	SWI/SNF2-related matrix-associated actin-dependent- regulator of chromatin subfamily A member 6 Proliferation-associated SNF2-like protein	<i>HELLS</i> <i>PASG, SMARCA6</i>
22	Q9ULG1	hINO80	DNA helicase INO80 complex homolog 1	<i>INOC1</i>
23	Q9Y265	RuvB-like 1	INO80H, NMP238, TIP49, TIP49A	<i>RUVBL1</i>

Table 2. Human DNA helicases (continued)

24	Q9Y230	RuvB-like 2	INO80J, TIP48, TIP49B	<i>RUVBL2</i>
25	Q9H611	PIF1	RRM3 DNA helicase	<i>Pif1</i>
26	Q96RR1	Twinkle	T7 gp4-like protein	<i>PEO1</i>
Progressive external ophthalmoplegia 1 protein				
T7-like mitochondrial DNA helicase				
27	AAI01475*	BACH1	BRCA1 interacting protein C-terminal helicase	<i>BRIP1; BACH1</i>
28	BAA95952*	RecQ5 alpha	DNA helicase recQ5 alpha	<i>RecQ5a</i>
29	BAA95953*	RecQ5 beta	DNA helicase recQ5 beta	<i>RecQ5b</i>
30	BAA95954*	RecQ5 gamma	DNA helicase recQ5 gamma	<i>RecQ5c</i>
31	Q9NZ71	RTEL1	Novel helicase-like	<i>RTEL1</i>

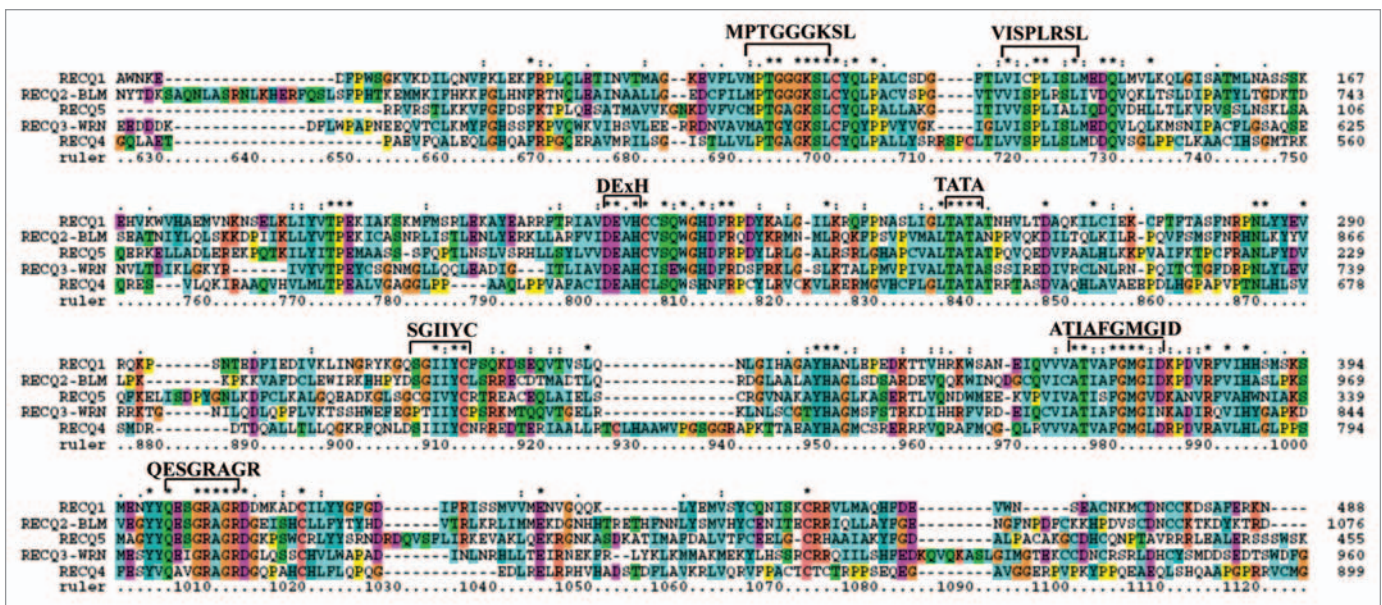


Figure 4. Protein alignment of five RecQ helicases in human. The five RecQ members are identified on left and sequence position of amino acid residues are shown on right. The seven conserved RecQ motifs are labeled on top of the alignment. Asterisks and dots drawn on top of sequence indicate identical residues and conservative amino acid changes, respectively. Gaps in the amino acid sequences are introduced to improve the alignment. Only a part of protein alignment with conserved helicase motifs is shown.

the genome of *E. coli* (*RecQ*), budding yeast (*Sgs1*) and fission yeast (*Rqh1*), whereas, 5 different homologs are found in the human genome.⁵⁸ The genomes of *A. thaliana* and *O. sativa* each encode 6 different RecQ family helicases.^{2,5} The RecQ enzymes can unwind various DNA substrates such as standard B-form DNA duplexes, triple helices, D-loops, 3- or 4-way junctions and G-quadruplex DNA.⁶³

RecQ helicases are also known as guardians of the genome because these DNA helicases have diverse roles in various DNA metabolic pathways which include DNA recombination, replication, repair, transcription and telomere maintenance. The high levels of genomic instability leading to premature aging and a high incidence of cancer result from mutations in 3 of the human RecQ genes.⁵⁸ The mutations in WRN, BLM and RECQ4 are linked to three recessive genetic disorders: Werner's, Bloom and Rothmund-Thomson syndromes.⁵⁸ To the best of our knowledge up to now there are no reports of association of any disease with RecQ1 and

RecQ5. The human RecQ4 can unwind forked duplexes and the unwinding is independent of strand annealing.⁶⁴ It was also demonstrated that a point mutation of conserved lysine in Walker A abolished human RecQ4 helicase activity.⁶⁴ Distinct roles for human RecQ1 and RecQ4 were proposed.⁶⁵ Followed with origin recognition complex (ORC) and MCM complex assembly, the human RecQ4 is recruited to origins at late G₁ phase while, RecQ1 and additional RecQ4 are loaded at origins at the onset of S phase during origin firing.⁶⁵ Depletion of RecQ1 from human cells by RNA interference resulted in elevated sister chromatid exchange and DNA damage sensitivity.⁶⁶ The RecQ5 gene in *Drosophila* and human undergoes alternative splicing and exists in different isoforms.⁶⁷ A proteomic approach has revealed the interaction of RecQ5 helicase with RNA polymerase II.⁶⁸ The recruitment of RecQ5 at the sites of DNA damage is mediated by the MRE11-RAD50-NBS1 (MRN) complex.⁶⁹ The RecQ5 physically interacts with the RAD51 recombinase and can disrupt RAD51 presynaptic

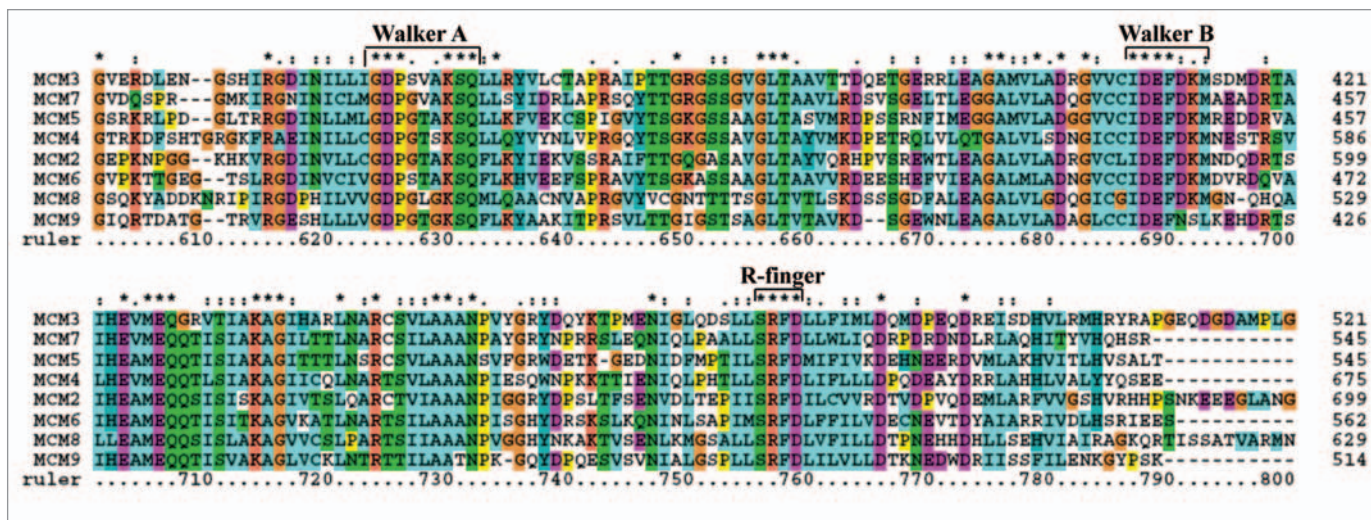


Figure 5. Protein alignment of MCM proteins in human. Amino acid alignment of canonical MCM2-7, and the unusual MCM8 and MCM9 proteins. The figure is constructed with ClustalX program, names of MCM proteins are given on left, and sequence positions on right. Asterisks and dots drawn on top of sequence indicate identical residues and conservative amino acid changes, respectively. Gaps in the amino acid sequences are introduced to improve the alignment. Only a part of protein alignment with conserved helicase motifs is shown. Note the presence of a glycine (G) instead of alanine (A) or serine (S) in the Walker A motif of MCM8 and MCM9 proteins.

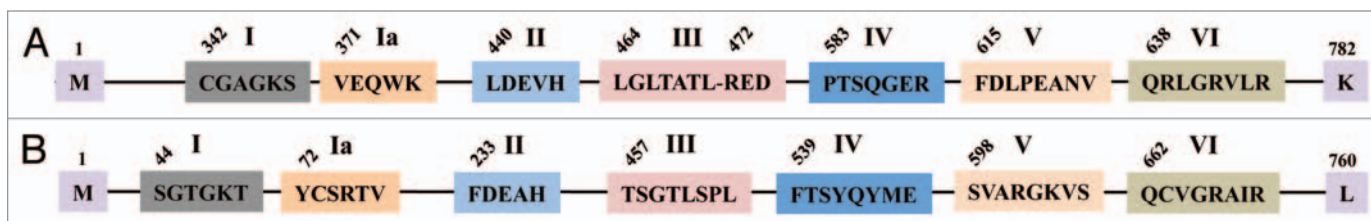


Figure 6. Schematic representation of seven conserved helicase motifs found in DNA repair helicases, XPB and XPD. The numbers given on top of the diagram indicate amino acid position of the first residue of a conserved motif for the entire protein sequence. (A) Amino acid consensus for motifs of XPB helicase. The sequences representing CGAGKS (motif I), VEQWK (motif Ia), LDEVH (motif II), LGLTATLxRED (motif III), PTSQGER (motif IV), FDLPEANV (motif V) and QRLGRVLR (motif VI) are shown. The first and last protein residues are methionine (M) and lysine (K) respectively are indicated. (B) Amino acid consensus for motifs of XPD helicase. The consensus sequence for these motifs is shown, SGTGKT (motif I), YCSRTV (motif Ia), FDEAH (motif II), TSGTLSPL (motif III), FTSYQYME (motif IV), SVARGKVS (motif V) and QCVGRAIR (motif VI). The first and last protein residues are methionine (M) and leucine (L) are shown.

filaments.⁷⁰ RecQ5 is involved in stabilization of replication fork wherein overexpressed RecQ5 promotes survival of human cells during thymidine-induced slowing of replication forks.⁷¹ In other species, mutants in the RecQ genes lead to unstable chromosomes.⁵⁸ Collectively, these findings indicate that RecQ helicases play a central role in the maintenance of genomic stability.

Sgs1 is a yeast homolog of the human BLM helicase. The loss of Sgs1 also causes elevated recombination, increased chromosome instability and DNA-damage sensitivity.⁷² Some of the associated phenotypes in *sgs1* deficient cell are suppressed by the expression of the human *BLM* or *WRN* gene.⁵⁸ By the process of unwinding the D-loops and/or by dissolving double Holliday junctions (dHJs) the Sgs1 is thought to disassemble joint molecules (JMs) as well.⁷³ The RecQ helicases from both yeast and human can interact with DNA topoisomerase III.^{74,75} The proteins BLM and Rad51 can also interact suggesting a possible link between RecQ helicases and homologous recombination in human cells.⁷⁶ The transcription analysis in a large array

of human tumor types showed RECQL4 to be the sixth most highly expressed gene.⁷⁷

MCM family genes have been found in all eukaryotes. The MCM-related proteins are present in archaea but MCM proteins do not exist in prokaryotes.⁷⁸ Our search criterion has identified 9 MCMs (MCM2-7, MCM8, MCM9 and MCM10) in human genome. Only 6 proteins, MCM2 through MCM7, are members of the MCM family. These six essential subunits are each AAA⁺ ATPases and form a heterohexameric complex.⁷⁹ The MCM2-7 complex acts as replicative helicase in vivo.⁸⁰ Both the replication initiation and fork regression requires MCM2-7 complex.⁸¹ In *Drosophila*, MCM4 is called disc proliferation abnormal,⁸² while *A. thaliana* MCM7 is also called as PROLIFERA.⁸³

Similar to MCM2-7, the MCM8 and MCM9 proteins also occur only in eukaryotes.⁸⁴ The MCM8 has been identified in humans and it contains the consensus IDEFDKM and arginine finger motifs.^{85,86} The occurrence of GKS motif in Walker A sequence is unique to MCM8. In addition to proliferating cells,

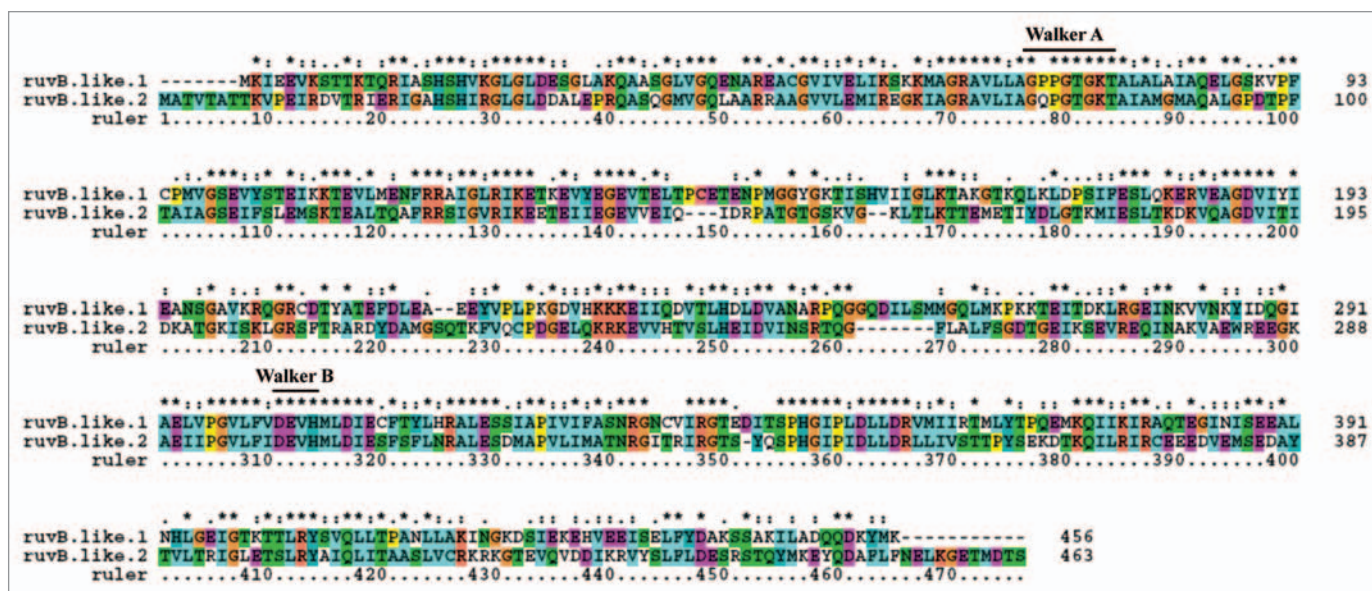


Figure 7. Protein alignment of members of human ‘RuvB-like’ family of helicases. The Walker A (G(x)₄GKT) and Walker B consensus (DEXH) are indicated on top of the alignment. The sequences are aligned using ClustalX program. Names of RuvB-like proteins are given on left, and sequence positions on right. Asterisks and dots drawn on top of sequence indicate identical residues and conservative amino acid changes, respectively. Gaps in the amino acid sequences are introduced to improve the alignment. Only a part of protein alignment with conserved helicase motifs is shown.

MCM8 is also expressed in a variety of tissues and throughout the cell cycle.^{85,86} Recombinant MCM8 displays both DNA helicase and ATPase activities in vitro.⁸⁷ The results also indicated involvement of human MCM8 in G₁ phase progression.⁸⁸ The human MCM8 has been proposed to be a crucial component of the pre-RC and its interaction in vivo with two components of the pre-RC, namely, cdc6 and ORC2 has been identified.⁸⁸ Depletion of human MCM8 protein leads to reduced loading of cdc6 and the MCM2-7 complex.⁸⁸ Human MCM9 protein also contains Walker A motif, a partially conserved Walker B motif, and an arginine finger and the phylogenetic analysis shows that MCM9 is more closely related to MCM8. The full length human MCM9 was described.⁸⁹ It was demonstrated that MCM9 binds to chromatin and is required for recruitment of MCM2-7 helicase onto chromatin.⁹⁰ The MCM9 has been suggested to be an essential linker between licensing factor Cdt1 and the MCM2-7 complex.⁹⁰

In spite of its name, the MCM10 protein is not a member of the canonical MCM2-7 family. The MCM2-7 complex and other replication factors are known to interact with MCM10 protein.⁹¹ MCM10 was isolated in a screen for replication mutants and also in the same *S. cerevisiae* screen as the true MCM proteins.⁹² Experiments with fission yeast, *Xenopus* extracts and human cells have indicated that MCM10 acts after association of canonical MCM2-7 with the pre-RC.⁹³ Direct interaction of human MCM10 with RecQ4 regulating its DNA unwinding activity was demonstrated.⁹⁴ It was shown that degradation of the 26 S proteasome caused due to UV-specific damage leads to MCM10 downregulation in mammalian cells.⁹⁵

The BLAST search has identified DNA repair helicases, XPB and XPD, in the human genome. XPB and XPD play crucial roles in NER and in transcription initiation.⁹⁶ The core TFIIH,

a basal transcription factor, contains these two helicases which are implicated in human autosomal recessive disorders such as xeroderma pigmentosum (XP), Cockayne syndrome (CS) and trichothiodystrophy (TTD).⁹⁷ These syndromes are caused due to defects in TFIIH which in turn might affect not only nucleotide excision repair (NER), but also transcription.⁹⁸ The DEAD 2 domain is also present in XPD and mutations in this domain lead to TTD disease.⁹⁹ Similarly, the mutations in human ATR-X gene result in X-linked alpha-thalassaemia with mental retardation (ATR-X) syndrome and correlate with changes in methylation of repetitive DNA sequences.¹⁰⁰

Our analysis has also identified DNA helicases like Pif1, RecQ5 beta, Twinkle, BACH1, nucleolin and RTEL1 encoded by the human genome. The Pif1 is a member of SF1 helicase conserved from yeast to human involved in several aspects of DNA replication and repair.^{11,101} RecQ5 beta, one among the 3 different isoforms of RecQ5, is a 3'-5' DNA helicase which can promote migration of Holliday junctions.¹⁰² Twinkle is a mitochondrial DNA helicase with 5' to 3' helicase activity, and possesses specified substrate requirements.¹⁰³ The autosomal dominant progressive ophthalmoplegia caused due to multiple deletions in the human mitochondrial DNA results from mutations in Twinkle.¹⁰³ Twinkle colocalizes with mtDNA in mitochondrial nucleoids.¹⁰⁴ The BACH1 (for BRCA1-associated C-terminal helicase), a member of SF2 helicase superfamily, is a 5' to 3' helicase which directly interacts with the BRCA1 BRCT motifs and is implicated in double strand break repair.^{105,106} The mutant BACH1 participates in breast cancer development.¹⁰⁷ The helicase nucleolin can unwind DNA-DNA duplexes, as well as RNA-RNA, and DNA-RNA duplexes, and is involved in several metabolic processes.^{108,109} The helicase RTEL1 play role in telomere homeostasis and was shown to disrupt D-loops in vitro.¹¹⁰

Table 3. Homologs for human RNA helicases in different model organisms

S. no.	Human	Rice [*]	Arabidopsis	Yeast	<i>P. falciparum</i> [†]
1	ASCC3L1	U5.snRNP (LOC_Os02g01740)	RH13 (Q93Y39)	BRR2 (P32639)	U5.snRNP (PFD1060w)
2	ATRX	SNF2 (LOC_Os10g31970)	PICKLE (Q9S775)	RDH54 (P38086)	PfSNF2L (PF11_0053)
3	DDX1	DEAD-box (LOC_Os01g07080)	RH37 (Q84W89)	DBP2 (P24783)	DEAD-box (PFE1085w)
4	DDX2A	DEAD-box (LOC_Os06g48750)	eIF4A-3 (Q9CAI7)	eIF4A (P10081)	helicase 45 (PF14_0655)
5	DDX2B	DEAD-box (LOC_Os02g05330)	eIF4A-2 (P41377)	eIF4A (P10081)	helicase 45 (PF14_0655)
6	DDX3X	DEAD-box 52B (LOC_Os07g10250)	RH52 (Q9M2F9)	DBP1 (P24784)	RNA helicase (PF08_0096)
7	DDX3Y	DEAD-box (LOC_Os11g38670)	RH52 (Q9M2F9)	DBP1 (P24784)	helicase (PF14_0437)
8	DDX4	rhIE (LOC_Os01g08930)	RH39 (Q56X76)	DBP2 (P24783)	helicase (PF14_0437)
9	DDX5	DEAD-box (LOC_Os01g10050)	RH20 (Q9C718)	DBP2 (P24783)	helicase (PF14_0437)
10	DDX6	dhh1 (LOC_Os02g42860)	RH8 (Q8RXK6)	DHH1 (P39517)	RNA helicase (PFC0915w)
11	DDX10	DBP4 (LOC_Os02g57980)	RH32 (Q9FFT9)	DBP4 (P20448)	DEAD/DEAH-box (PFF1500c)
12	DDX11	helicase (LOC_Os05g13300)	UVH6 (Q8W4M7)	CHL1 (P22516)	DNA-repair helicase (PF14_0081)
13	DDX12	helicase (LOC_Os05g13300)	UVH6 (Q8W4M7)	CHL1 (P22516)	DEAD-box (MAL13P1.134)
14	DDX13	DSHCT (LOC_Os02g06500)	RH29 (O49289)	SKI2 (P35207)	helicase (PFI0480w)
15	DDX17	recG (LOC_Os02g48100)	RH2 (Q94A52)	DBP5 (P20449)	helicase (PF14_0437)
16	DDX18	DDX18 (LOC_Os06g34420)	RH51 (Q9LIH9)	HAS1 (Q03532)	DEAD/DEAH-box (PFF1500c)
17	DDX19A	DBP5 (LOC_Os03g06220)	RH38 (Q93ZG7)	DBP5 (P20449)	helicase 45 (PF14_0655)
18	DDX19B	DEAD-box (LOC_Os03g06220)	RH38 (Q93ZG7)	DBP5 (P20449)	helicase 45 (PF14_0655)
19	DDX20	PICKLE (LOC_Os01g65850)	eIF4A-1 (P41376)	DHH1 (P39517)	helicase 45 (PF14_0655)
20	DDX21	NRH2 (LOC_Os03g61220)	RH3 (Q8L7S8)	DBP1 (P24784)	helicase (PFE0215w)
21	DDX23	DDX23 (LOC_Os03g50090)	RH21 (P93008)	PRP28 (P23394)	snrnp protein (PFE0925c)
22	DDX24	MAK5 (LOC_Os04g43140)	RH13 (Q93Y39)	MAK5 (P38112)	DEAD/DEAH-box (PFB0860c)
23	DDX25	DEAD-box (LOC_Os03g06220)	RH38 (Q93ZG7)	DBP5 (P20449)	helicase 45 (PF14_0655)
24	DDX27	DRS1 (LOC_Os12g29660)	RH28 (Q9ZRZ8)	DRS1 (P32892)	DEAD/DEAH-box (PFL2475w)
25	DDX28	DEAD-box (LOC_Os01g08930)	RH39 (Q56X76)	DBP8 (P38719)	snrnp protein (PFE0925c)
26	DDX31	DEAD-box (LOC_Os03g58810)	RH17 (Q7XJN0)	DBP7 (P36120)	DEAD-box (MAL7P1.113)
27	DDX39	DEAD-box (LOC_Os01g36920)	RH56 (Q9LFN6)	SUB2 (Q07478)	UAP56 (PFB0445c)
28	DDX41	DDX41 (LOC_Os06g48210)	RH35 (Q9LU46)	DBP2 (P24783)	RNA-helicase 1 (PFE1390w)
29	DDX42	DEAD-box (LOC_Os03g19530)	RH24 (Q22907)	DBP2 (P24783)	helicase (PF14_0437)
30	DDX43	DEAD-box (LOC_Os01g10050)	RH20 (Q9C718)	DBP2 (P24783)	helicase (PF14_0437)
31	DDX46	DDX46 (LOC_Os08g05810)	RH42 (Q8H0U8)	PRP5 (P21372)	RNA helicase (PFE0430w)
32	DDX47	DDX47 (LOC_Os03g46610)	RH10 (Q8GY84)	RRP3 (P38712)	DEAD/DEAH-box (PFB0860c)
33	DDX48	DEAD-box (LOC_Os03g36930)	RH2 (Q94A52)	FAL1 (Q12099)	eIF (PFD1070w)
34	DDX49	DDX49 (LOC_Os07g43980)	RH36 (Q9SA27)	DBP8 (P38719)	DEAD/DEAH-box (PFB0860c)
35	DDX50	DEAD-box 7 (LOC_Os09g34910)	RH7 (Q39189)	DBP1 (P24784)	helicase (PFE0215w)
36	DDX51	DDX51 (LOC_Os02g55260)	RH1 (Q7FGZ2)	DBP6 (P53734)	DEAD/DEAH-box (PFB0860c)
37	DDX52	DDX52 (LOC_Os07g45360)	RH57 (Q84TG1)	ROK1 (P45818)	helicase (PF14_0437)
38	DDX53	DEAD-box 30 (LOC_Os01g68320)	RH30 (Q8W4R3)	DBP2 (P24783)	helicase (PF14_0437)
39	DDX54	DDX54 (LOC_Os08g32090)	RH29 (O49289)	DBP10 (Q12389)	RNA helicase (MAL8P1.19)
40	DDX55	DEAD-box (LOC_Os01g07080)	RH18 (Q9FLB0)	SPB4 (P25808)	DEAD/DEAH-box (PFF1500c)
41	DDX56	DDX56 (LOC_Os03g51900)	RH16 (Q9SW44)	DBP9 (Q06218)	DEAD/DEAH-box (PFB0860c)
42	DDX58	Dicer (LOC_Os03g02970)	Dicer (Q9SP32)	MPH1 (P40562)	helicase (PF14_0437)
43	DDX59	DEAD-box 41 (LOC_Os02g10770)	RH41 (Q3EBD3)	DBP2 (P24783)	helicase (PF14_0437)
44	DDX60	DSHCT (LOC_Os02g06500)	RH57 (Q84TG1)	DOB1 (P47047)	DEAD/DEAH-box (PFI0165c)
45	DHX8	DHX8 (LOC_Os06g09280)	sf (Q22899)	PRP22 (P24384)	RNA helicase (PF10_0294)
46	DHX9	DEAD/DEAH-box (LOC_Os10g33275)	DHX15 (O22899)	YLR419W (Q06698)	RNA helicase (PFI0860c)

Table 3. Homologs for human RNA helicases in different model organisms (continued)

47	DHX15	pre-mRNA sf (LOC_Os03g19960)	DHX15 (O22899)	PRP43 (P53131)	RNA helicase (PFI0860c)
48	DHX16	DHX16 (LOC_Os03g17432)	DHX8 (Q38953)	PRP22 (P24384)	RNA helicase (PF10_0294)
49	DHX29	DEAD/DEAH-box (LOC_Os04g35260)	DHX8 (Q38953)	YLR419W (Q06698)	RNA helicase (PFI0860c)
50	DHX30	DEAD/DEAH-box (LOC_Os10g33275)	DHX15 (O22899)	YLR419W (Q06698)	RNA helicase (PF10_0294)
51	DHX32	pre-mRNA sf (LOC_Os03g19960)	DHX15 (O22899)	PRP43 (P53131)	RNA helicase (PFI0860c)
52	DHX33	DHX8 (LOC_Os06g09280)	DHX8 (Q38953)	PRP22 (P24384)	RNA helicase (PF10_0294)
53	DHX34	DEAD/DEAH-box (LOC_Os10g33275)	DHX8 (Q38953)	PRP2 (P20095)	RNA helicase (PF10_0294)
54	DHX35	DHX35 (LOC_Os01g11370)	DHX8 (Q38953)	PRP22 (P24384)	RNA helicase (PFI0860c)
55	DHX36	DHX36 (LOC_Os01g02884)	sf (O22899)	YLR419W (Q06698)	RNA helicase (PF10_0294)
56	DHX37	helicase (LOC_Os02g50370)	DHX8 (Q38953)	DHR1 (Q04217)	RNA helicase (PF10_0294)
57	DHX38	PRP16 (LOC_Os07g32430)	DHX8 (Q38953)	PRP16 (P15938)	splicing factor (MAL13P1.322)
58	DHX40	PRP16 (LOC_Os07g32430)	DHX8 (Q38953)	PRP22 (P24384)	RNA helicase (PFI0860c)
59	DHX57	DEAD/DEAH-box (LOC_Os10g33275)	DHX8 (Q38953)	YLR419W (Q06698)	RNA helicase (PF10_0294)
60	DHX58	Dicer (LOC_Os03g02970)	Dicer (Q9SP32)	MPH1 (P40562)	DEAD/DEAH-box (PFL2475w)
61	BAT1	DEAD-box (LOC_Os01g36920)	RH56 (Q9LFN6)	SUB2 (Q07478)	UAP56 (PFB0445c)
62	SUPV3L1	SUV3 (LOC_Os03g53500)	RH53 (Q9LUW5)	SUV3 (P32580)	DEAD-box (PFF1140c)
63	SKIV2L2	SKIV2L2 (LOC_Os12g18140)	ISE2 (B9DFG3)	DOB1 (P47047)	RNA helicase (PFF0100w)
64	FANCI	helicase (LOC_Os09g37920)	UVH6 (Q8W4M7)	CHL1 (P22516)	DNA-repair helicase (PF14_0081)

The human cells with siRNA-mediated knockdown of RETL1 show signs of elevated recombination.¹¹¹

Materials and Methods

Database search and enlistment of RNA and DNA helicases.

The following methodology is adopted for identification of almost all the RNA and DNA helicases that are encoded by the human genome. The *Homo sapiens* RefSeq protein database available as a link to NCBI/BLAST home (<http://www.ncbi.nlm.nih.gov>; <http://blast.ncbi.nlm.nih.gov>) is searched using the program BLASTP-Compare Protein Sequences against 'BLAST human sequences' resource. As an input sequence, the human prototype for the DEAD-box proteins, the translation initiation factor eIF4A-1 (Swiss-Prot Id: P60842) is used. A tentative list of putative candidates for human helicases, DDX and/or DDH RNA helicases is compiled. Specific protein sequences that belong to the category of DExD/H RNA helicases are further used with the progress of the search. The search is further refined with enlisting all the helicase proteins which belong to DExD/H RNA helicases. A comprehensive list is prepared with the Swiss-Prot Ids, name of the protein, synonyms and gene name for these RNA helicases.

A dataset representing DNA helicases from rice (<http://rice.plantbiology.msu.edu>)⁵ is used to identify DNA helicases encoded by the human genome. A BLAST against human RefSeq protein database on NCBI with individual sequences of rice DNA helicases is performed. The best hit is enlisted with its Swiss-Prot Ids, name of the protein, synonyms and gene name for these DNA helicases.

Validation of helicases. Protein sequences of human RNA and DNA helicases are validated based on the occurrence of conserved helicase motif signatures.^{3,5,7}

Protein alignment and phylogenetic studies. Protein sequences for RNA and DNA helicases are retrieved from NCBI database. These sequences are aligned with ClustalX using default parameters. The helicase motif signatures are verified based on alignment comparison and manual assignment of the motifs. Phylogenetic analysis is performed based on entire protein sequences with MEGA4 software. Phylogenies are built using the Neighbor-Joining (NJ) method.

Identification of homologs for helicase proteins in model organisms. To identify the homologs of human RNA and DNA helicases in yeast and Arabidopsis, the protein sequences are used for the BLAST search on NCBI against the Swiss-Prot protein sequences database with blastp (protein-protein BLAST) algorithm. The homologs of human helicases in rice are traced in the Rice Genome Annotation Project (RAP) database (<http://rice.plantbiology.msu.edu>). Similarly, homologs for these helicases in *P. falciparum* are identified in PlasmoDB: Malaria Parasite Genome Project (<http://plasmodb.org/plasmo>). A list of homologs is prepared with Swiss-Prot Ids for yeast and *A. thaliana*, Locus Ids for *O. sativa* and Gene Ids for *P. falciparum*.

Conclusions

The collections of RNA and DNA helicases identified during the present study will provide an invaluable resource for elaborate biological research on these helicases. We have also shown that in combination with protein alignment and phylogenetic studies, the identification of homologs using comparative genomic

Table 4. Homologs for DNA helicases in different model organisms

S. no.	Human	Rice*	Arabidopsis	Yeast	<i>P. falciparum</i> [†]
1	RECQ1	RECQ4A (LOC_Os11g48090)	RH20 (Q9C718)	SGS1 (P35187)	DNA helicase (PF14_0278)
2	BLM	RECQ4B (LOC_Os04g35420)	RH20 (Q9C718)	SGS1 (P35187)	DNA helicase (PFI0910w)
3	WRN	RecQSim (LOC_Os05g05810)	RH41 (Q3EBD3)	SGS1 (P35187)	DNA helicase (PF14_0278)
4	RECQ4	RH40 (LOC_Os04g40970)	RH30 (Q8W4R3)	SGS1 (P35187)	DNA helicase (PF14_0278)
5	RECQ5	RecQI3 (LOC_Os02g54020)	RH14 (Q8H136)	SGS1 (P35187)	DNA helicase (PFI0910w)
6	KU70	ku80 (LOC_Os07g08729)	ARP (P45951)	KU70 (P32807)	n/f
7	MCM2	MCM2 (LOC_Os11g29380)	MCM3 (Q9FL33)	MCM2 (P29469)	MCM2 (PF14_0177)
8	MCM3	MCM3 (LOC_Os05g39850)	MCM3 (Q9FL33)	MCM3 (P24279)	MCM3 (PFE1345c)
9	MCM4	MCM4 (LOC_Os01g36390)	PROL (P43299)	CDC54 (P30665)	MCM4-related (PF13_0095)
10	MCM5	MCM5 (LOC_Os02g55410)	MCM3 (Q9FL33)	MCM5 (P29496)	MCM5 (PFL0580w)
11	MCM6	MCM6 (LOC_Os05g14590)	PROL (P43299)	MCM6 (P53091)	Replication licensing factor (PF13_0291)
12	MCM7	MCM7 (LOC_Os12g37400)	PROL (P43299)	CDC47 (P38132)	MCM7 (PF07_0023)
13	MCM8	MCM8 (LOC_Os05g38850)	PROL (P43299)	MCM6 (P53091)	MCM protein (PFL0560c)
14	MCM9	MCM9 (LOC_Os06g11500)	MCM3 (Q9FL33)	CDC54 (P30665)	Replication licensing factor (PF13_0291)
15	MCM10	MCM10 (LOC_Os09g36820)	MCM10 (Q5XVE2)	MCM10 (P32354)	conserved Plasmodium protein (PF11_0207)
16	NCL	NCL (LOC_Os04g52960)	AT1G48920#	PUB1 (P32588)	RNA binding protein (PFI1175c)
17	CHD2	SNF2 (LOC_Os07g46590)	PKL (Q9S775)	CHD1 (P32657)	CHD1 (PF10_0232)
18	ERCC3	XPB2 (LOC_Os01g49680)	XPB2 (Q9FUG4)	RAD25 (Q00578)	DNA repair helicase rad25 (PF10_0369)
19	ERCC2	UVH6 (LOC_Os05g05260)	UVH6 (Q8W4M7)	RAD3 (P06839)	DNA excision-repair helicase (PFI1650w)
20	CHD7	CHD3 (LOC_Os06g08480)	ISW2 (Q8RWY3)	ISW2 (Q08773)	CHD1 (PF10_0232)
21	HELLS	SNF2 (LOC_Os03g22900)	BRM (Q6EVK6)	P43610	DEAD/DEAH box helicase (PFB0730w)
22	INOC1	SNF2 (LOC_Os03g22900)	INO80 (Q8RXS6)	INO80 (P53115)	ATP-dependent helicase (PF08_0048)
23	RuvB-like 1	ruvB-like (LOC_Os07g08170)	AT5G67630#	RUVBL1 (Q03940)	RuvB DNA helicase (PF11_0071)
24	RuvB-like 2	ruvB-like 2 (LOC_Os06g08770)	AT3G49830#	RUVBL2 (Q12464)	DNA helicase (PF13_0330)
25	PIF1	n/f	PIF1 (AT3G51690#)	PIF1 (P07271)	n/f
26	Twinkle	n/f	n/f	n/f	n/f
27	BACH1	helicase (LOC_Os09g37920)	UVH6 (Q8W4M7)	CHL1 (P22516)	DNA repair helicase (PF14_0081)
28	RecQ5 alpha	DEAD-box (LOC_Os02g54020)	AT4G35740#	SGS1 (P35187)	DNA helicase (PFI0910w)
29	RecQ5 beta	DEAD-box (LOC_Os02g54020)	AT4G35740#	SGS1 (P35187)	DNA helicase (PFI0910w)
30	RecQ5 gamma	DEAD-box (LOC_Os02g54020)	AT4G35740#	SGS1 (P35187)	DNA helicase (PFI0910w)
31	RTEL1	helicase (LOC_Os01g40980)	UVH6 (Q8W4M7)	RAD3 (P06839)	DNA repair helicase (PF14_0081)

Swiss-Prot Ids are given in brackets. n/f, not found. *Locus Ids are indicated in brackets for rice homologs (Rice Genome Annotation Project; rice.plantbiology.msu.edu). †Gene Ids are indicated in brackets for *P. falciparum* homologs (PlasmoDB: Malaria Parasite Genome Project; plasmodb.org/plasmo). #TAIR Locus Ids (www.arabidopsis.org).

approach in model organisms can be used effectively to carry out rapid characterization of these helicases as well. The phylogenetic analysis of different helicase categories could also represent an evolutionary upgrade in humans. In addition, the identified homologs in sequenced genomes such as *A. thaliana*, *O. sativa*, *S. cerevisiae* and *P. falciparum* will provide valuable insights on the evolutionary history of these helicases. It is interesting to note that no clear homologs for some of the putative human RNA and/or DNA helicases could be found in yeast suggesting that these proteins might perform distinct functions that are not present in this lower eukaryote. The study on comparing these homologs will spread light on the organization and complexity of helicase gene family in model organisms.

Acknowledgements

This work is partially supported by Department of Science and Technology, Department of Biotechnology and Defence Research and Development Organization grants. Infrastructural support from the Department of Biotechnology, Government of India is gratefully acknowledged.

Note

Supplementary materials can be found at: <http://www.landesbioscience.com/journals/cib/article/13844/>

References

- Compton SA, Tolun G, Kamath-Loeb AS, Loeb LA, Griffith JD. The Werner syndrome protein binds replication fork and holliday junction DNAs as an oligomer. *J Biol Chem* 2008; 283:24478-83.
- Hartung F, Plchova H, Puchta H. Molecular characterization of RecQ homologues in *Arabidopsis thaliana*. *Nucleic Acids Res* 2000; 28:4275-82.
- Tuteja N, Tuteja R. Prokaryotic and eukaryotic DNA helicases. Essential molecular motor proteins for cellular machinery. *Eur J Biochem* 2004; 271:1835-48.
- Tuteja R. Genome wide identification of *Plasmodium falciparum* helicases: a comparison with human host. *Cell Cycle* 2010; 9:104-20.
- Umate P, Tuteja R, Tuteja N. Genome-wide analysis of helicase gene family from rice and Arabidopsis: a comparison with yeast and human. *Plant Mol Biol* 2010; 73:449-65.
- Tanner NK, Cordin O, Banroques J, Doere M, Linder P. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol Cell* 2003; 11:127-38.
- Tuteja N, Tuteja R. Unravelling DNA helicases. Motif, structure, mechanism and function. *Eur J Biochem* 2004; 271:1849-63.
- Tanner NK, Linder P. DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol Cell* 2001; 8:251-62.
- Caruthers JM, McKay DB. Helicase structure and mechanism. *Curr Opin Struct Biol* 2002; 12:123-33.
- Hogbom M, Collins R, van den Berg S, Jenvert RM, Karlberg T, Kotenyova T, et al. Crystal structure of conserved domains 1 and 2 of the human DEAD-box helicase DDX3X in complex with the mononucleotide AMP. *J Mol Biol* 2007; 372:150-9.
- Gorbalenya AE, Koonin EV. Helicases: amino acid sequence comparisons and structure-function relationships. *Curr Opin Struct Biol* 1993; 3:419-29.
- Laggerbauer B, Achsel T, Luhrmann R. The human U5-200 kD DEXH-box protein unwinds U4/U6 RNA duplexes in vitro. *Proc Natl Acad Sci USA* 1998; 95:4188-92.
- Raghuathan PL, Guthrie C. RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr Biol* 1998; 8:847-55.
- Fairman ME, Maroney PA, Wang W, Bowers HA, Gollnick P, Nilsen TW, et al. Protein displacement by DExH/D "RNA helicases" without duplex unwinding. *Science* 2004; 304:730-4.
- Shen H, Zheng X, Shen J, Zhang L, Zhao R, Green MR. Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. *Genes Dev* 2008; 22:1796-803.
- Rozen F, Edery I, Meerovitch K, Dever TE, Merrick WC, Sonenberg N. Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4E. *Mol Cell Biol* 1990; 10:1134-44.
- Thomas AA, van Aalzum L, Voorma HO. RNA unwinding by eukaryotic initiation factor 4A and nucleotide modification. *Biochem Int* 1992; 27:17-23.
- Wang Y, Wagner JD, Guthrie C. The DEAH-box splicing factor Ppp16 unwinds RNA duplexes in vitro. *Curr Biol* 1998; 8:441-51.
- Yang Q, Jankowsky E. ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry* 2005; 44:13591-601.
- Rocak S, Linder P. DEAD-box proteins: The driving forces behind RNA metabolism. *Nat Rev Mol Cell Biol* 2004; 5:232-41.
- Cordin O, Banroques J, Tanner NK, Linder P. The DEAD-box protein family of RNA helicases. *Gene* 2006; 367:17-37.
- Linder P. Dead-box proteins: A family affair-active and passive players in RNP-remodeling. *Nucleic Acids Res* 2006; 34:4168-80.
- Tuteja N, Tuteja R. DNA helicases as molecular motors: an insight. *Physica Acta* 2006; 372:70-83.
- de la Cruz J, Kressler D, Linder P. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends Biochem Sci* 1999; 24:192-8.
- Daugeron MC, Linder P. Are DEAD-box proteins becoming respectable helicases? *Nat Struct Biol* 2000; 7:97-9.
- Jankowsky E, Gross CH, Shumann S, Pyle AM. The DEXH protein NPH-II is a processive and directional motor for unwinding RNA. *Nature* 2000; 403:447-51.
- Aubourg S, Kreis M, Lecharny A. The DEAD-box RNA helicase family in *Arabidopsis thaliana*. *Nucleic Acids Res* 1999; 27:628-36.
- Boudet N, Aubourg S, Toffano-Nioche C, Kreis M, Lecharny A. Evolution of intron/exon structure of DEAD helicase family genes in Arabidopsis, Caenorhabditis and Drosophila. *Genome Res* 2001; 11:2101-14.
- Abdelhaleem M, Maltais L, Wain H. The human DDX and DHX gene families of putative RNA helicases. *Genomics* 2003; 81:618-22.
- Staley JP, Guthrie C. Mechanical devices of the spliceosome: Motors, clocks, springs and things. *Cell* 1998; 92:315-26.
- Gencheva M, Kato M, Newo AN, Lin RJ. Contribution of DEAH-box protein DHX16/hPRP2 in human pre-mRNA splicing. *Biochem J* 2010; 429:25-32.
- Pisareva VP, Pisarev AV, Komar AA, Hellen CU, Pestova TV. Translation initiation on mammalian mRNAs with structured 5' UTRs requires DExH-box protein DHX29. *Cell* 2008; 135:1237-50.
- Chakraborty P, Grosse F. WRN helicase unwinds Okazaki fragment-like hybrids in a reaction stimulated by the human DHX9 helicase. *Nucleic Acids Res* 2010; 38:4722-30.
- Huang M, Mitchell BS. Guanine nucleotide depletion mediates translocation of nucleolar proteins, including RNA helicase A (DHX-9). *Nucleosides Nucleotides* 2008; 27:704-11.
- Alli Z, Ackerley C, Chen Y, Al-Saud B, Abdelhaleem M. Nuclear and mitochondrial localization of the putative RNA helicase DHX32. *Exp Mol Pathol* 2006; 81:245-8.
- Chen Y, Alli Z, Ackerley C, Al-Saud B, Abdelhaleem M. Altered distribution of heat shock protein 60 (Hsp60) with dysregulated expression of DHX32. *Exp Mol Pathol* 2007; 82:256-61.
- Fu JJ, Li LY, Lu GX. Molecular cloning and characterization of human DDX36 and mouse Ddx36 genes, New members of the DEAD/H box superfamily. *Acta Biochim Biophys Sinica* 2002; 34:655-61.
- Vaughn JP, Creacy SD, Routh ED, Joyner-Burt C, Jenkins GS, Pauli S, et al. The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates. *J Biol Chem* 2005; 280:38117-20.
- Lee JH, Rho SB, Chun T. GABAA receptor-associated protein (GABARAP) induces apoptosis by interacting with DEAD (Asp-Glu-Ala-Asp/His) box polypeptide 47 (DDX 47). *Biotechnol Lett* 2005; 27:623-8.
- Uhlmann-Schiffler H, Jalal C, Stahl H. Ddx42p—a human DEAD box protein with RNA chaperone activities. *Nucleic Acids Res* 2006; 34:10-22.
- Mulhern O, Bowie AG. Unexpected roles for DEAD-box protein 3 in viral RNA sensing pathways. *Eur J Immunol* 2010; 40:933-5.
- Chen CY, Yedavalli VR, Jeang KT. A method to study the role of DDX3 RNA helicase in HIV-1. *Methods Mol Biol* 2010; 587:281-9.
- Lee CS, Dias AP, Jedrychowski M, Patel AH, Hsu JL, Reed R. Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. *Nucleic Acids Res* 2008; 36:4708-18.
- Franca R, Belfiore A, Spadari S, Maga G. Human DEAD-box ATPase DDX3 shows a relaxed nucleoside substrate specificity. *Proteins* 2007; 67:1128-37.
- Chao CH, Chen CM, Cheng PL, Shih JW, Tsou AP, Lee YH. DDX3, a DEAD box RNA helicase with tumor growth-suppressive property and transcriptional regulation activity of the p21^{waf1/cip1} promoter, is a candidate tumor suppressor. *Cancer Res* 2006; 66:6579-88.
- Rosner A, Rinkevich B. The DDX3 subfamily of the DEAD box helicases: divergent roles as unveiled by studying different organisms and in vitro assays. *Curr Med Chem* 2007; 14:2517-25.
- Botlagunta M, Vesuna F, Mironchik Y, Raman A, Lisok A, Winnard P Jr, et al. Oncogenic role of DDX3 in breast cancer biogenesis. *Oncogene* 2008; 27:3912.
- Yedavalli VS, Neuveut C, Chi YH, Kleiman L, Jeang KT. Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. *Cell* 2004; 119:381-92.
- Abdelhaleem M. RNA helicases: regulators of differentiation. *Clin Biochem* 2005; 38:499-503.
- Rauschendorf MA, Zimmer J, Hanstein R, Dickemann C, Vogt PH. Complex transcriptional control of the AZFa gene DDX3Y in human testis. *Int J Androl* 2011; 34:84-96.
- Margossian SP, Li H, Zassenhaus HP, Butow RA. The DEXH box protein Suv3p is a component of yeast mitochondrial 3'-to-5' exoribonuclease that suppresses group I intron toxicity. *Cell* 1996; 84:199-209.
- Minczuk M, Piwowarski J, Papworth MA, Awiszus K, Schalinski S, Dziembowski A, et al. Localisation of the human Suv3p helicase in the mitochondrial matrix and its preferential unwinding of dsDNA. *Nucleic Acids Res* 2002; 30:5074-86.
- Dziembowski A, Piwowarski J, Hoser R, Minczuk M, Dmochowska A, Siep M, et al. The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism. *J Biol Chem* 2003; 278:1603-11.
- Khidr L, Wu G, Davila A, Procaccio V, Wallace D, Lee WH. Role of SUV3 helicase in maintaining mitochondrial homeostasis in human cells. *J Biol Chem* 2008; 283:27064-73.
- Shu Z, Vijayakumar S, Chen CF, Chen PL, Lee WH. Purified human SUV3p exhibits multiple-substrate unwinding activity upon conformational change. *Biochemistry* 2004; 43:4781-90.
- Pereira M, Mason P, Szczesny RJ, Maddukuri L, Dziwura S, Jedrzejczak R, et al. Interaction of human SUV3 RNA/DNA helicase with BLM helicase; loss of the SUV3 gene results in mouse embryonic lethality. *Mech Ageing Dev* 2007; 128:609-17.
- Dziembowski A, Malewicz M, Minczuk M, Golik P, Dmochowska A, Stepien PP. The yeast nuclear gene DSS1, which codes for a putative RNase II, is necessary for the function of the mitochondrial degradosome in processing and turnover of RNA. *Mol Gen Genet* 1998; 260:108-14.
- Bjergbaek L, Cobb JA, Gasser SM. RecQ helicases and genome stability: lessons from model organisms and human disease. *Swiss Med Wkly* 2002; 132:433-42.
- Shen JC, Gray MD, Oshima J, Kamath-Loeb AS, Fry M, Loeb LA. Werner syndrome protein. I. DNA helicase and dna exonuclease reside on the same polypeptide. *J Biol Chem* 1998; 273:34139-44.
- Puranam KL, Blackshear PJ. Cloning and characterization of RECQL, a potential human homologue of the *Escherichia coli* DNA helicase RecQ. *J Biol Chem* 1994; 269:29838-45.
- Yan H, Chen CY, Kobayashi R, Newport J. Replication focus-forming activity 1 and the Werner syndrome gene product. *Nat Genet* 1998; 19:375-8.
- Cogoni C, Macino G. Posttranscriptional gene silencing in Neurospora by a RecQ DNA helicase. *Science* 1999; 286:2342-4.
- Pike ACW, Shrestha B, Popuri V, Burgess-Brown N, Muzzolini L, Costantini S, et al. Structure of the human RECQ1 helicase reveals a putative strand-separation pin. *Proc Natl Acad Sci USA* 2009; 106:1039-44.

64. Rossi ML, Ghosh AK, Kulikowicz T, Croteau DL, Bohr VA. Conserved helicase domain of human RecQ4 is required for strand annealing-independent DNA unwinding. *DNA Repair (Amst)* 2010; 9:796-804.
65. Thangavel S, Mendoza-Maldonado R, Tissino E, Sidorova JM, Yin J, Wang W, et al. Human RECQ1 and RECQ4 helicases play distinct roles in DNA replication initiation. *Mol Cell Biol* 2010; 30:1382-96.
66. Bugreev DV, Brosh RM Jr, Mazin AV. RECQ1 possesses DNA branch migration activity. *J Biol Chem* 2008; 283:20231-42.
67. Sekelsky JJ, Brodsky MH, Rubin GM, Hawley RS. *Drosophila* and human RecQ5 exist in different isoforms generated by alternative splicing. *Nucleic Acids Res* 1999; 27:3762-9.
68. Aygün O, Svejstrup J, Liu Y. A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proc Natl Acad Sci USA* 2008; 105:8580-4.
69. Zheng L, Kanagaraj R, Mihaljevic B, Schwendener S, Sartori AA, Gerrits B, et al. MRE11 complex links RECQ5 helicase to sites of DNA damage. *Nucleic Acids Res* 2009; 37:2645-57.
70. Schwendener S, Raynard S, Paliwal S, Cheng A, Kanagaraj R, Shevelev I, et al. Physical interaction of RECQ5 helicase with RAD51 facilitates its anti-recombinase activity. *J Biol Chem* 2010; 285:15739-45.
71. Blundred R, Myers K, Helleday T, Goldman AS, Bryant HE. Human RECQL5 overcomes thymidine-induced replication stress. *DNA Repair (Amst)* 2010; 9:964-75.
72. Watt PM, Hickson ID, Borts RH, Louis EJ. SGS1, a homologue of the Bloom's and Werner's syndrome genes, is required for maintenance of genome stability in *Saccharomyces cerevisiae*. *Genetics* 1996; 144:935-45.
73. van Brabant AJ, Ye T, Sanz M, German IJ, Ellis NA, Holloman WK. Binding and melting of D-loops by the Bloom syndrome helicase. *Biochemistry* 2000; 39:14617-25.
74. Angloff S, McDonald JP, Bendixen C, Arthur L, Rothstein R. The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol Cell Biol* 1994; 14:8391-8.
75. Wu L, Davies SL, North PS, Goulaouic H, Riou JF, Turley H, et al. The Bloom's syndrome gene product interacts with topoisomerase III. *J Biol Chem* 2000; 275:9636-44.
76. Wu L, Davies SL, Levitt NC, Hickson ID. Potential role for the BLM helicase in recombinational repair via a conserved interaction with RAD51. *J Biol Chem* 2001; 276:19375-81.
77. Velculescu VE, Madden SL, Zhang L, Lash AE, Yu J, Rago C, et al. Analysis of human transcriptomes. *Nat Genet* 1999; 23:387-8.
78. Kearsley SE, Labib K. MCM proteins: evolution, properties and role in DNA replication. *Biochim Biophys Acta* 1998; 1398:113-36.
79. Forsburg SL. Eukaryotic MCM proteins: beyond replication initiation. *Microbiol Mol Biol Rev* 2004; 68:109-31.
80. Bell SP, Dutta A. DNA replication in eukaryotic cells. *Annu Rev Biochem* 2002; 71:333-74.
81. Labib K, Kearsley SE, Diffley JF. MCM2-7 proteins are essential components of prereplicative complexes that accumulate cooperatively in the nucleus during G₁-phase and are required to establish, but not maintain, the S-phase checkpoint. *Mol Biol Cell* 2001; 12:3658-67.
82. Feger G, Vaessin H, Su TT, Wolff E, Jan LY, Jan YN. dpa, a member of the MCM family, is required for mitotic DNA replication but not endoreplication in *Drosophila*. *EMBO J* 1995; 14:5387-98.
83. Springer PS, McCombie WR, Sundaresan V, Martienssen RA. Gene trap tagging of PROLIFERA and essential MCM2-3-5 like gene in *Arabidopsis*. *Science* 1995; 268:877-80.
84. Maiorano D, Lutzmann M, Méchali M. MCM proteins and DNA replication. *Curr Opin Cell Biol* 2006; 18:130-6.
85. Gozuacik D, Chami M, Lagorce D, Faivre J, Murakami Y, Poch O, et al. Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. *Nucleic Acids Res* 2003; 31:570-9.
86. Johnson EM, Kinoshita Y, Daniel DC. A new member of the MCM protein family encoded by the human MCM8 gene, located contrapodal to GCD10 at chromosome band 20p12.3-13. *Nucleic Acids Res* 2003; 31:2915-25.
87. Maiorano D, Cuvier O, Danis E, Méchali M. MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell* 2005; 120:315-28.
88. Volkening M, Hoffmann I. Involvement of human MCM8 in prereplication complex assembly by recruiting hcdc6 to chromatin. *Mol Cell Biol* 2005; 25:1560-8.
89. Lutzmann M, Maiorano D, Méchali M. Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family. *Gene* 2005; 362:51-6.
90. Lutzmann M, Méchali M. MCM9 binds Cdt1 and is required for the assembly of prereplication complexes. *Mol Cell* 2008; 31:190-200.
91. Homesley L, Lei M, Kawasaki Y, Sawyer S, Christensen T, Tye BK. Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins. *Genes Dev* 2000; 14:913-26.
92. Maine GT, Sinha P, Tye BK. Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics* 1984; 106:365-85.
93. Wohlschlegel JA, Dhar SK, Prokhorova TA, Dutta A, Walter JC. Xenopus Mcm10 binds to origins of DNA replication after Mcm2-7 and stimulates origin binding of Cdc45. *Mol Cell* 2002; 9:1-20.
94. Xu X, Rochette PJ, Feyissa EA, Su TV, Liu Y. MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. *EMBO J* 2009; 28:3005-14.
95. Sharma A, Kaur M, Kar A, Ranade SM, Saxena S. Ultraviolet radiation stress triggers the downregulation of essential replication factor MCM10. *J Biol Chem* 2010; 285:8352-62.
96. Guzder SN, Sung P, Bailly V, Prakash L, Prakash S. RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. *Nature* 1994; 369:578-81.
97. Sung P, Bailly V, Weber C, Thompson LH, Prakash L, Prakash S. Human xeroderma pigmentosum group D gene encodes a DNA helicase. *Nature* 1993; 365:852-5.
98. Zurita M, Merino C. The transcriptional complexity of the TFIIH complex. *Trends Genet* 2003; 19:578-84.
99. Dubaele S, Proietti De Santis L, Bienstock RJ, Keriel A, Stefanini M, Van Houten B, et al. Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Mol Cell* 2003; 11:1635-46.
100. Gibbons R. Alpha thalassaemia-mental retardation, X linked. *Orphanet J Rare Dis* 2006; 1:15.
101. Ivessa AS, Zhou JQ, Zakian VA. The *Saccharomyces Pif1p* DNA helicase and the highly related Rrm3p have opposite effects on replication fork progression in ribosomal DNA. *Cell* 2000; 100:479-89.
102. Garcia PL, Liu Y, Jiricny J, West SC, Janscak P. Human RECQ5b, a protein with DNA helicase and strand-annealing activities in a single polypeptide. *EMBO J* 2004; 23:2882-91.
103. Korhonen JN, Gaspari M, Falkenberg M. TWINKLE has 5'-3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J Biol Chem* 2003; 278:48627-32.
104. Spelbrink JN, Li FY, Tiranti V, Nikali K, Yuan QP, Tariq M, et al. Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nat Genet* 2001; 28:223-31.
105. Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, et al. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* 2001; 105:149-60.
106. Gupta R, Sharma S, Doherty KM, Sommers JA, Cantor SB, Brosh RM Jr. Inhibition of BACH1 (FANCD1) helicase by backbone discontinuity is overcome by increased motor ATPase or length of loading strand. *Nucleic Acids Res* 2006; 34:6673-83.
107. Cantor S, Drapkin R, Zhang F, Lin Y, Han J, Pamidi S, et al. The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proc Natl Acad Sci USA* 2004; 101:2357-62.
108. Tuteja N, Huang NW, Skopac D, Tuteja R, Hrvatic S, Zhang J, et al. Human DNA helicase IV is nucleolin, an RNA helicase modulated by phosphorylation. *Gene* 1995; 160:143-8.
109. Tuteja R, Tuteja N. Nucleolin: a multifunctional major nucleolar phosphoprotein. *Crit Rev Biochem Mol Biol* 1998; 33:407-36.
110. Barber LJ, Youds JL, Ward JD, McIlwraith MJ, O'Neil NJ, Petalcorin MI, et al. RTEL1 maintains genomic stability by suppressing homologous recombination. *Cell* 2008; 135:261-71.
111. Chavez A, Tsou AM, Johnson FB. Telomeres do the (un)twist: Helicase actions at chromosome termini. *Biochim Biophys Acta* 2009; 1792:329-40.