An RNA Metabolism and Surveillance Quartet in the Major Histocompatibility Complex

Danlei Zhou [1,2,3,*] **, Michalea Lai** [2,3]**, Aiqin Luo** [1,*] **and Chack-Yung Yu** [2,3,*]

1   School of Life Science, Beijing Institute of Technology, Beijing 100081, China
2   The Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH 43205, USA
3   Department of Pediatrics, The Ohio State University, Columbus, OH 43205, USA
*   Correspondence: Danlei.Zhou@nationwidechildrens.org (D.Z.); Bitluo@bit.edu.cn (A.L.);
    Chack-Yung.Yu@nationwidechildrens.org (C.-Y.Y.)

Abstract: At the central region of the mammalian major histocompatibility complex (MHC) is a complement gene cluster that codes for constituents of complement C3 convertases (C2, factor B and C4). Complement activation drives the humoral effector functions for immune response. Sandwiched between the genes for serine proteinase factor B and anchor protein C4 are four less known but critically important genes coding for essential functions related to metabolism and surveillance of RNA during the transcriptional and translational processes of gene expression. These four genes are NELF-E (RD), SKIV2L (SKI2W), DXO (DOM3Z) and STK19 (RP1 or G11) and dubbed as NSDK. NELF-E is the subunit E of negative elongation factor responsible for promoter proximal pause of transcription. SKIV2L is the RNA helicase for cytoplasmic exosomes responsible for degradation of de-polyadenylated mRNA and viral RNA. DXO is a powerful enzyme with pyro-phosphohydrolase activity towards 5′ triphosphorylated RNA, decapping and exoribonuclease activities of faulty nuclear RNA molecules. STK19 is a nuclear kinase that phosphorylates RNA-binding proteins during transcription. STK19 is also involved in DNA repair during active transcription and in nuclear signal transduction. The genetic, biochemical and functional properties for NSDK in the MHC largely stay as a secret for many immunologists. Here we briefly review the roles of (a) NELF-E on transcriptional pausing; (b) SKIV2L on turnover of deadenylated or expired RNA 3′→5′ through the Ski-exosome complex, and modulation of inflammatory response initiated by retinoic acid-inducible gene 1-like receptor (RLR) sensing of viral infections; (c) DXO on quality control of RNA integrity through recognition of 5′ caps and destruction of faulty adducts in 5′→3′ fashion; and (d) STK19 on nuclear protein phosphorylations. There is compelling evidence that a dysregulation or a deficiency of a NSDK gene would cause a malignant, immunologic or digestive disease.

Keywords: DXO; DOM3Z; NELF-E; RD; SKIV2L; SKI2W; STK19; RP1; NSDK; RLR; miR1236; SVA; RNA quality control; 5′→3′ RNA decay; 3′→5′ mRNA turnover; antiviral immunity; interferon β; promoter-proximal transcriptional pause; exosomes; nuclear kinase; hepatocellular carcinoma; Ski complex; trichohepatoenteric syndrome; melanoma

1. Introduction

In 1984, Professor Rodney Porter's team in Oxford reported the cloning and confirmed the physical linkage of the genes for human complement C4A and C4B, factor B and C2 through overlapping cosmids [1]. These genes are located at the central region of the human major histocompatibility complex (MHC), a region that is associated with numerous complex autoimmune

diseases [2–7]. They code for the essential constituents of the classical, lectin and alternative pathways C3 convertases. The complement system is a major effector arm for both innate and adaptive immune systems [8]. While C2 and factor B are only separated by an intergenic distance of 421 bp [9], it was a bit puzzling that the genes for C4 and factor B are separated by a large gap of 30 kb. Thus, it was satisfying when a genomic DNA probe upstream of the C4 promoter region successfully fished out an 1.1 kb cDNA clone that was dubbed as RP (to memorialize the late Professor Rodney Porter) [10,11]. Working from the other end, Meo's team reported the cloning of RD that is up to 149 bp tail-to-tail with factor B [12]. The protein sequence of RD is fascinating because of its 23 consecutive copies of positively charged arginine (R) and negatively charged aspartic acid (D), not to mention its leucine zipper and RNA binding motif. Between RD and RP, we found two more genes, SKI2W and DOM3Z [13–15].

Through biochemical and molecular biologic research to decipher their functions, these four "novel" genes turn out to be extremely old in evolution. Except for RP, structurally and functionally related genes and proteins can be traced all the way back to yeast or metazoans. The modern names for these four genes are NELF-E, SKIV2L, DXO and STK19 (NSDK), respectively. Unlike the Class I and Class II and complement that have distinct immunologic functions involving proteins, NSDK seems to play fundamental life functions in eukaryotes and actively engaged in the surveillance of RNA integrity including the 5' cap and 3' tail. In the process of cell metabolism, these four genes play critical roles on transcriptional control, RNA turnover and signaling. They may also fine-tune the innate immune response or type I interferon-mediated inflammation elicited by infections of RNA viruses. Here we give a brief account on the molecular genetics and biology of this RNA metabolism and surveillance "quartet" in the MHC (Figure 1, Table 1).
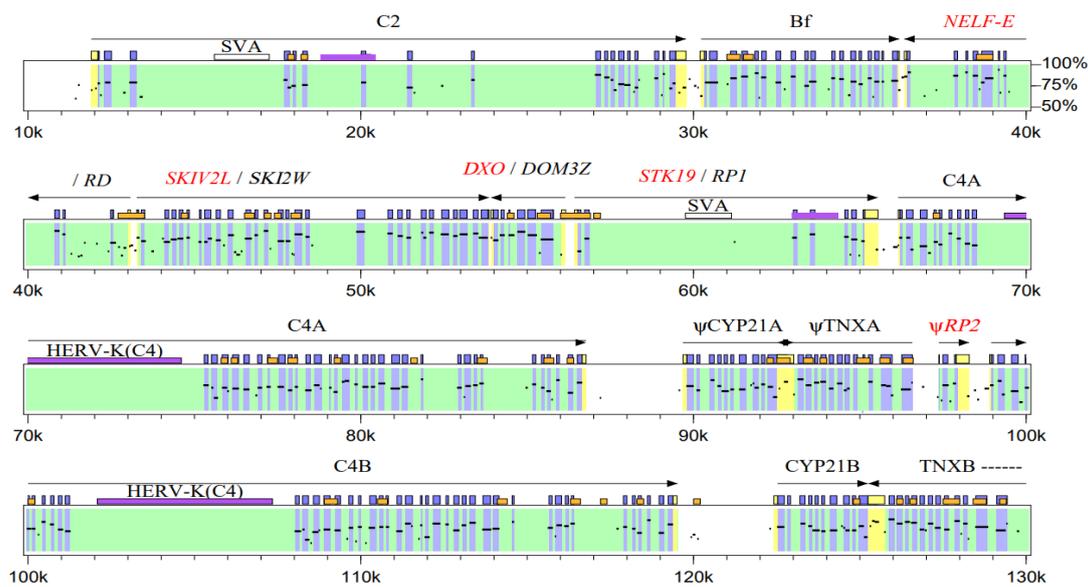
(A)

Figure 1. *Cont.*
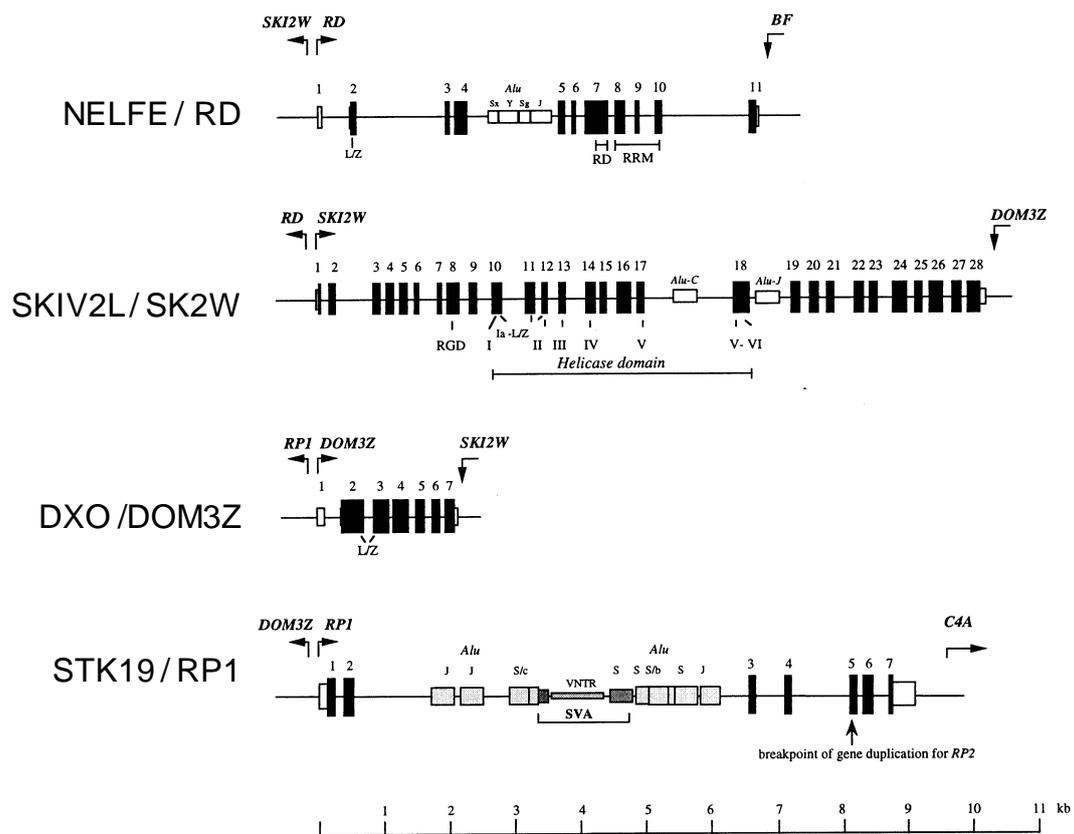
(B)

Figure 1. Map and gene structures of the NSDK (NELF-E, SKIV2L, DOX and STK19) quartet in the major histocompatibility complex (MHC) complement gene cluster. (**A**) A molecular map comparing the gene organization of human and mouse MHC complement gene cluster with RNA surveillance quartet consisting of NELF-E, SKIV2L, DXO and STK19 at the intergenic region between complement factor B and complement C4A gene. Original gene names are shown as alternatives. The mouse sequences with identities from 50–100% are plotted below the human sequences. Exons for human genes are shown as solid, purple boxes. The genomic region spanned for each gene is shaded green. The 5′ and 3′ untranslated sequences are in yellow boxes. The conserved coding sequences between human and mouse genes are highlighted in purple. Locations of CpG rich dinucleotides are shown as orange boxes. Large, repetitive retroelements SVA and endogenous retrovirus HERV-K(C4) are shown in burgundy. Arrows with solid lines represent configurations of structural genes. *RP2* and *TNXA* are partially duplicated gene fragments. *CYP21A* is a pseudogene in human. Pseudogenes and gene fragments are labeled with ψ in front of gene names. Genomic DNA sequences are obtained from the following accession numbers: Human, U89335-U89337, AF019413, M59815, M59816, L26260-L26263, U07856, AF059675 and AF077974; mouse, AF030001, AF049850 and AF109906. Numberings below box represents length in human genomic DNA in kb (2k = 2 kb). The 5′ region of complement C2 gene is not well defined. Located approx. 26.4 kb upstream of the gene for C2 isoform 5 and in opposite orientation is the zinc finger and BTB domain containing protein 12 (ZBTB12). Present at the 3′ end of the C2 gene and in opposite orientation is a genetic element for long non-coding RNA (lncRNA C2-AS1, not shown). (**B**) Exon-intron structures of human NELF-E, SKIV2L, DXO and STK19 (modified from refs [15,16]).

Table 1. Characteristics of the NSDK genes for RNA metabolism and surveillance quartet in the class III region of the human major histocompatibility complex.

Official Symbol	Original Name; Other Earlier Names	Gene, Transcripts and Proteins (Location, cDNA, aa)	Expression	Specific Features	Function	Website
NELF-E	**RD** RDP; RDBP; D6S45	11 exons 6.7 kb 380 aa	Ubiquitous; highest in testis	CpG rich, leucine zipper motif, RD repeats, RNA recognition motif (RRM); miR-1236 in intron 3	A subunit of the negative RNA elongation factor that represses transcriptional elongation by RNA polymerase II	https://www.ncbi.nlm.nih.gov/gene/7936
SKIV2L	**SKI2W** HLP, SKI2, 170A, DDX13, DDX49, SKIV2, SKIV2L1, THES2	28 exons 11 kb 1246 aa	Ubiquitous; highest in spleen	CpG rich, leucine zipper motif, helicase domain, RGD and Alu element	Unwinds RNA secondary structures, RNA turnover, antiviral defense, modulates type 1 interferon expression	https://www.ncbi.nlm.nih.gov/gene/6499
DXO	**DOM3Z** DOM3L, NG6, RAI1	7 exons 2.2 kb 396 aa	Ubiquitous; highest in testis and adrenal gland	CpG rich; leucine zipper motif	RNA quality control, decapping and 5′→3′ RNA decay	https://www.ncbi.nlm.nih.gov/gene/1797
STK19	**RP1** and **RP2** HLA-RP1, G11, D6S60, D6S60E	9 exons 11 kb 364 aa or 7 exons 9.1 kb 254 aa	Ubiquitous; highest in adrenal gland	Partial gene duplication in the RCCX modules; SVA (SINE - 21 CpG rich VNTRs-Alu) in intron; 8 Alu elements	Nuclear Serine/Threonine kinase	https://www.ncbi.nlm.nih.gov/gene/8859

2. NELF-E or RD

To Start a Long Journey, Get Prepared, Do a Test Run, Pause to Check Things Out, Modify and Let Go:
The Amazing RD or Negative Elongation Factor Subunit E (NELF-E) in Gene Transcription

The human RD gene, also known as NELF-E and D6S45, is located between complement factor B (Bf) and the SKIV2L gene. This gene was discovered by Meo and colleagues in 1988 [12]. RD is organized in a tail-to-tail configuration with Bf (Figure 1). There are 205 nucleotides between the 3′ ends of RD and Bf. The 5′ regulatory region of the RD gene is rich in CpG sequences [17]. Transcripts for RD are ubiquitously expressed but the highest expression level is found in the testis.

The RD gene is 6.7 kb in size and consists of 11 exons. It encodes 380 amino acids. The calculated molecular weight of the RD protein is 46 kDa [17,18]. The RD protein contains a leucine-zipper motif at the N terminal region, which facilitates interactions with other proteins [12,17,19]. A striking feature of RD, however, is a tract of tandem repeats with positively and negatively charged amino acids, which are predominantly Arg-Asp (RD), from residues 184 to 243, which are encoded by exon 7. Close to the carboxyl terminal region is an RNA recognition motif (RRM), which is located between residues 264 and 327 and encoded by exons 8–10. Thus, RD belongs to ribonucleoprotein family. The high degree of sequence similarities between human and murine RD both at DNA and amino acid levels suggested functional importance [12,15,17,20].

One feature for the regulation of gene expression in metazoans is the promoter-proximal transcriptional pausing of RNA polymerase II. This phenomenon is carried out by two complexes known as the negative elongation factor (NELF) and DRB (5,6-dichloro-1-β-ᴅ-ribofuranosyl-benzimidazole) sensitivity inducing factor (DSIF). NELF acts cooperatively with DSIF to strongly represses transcriptional elongation by RNA polymerase II (RNAPII), about 25–50 nucleotides downstream of transcriptional initiation, particularly in the absence of P-TEFb (positive transcription elongation factor b) [21]. Promoter proximal pausing of transcription is a universal process for RNA polymerase II transcripts that allows modifications of nascent transcripts with 5′ m^7G caps, proper transcript processing and elongation checkpoint control. In an active transcription process, the cyclin dependent kinase 9 (Cdk9) in P-TEFb may phosphorylate the carboxyl domain of the RNAPII, DSIF and NFLF-E at sites next to its RNA recognition motif, and release the paused transcripts for processive elongation [22] (Figure 2).

NELF consists of four subunits, A, B, C or D, and E. The C and D subunits of NELF are alternatively spliced products of the same gene [23]. Remarkably, the RD protein is identified as the E-subunit or the smallest subunit of the NELF complex [21]. NELF-E is important for NELF function because it has a functional RNA recognition and binding (RRM) domain. Mutations of the RRM domain in NELF-E impair transcription repression without affecting protein-protein interactions with other NELF subunits [23].

The biochemical activities for various NELF subunits were determined mainly by Hiroshi Handa's team in Yokohama, Japan. They reconstituted NELF-like complex with epitope-tagged NELF-A, NELF-B, NELF-D, and NELF-E expressed in insect Sf9 cells and investigated their interactions [24]. It was shown that NELF binds to DSIF-RNA Polymerase II (RNAPII) complex [21,23]. The leucine zipper of NELF-E interacts directly with NELF-B but not with NELF-A or NELF-D. NELF-D interacts directly with NELF-A and NELF-B. However, NELF-B does not interact directly with NELF-A. Consistent with the model that NELF-B and NELF-D (or NELF-C) that bring NELF-A and NELF-E together via 3 protein-protein interactions, and NELF-A binds to RNAPII directly. The NELF complex needs interactions of 4 subunits, A–(D or C)–B–E. NELF-C and NELF-D are present in distinct NELF complexes [24].

In the nucleus, NELF interacts with the nuclear cap-binding complex (CBC) and participates in the 3′ end processing of mRNA. Both cap-binding proteins CBP-80 and CBP-20 of CBC bind to NELF-E. The region from amino acids 244 to 380 of NELF-E containing an RNA-recognition motif is necessary and sufficient for the interaction between NELF and CBC-RNA. Knockdowns of NELF-E and CBP80 in

HeLa cells rendered the free NELF subunits and CBP20 unstable, which are subjected to degradation rapidly, and abolished functions of the NELF holocomplex and the CBC [25].

On the transcription and processing of small nuclear RNA (snRNAs) that are essential for the assembly and functions of spliceosomes, and small nucleolar RNAs (snoRNAs) that are integral to the biosynthesis of ribosomes by RNAPII, NELF-E and ARS2 (arsenite resistance protein 2) play their respective roles progressively through binding to the cap-binding complex CBC. The binding of NELF-E allows promoter proximal pausing, and the subsequent replacement of NELF-E by ARS2 to the same binding site on CBC plus concurrent binding by PHAX (phosphorylated adaptor for RNA export) enable a continuous process to produce these distinct groups of RNA and prepare them for further processing in the cytoplasm or the nucleus [26].

The mRNAs for histones are special as they may not contain polyadenylated 3' ends. The RRM of NELF-E may bind to the stem-loop structure in the 3' end of the "replication-dependent" histone mRNAs and enhance their stability. Notably, NELF appears to physically associate with histone gene loci and forms the so-called "NELF-bodies", which are probably engaged in the processing of histone mRNA. Knockdowns of NELF-E and CBP-80 in HeLa cells led to increased accumulation of the aberrant, polyadenylated form of histone mRNAs [25].

NELF-E was studied on the regulation of gene expression for human immunodeficiency virus HIV. Recombinant protein with the RNA-recognition motif (RRM) of NELF-E was shown to bind the stem of viral RNA with TAR (transactivation response element) [27]. Such binding (between NELF-E and TAR) inhibits the human immunodeficiency virus type 1 basal transcription from the long terminal repeat (HIV-LTR) [28]. Cyclin-dependent kinase 9 (Cdk9) in P-TEFb can phosphorylate NELF-E at sites next to its RNA recognition motif, and thereby making NELF-E no longer able to bind to TAR nor to repress HIV transcription [22].

Poly-(ADP-ribose) polymerase 1 (PARP1) can covalently transfer of ADP-ribose from NAD^+ onto substrate proteins, consequently modulates diverse biological processes of ribosylated proteins. In a human cell line, Gibson et al. reported that PARP1 interacted with NELF and mediated ADP ribosylation of NELF-E [29,30]. In a manner similar to Cdk9 of P-TEFb-mediated phosphorylation, ADP-ribosylation near the RRM of NELF-E ablates its ability to bind RNA, and therefore releases RNAPII from NELF-dependent pausing. ADP-ribosylation of NELF-E is dependent on phosphorylation by Cdk9/P-TEFb. PARP1-dependent ADP-ribosylation of NELF-E reinforces P-TEFb-mediated RNAPII pause release. Depletion or inhibition of PARP1, or mutagenesis of the four negatively charged glutamate residues in NELF-E (E122, E151, E172, and E374) to glutamines (Q), resulted in a substantial reduction in NELF-E modification by PARP-1. Thus, it was concluded that PARP1-dependent ADP ribosylation of phosphorylated NELF-E (and NELF-A) is necessary for efficient release of RNA polymerase II into productive elongation [29,30].

In a study of double-stranded (ds) DNA break and DNA repair, NELF-E is found preferentially recruited to nearby transcriptionally active genes at sites of DNA ds-breaks to underpin transcriptional repression in a PARP-1 dependent manner. The transcriptional pause allows the machinery for DNA repair to carry out its function and avoid a potential collision of the two systems (DNA repair and transcription) at the damage site. A slight decrease in ribosylation of NELF-E (by PARP-1) at the site of DNA breaks also ensures the pause of transcription imposed by NELF [31].

In a different study of UV-light induced DNA damage (but not ds-breaks) and changes of RNA metabolism, it was found that phosphorylation of NELF-E at serine 115 by MK2 (MAPK activated protein kinase II) promotes the recruitment of 14-3-3 dimers and rapid dissociation of the NELF complex from chromatin. This process is accompanied by elongation of transcription through RNA polymerase II [32].
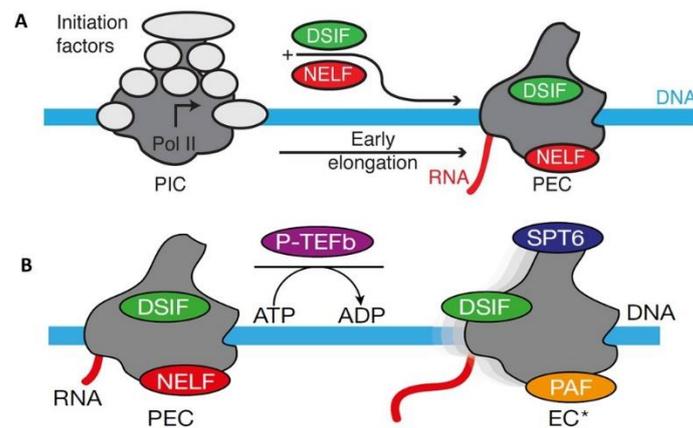
Figure 2. Pause and go of RNA polymerase II during transcription. (**A**) Schematic of conversion of the RNAPII pre-initiation complex (PIC) to a promoter proximally paused RNAPII-DSIF-NELF elongation complex (PEC) [33]. (**B**) Schematic showing conversion of the paused PEC to the activated RNA Pol II-DSIF-PAF-SPT6 elongation complex (EC*). SPT6, homolog of yeast *suppressor* of *Ty6* (adapted from references [33,34], with permission from Dr. Patrick Cramer, Max Planck Institute for Biophysical Chemistry, Germany).

A cryo-electron microscopy structure with paused transcription complex consisting of RNAPII purified from porcine, human DSIF expressed in bacteria, and human NELF complex expressed in insect cells has been presented with 3.2 Å resolution (Figure 2A). This paused RNAPII–DSIF–NELF elongation complex structure contains a tilted DNA–RNA hybrid that is incompatible with binding of nucleoside triphosphate substrates. The NELF complex binds to the funnel structure of the RNAPII and restricts its mobility. During this paused state, through its two tentacle-like structures, NELF has flexible contacts with the DSIF and the exiting RNA. NELF also prevents the binding of the anti-pausing transcription elongation factor IIS [33].

The same research team continued to show cryo-EM structures of the elongation complex with an active transcription process. To re-activate the paused complex, P-TEFb phosphorylates RNAPII, DSIF and NELF. NELF is displaced by Pol II associated factor PAF. Homolog of suppressor for transposon Ty insertion 6 (SPT6) binds to RNAPII and then set off the transcriptional elongation process (Figure 2B) [34].

Considering the fundamental relevance of NELF-E on the cellular process such as gene expression, dysregulated production or aberrant function of NELF-E could have detrimental effects on life processes. Reports on disease associations of NELF-E are relatively scarce but there are recent articles related to NELF-E in carcinogenesis. NELF-E was preferentially overexpressed in hepatocellular carcinoma (HCC), especially in HCC with portal vein invasion (PVI) [35]. NELF-E via somatic copy number alterations appeared to enhance oncogene MYC signaling and activation, which promoted HCC progression [36]. NELF-E protein levels could be an independent risk factor for early intra-hepatic recurrence of HCC within two years of surgery [35].

A Gene Within a Gene: miR-1236 in Intron 3 of NELF-E

While overexpression of NELF-E may be associated with carcinogenesis, a genetic element embedded in the NELF-E gene serves as a tumor suppressor. Between nucleotides 2104 and 2205 of the human NELF-E gene is the precursor for a micro-RNA, miR-1236. This is a splice-dependent mirtronic miRNA that co-expresses with NELF-E [37,38]. miR-1236 is transcribed by RNA polymerase II. The promoter region of miR-1236 is the same as the NELF-E promoter.

miR-1236 can inhibit tumors migration, proliferation and invasion activity by binding to target mRNAs and down-regulating their protein synthesis. It can also induce target gene expression by binding to gene promoter in the nucleus. Binding of miR-1236 to target mRNAs can lead to

abolition of hypoxia-induced epithelial–mesenchymal transition (EMT) and inhibition of migration and invasion of tumor cells. miR-1236-3p can suppress ovarian cancer metastasis by binding to the 3′ untranslated region of zinc-finger E-box binding homeobox 1 (ZEB1) mRNA, and downregulate ZEB1 gene expression [39]. ZEB1 is an EMT inducers. miR-1236 can inhibit the invasion and metastasis of gastric cancer possibly through decreasing phosphorylation by Ak transforming (AKT) kinase and binding to metastasis-associated protein MTA2 [40]. miR-1236 may bind to the promoter of p21 and induces nuclear p21 expression. p21 is a cyclin dependent kinase inhibitor and a regulator of cell cycle progression. miR-1236 and p21 expression are significantly decreased in renal cell carcinoma. Targeted activation of p21 by miR-1236 can inhibit cell proliferation, may have therapeutic potential in the treatment of renal cell carcinoma [41].

3. SKIV2L or SKI2W

Quality Controls I—Clearing Broken, Bad, Used and Viral RNA Products in the Cell Requires An Engine to Unwind Helical structures of RNA for Degradation—The Helicase for Cytoplasmic Exosomes SKIV2L (Ski2w or DDX13)

The Ski complex is a heterotetrameric complex consisting of Ski2, Ski3 and Ski8 in 1:1:2 stoichiometry that is associated with the cytoplasmic RNA exosome for 3′→5′ mRNA decay [42]. Early investigations revealed that yeast Ski2p is engaged in the inhibition of translation from polyadenylation-negative RNA, and has antiviral activities as yeast viruses are RNA structures without poly-A tails [43]. Mutations in these Ski genes in yeasts gave rise to the super-killing phenotype with viral infections [44,45].

Exosomes are highly conserved complex structures responsible for mRNA turnover and therefore fundamental for regulation of gene expression and cell differentiation in eukaryotes. Each exosome consists of 9 or 10 structural proteins that forms a channel structure by which mRNA molecules thread through and get degraded 3′→5′ by an exoribonuclease [46]. The Ski complex serves to unwind the secondary or helical structures of single stranded mRNA or double stranded RNA from some viruses. The Ski3 and Ski8 proteins form the scaffold structure for the Ski complex by which the Ski2 RNA helicase performs its unwinding activity to prepare RNA to thread through the exosome for degradation in the cytoplasm [47,48]. In the nucleus, a similar exosome structure exists but the corresponding RNA helicase is the Mtr4 or KIAA0052 [49,50].

The functional human SKI2W gene in the HLA was discovered in 1995 [13]. It also documented the existence of a SKI2W-related gene, KIAA0052, in the human genome. Because of the striking sequence similarities with the yeast antiviral gene SKI2, SKI2W was renamed to SKIV2L, and KIAA0052 (MTR4) was renamed to SKIV2L2. The human SKIV2L gene is located between RD (NELF-E) and DOM3Z (DXO) at the MHC class III region. It spans 11 kb and consists of 28 exons that codes for a 4 kb transcript. SKIV2L transcripts are ubiquitously expressed in various tissues, especially in the spleen, lymph nodes, duodenum, appendix and gall bladder. The intergenic region between NELF-E and SKIV2L is close to minimum and it was shown that a 176 bp DNA fragment from this region had bidirectional promoter activity for both NELF-E and SKIV2L [51].

The Skiv2l protein has a molecular weight ~140 kDa that consists of 1,246 amino acids [13,17,52]. The Skiv2l protein has structural motifs characteristic of an RNA helicase, which include motifs for binding and hydrolysis of nucleotide triphosphates such as ATP, RNA binding and a DEVH-box [13,53]. Thus, another name for Skiv2l is DDX13 [54]. Skiv2l uses ATP as an energy source to unwind helical or secondary structures of RNA molecules. In addition, Skiv2l contains a leucine zipper motif for protein-protein interaction [13,54], an RGD motif that may be a ligand for adhesion molecules, and an acidic motif with KDLL sequence for endoplasmic reticulum targeting (possibly for unfold protein response). The Skiv2l protein also consists of four potential N-linked glycosylation sites and two potential sulfation sites [13,17].

Fusion proteins with human Ski2w immobilized on polyvinylidene fluoride (PVDF) membrane were able to hydrolyze γ-^{32}P-labeled ATP, which would be an energy source for helicase activities [13].

Immunocytochemistry experiments using specific rabbit antiserum showed that Skiv2l proteins in HeLa cells are localized to the nucleolus and the cytoplasm, a phenomenon that was independently confirmed in transfectants expressing tagged-Skiv2l protein (Figure 3) [52]. Co-sedimentation experiments for ribosome profile through sucrose gradient centrifugation and fractionations revealed that Ski2w is associated with the 40S subunit of ribosomes [52].
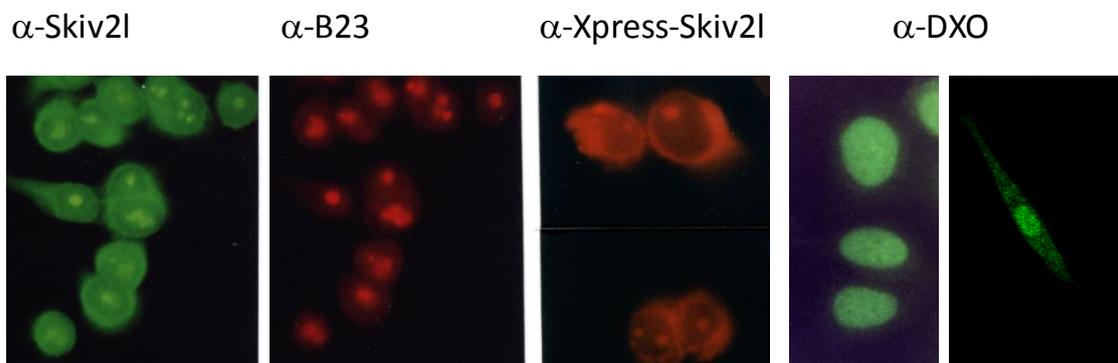
α-Skiv2l   α-B23   α-Xpress-Skiv2l   α-DXO

Figure 3. Cellular localization of Skiv2l and DXO in HeLa cells under immunofluorescent microscopy. Human HeLa cells were stained by antibodies against human Skiv2l (Ski2w), nucleoli-specific B23 (nucleophosmin), anti-Xpress for HeLa transfectants with Xpress-Skiv2l fusion protein, and antibodies against DXO (Dom3z). The first three panels illustrate the presence of Skiv2l in the nucleoli and cytoplasm. The fourth and fifth panels show the dominant location of DXO (Dom3z) in the nuclei of HeLa cells. Photographs were originally taken from UV-microscope with 10 × 40 magnifications (modified from [15,52,55]).

The roles of human Skiv2l on the fine regulation of antiviral defense and RNA turnover was examined recently [55]. Discrimination of foreign RNA such as viral RNA with 5′ triphosphate instead of a m^7G cap for host mRNA, or double-stranded RNA that seldom exists in a eukaryotic cell, are achieved by RNA helicases such as retinoic acid activating gene 1 (RIG1) or melanoma differentiation-associated protein 5 (MDA5) [56,57]. Stimulation with RIG1/MDA5 ligands [such as thapsigargin or poly(I:C)] on primary macrophages with *knockdown* of Skiv2l, but not Xrn1 or Ski3, led to substantial increase of transcription for interferon-β. Thus, Skiv2l specifically suppressed or regulated the production of type I interferon from macrophages in response to ligands that mimic viral infections. This is because Xrn1 knockdown or Ski3 (TTC37) knockdown in similar experiments had no effects on IFN-β expression [56]. Skiv2l probably fine-tunes the antiviral response by RIG1/MDA5 to prevent the over inflammation that can cause tissue injuries.

During a viral infection or an inflammation, the host cells are under stress and endoplasmic reticula in cells are subjected to excessive protein synthesis that overwhelms the posttranslational machinery to properly fold and modify protein molecules. Such stress triggers the unfolded protein response (UPR) through inositol-requiring enzyme 1 (IRE-1) for endonuclease cleavage of specific cellular mRNAs. One of them is a premature mRNA, which retains an unspliced intron, for the *X*-box *b*inding *p*rotein 1 (XBP-1). Splicing out the intronic sequences from the premature mRNA of XBP-1 in the cytoplasm generates ligands for endogenous retinoic acid-inducible gene-like receptors (RLRs). In the absence of Skiv2l, such UPR would trigger elevated production of type I interferon (IFN) production. Therefore, SKIV2L is an important negative regulator of the RLR response to exogenous RNA ligands including RNA viruses [56]. Under such a scenario, the absence or partial functional deficiency of Skiv2l could generate unchecked or dysregulated production of ligands or sensors for RLR. The results would be autoinflammation in the absence of a viral infection, or increased susceptibility to systemic autoimmune disease. Table 2 lists results of recent attempts to investigate the potential roles of Skiv2l mutations in human diseases.

It was mind-boggling to find that a deficiency in Skiv2l or another Ski complex protein Ski3 or TTC37 causes the syndromic intractable diarrhea known as tricho-hepato-enteric syndrome or THES [58,59]. This extremely rare syndrome is characterized by (1) the presence of woolly and easily breakable scalp hair, (2) facial dysmorphism and hypertelorism of paired organs such as eyes, (3) intrauterine restriction in growth with low birth weights, (4) dysfunctional liver with hematochromatosis and coarse appearance, (5) intractable diarrhea that may initiate soon after birth or at neonatal stage, persists despite prolonged bowl rest. Many baby patients with THES require parenteral nutrition, become dehydrated and marasmic, and incapable to thrive [60–63].

Table 2. A list of human diseases associated with mutations of SKIV2L.

Disease	SNP	Relationship to Disease	Remarks
SLE (Systemic Lupus Erythematosus)	T allele of rs419788 in intron 6	Confer disease susceptibility in additive pattern	One copy confers a low risk of disease and two copies result in greater susceptibility [64]
AMD (Age-related macular degeneration)	R151Q (rs438999)	May exert a functional effect in AMD	Strong LD (linkage disequilibrium) with Bf R32Q (rs641153) [65]
	Intronic SNP (rs429608)	AMD genetic protective factors	Protective effect for AMD [66] in Han Chinese and Japanese populations [67,68]
	rs429608 and rs453821	Significantly associated with neovascular AMD	Not associated significantly with PCV [69]
PCV (Polypoidal Choroidal Vasculopathy)	3′UTR (rs2075702)	Significant association with PCV	Decreased risk of developing PCV [70]
SD/THES (Syndromic diarrhea/tricho-hepato-enteric syndrome)	c.1635_1636insA (p.Gly546Argfs*35) c.2266C>T (p.Arg756*) c.2442G>A (p.Trp814*) c.848G>A (p.Trp283*) c.1022T>G (p.Val341Gly) c.2572del (p.Val858*) c.2662_2663del (p.Arg888Glyfs*12) c.1434del (p.Ser479Alafs*3)	Deleterious mutations detected in six individuals with typical SD/THES	Molecular defects in SKIV2L cause SD/THES [58,71]
	c.1891G>A p.Gly631Ser c.3187C>T p.Arg1063*	Two new mutations found in a trichohepatoenteric syndrome patient	The patient was a Malaysian child [72]
	(c.1891G>A) (c.1120C>T)	Two variants identified in THES patients	These two mutations can cause THES [73]
	c.1420G>T (p.Q474*) c.3262G>T (p.E1088*)	Novel compound heterozygous nonsense mutations were identified in THES patient	Decreased levels of SKIV2L protein expression in blood mononuclear cells [74]
	25 exons (p.Glu1038 fs*7 (c.3112_3140del))	A rare mutation in THES patient	This patient died at three year old [75]
IBD (Inflammatory Bowel Disease)	c.354+5G>A	Identified a novel splicing mutation in a patient with IBD (ulcerative colitis)	This mutation was related to Inflammatory Bowel Disease [76]

Further analyses revealed that most THES patients had defects in the immune system with low levels of switched immunoglobulins, highly variable responses to vaccination including unresponsiveness of specific antigens, impaired T cell and NK cell functions such as low production of γ-interferon and abnormal T cell proliferation, and severe Epstein Barr Virus (EBV) infection [77]. It appears that 60% of the THES patients had a genetic deficiency of TTC37 [59] and 40% had a deficiency of SKIV2L [78], and many of those homozygous deficiencies were the results of consanguineous marriage. Since the discovery of THES, a spectrum of disease phenotypes and severity has emerged,

and the disease has been reported in multiple racial groups [74,75,79–81]. It has been noticed that THES patients with deficiencies of SKIV2L experienced more severe disease in terms of liver damage and impairment of prenatal growth than those with deficiencies of TTC37 [78]. THES patients with deficiency of Skiv2l but not deficiency of TTC37 exhibited a remarkable type I interferon stimulated gene expression profile [56].

It is not clear for the mechanism between the loss of function for SKIV2L or TTC37 and the detrimental phenotypes of THES, which include the failure to thrive, inability in food digestion, poor quality in hair texture, and impairments of liver and immune functions.

Intriguingly, elevated expression levels of type I interferon stimulated genes in the absence of apparent viral infections is a marked phenomenon of systemic autoimmune diseases including systemic lupus erythematosus and juvenile dermatomyositis [82–84]. Thus, it would be appropriate to examine if there were genetic or acquired deficiency, dysregulation or impaired functions of Skiv2l in autoimmune, inflammatory and infectious diseases [56].

4. DXO or DOM3Z

Quality Controls II—Destroy Disqualified Products—DXO (DOM3Z) to Degrade Miscapped or Uncapped RNA from the Start to the End

The discovery of human DOM3Z (DXO) was reported by Yang and colleagues in 1998 [14]. This gene is arranged head-to-head orientation with STK19 (RP1) and tail-to-tail with SKIV2L (SKI2W) [15,17]. The 5′ ends of DOM3Z and STK19 overlap but they are in opposite configurations. The exon 1 of STK19 is present at the intron 2 of DOM3Z genomic region. The 3′ ends of DOM3Z and SKI2W are separated by 59 base pairs.

The DOM3Z gene has CpG-rich sequences at its 5′ region, a feature conserved in housekeeping genes in simple and complex eukaryotes. Human DOM3Z is ubiquitously expressed, but the highest expression levels are found in the adrenals and the testes. The first exon (123 bp) of DOM3Z is noncoding because there is an in-frame TGA stop codon at nucleotides 19–21. The full length DOM3Z cDNA is 1386 bp and its genomic DNA is 2.2 kb in size, which consists of 7 exons encoding 396 amino acids. Between residues 103 to 124 of the Dom3z protein is a leucine zipper motif for protein-protein interactions [14]. Yang and colleagues also found that human DOM3Z has homologs in round worm *Caenorhabditis elegans* (DOM-3), baker's yeast, fission yeast and possibly flowering plant Arabidopsis [14,15,17]. However, no other known protein motifs were identified at that time.

Most progress in studying the biologic functions of DOM3Z homologs occurred in yeasts, which is named **Rai1**. Through genetic deletion and complementation analyses, Xue and coworkers showed in year 2,000 that Rai1 interacts Rat1 (also named Xrn2). The yeast Rat1 is a nuclear 5′→3′ exoribonuclease that is important in RNA processing and RNA degradation [85]. Rai1 stabilizes and activates the exonuclease activities of Rat1. Genetic mutants of Rai1 affects the proper (5′ and 3′) processing of the 5.8S ribosomal RNA. Deletion of Rai1 gene leads to defective biogenesis of the 60S ribosome but that can be suppressed by high copy of RAT1. Intriguingly, double mutants of Rai1 and putative RNA helicase gene Ski2 in yeast were synthetically lethal [85], suggesting Rai1 and Ski2 proteins have overlapping functions that cannot be both dysfunctional.

Xiang, Jiao, Tong, Kiledjian and colleagues in 2009 reported crystal structures of *S. pombe* Rat1 in complex with Rai1 in 2.2 Å resolution [86], and Rai1 and its murine homolog Dom3z in 2.0 Å resolution. It was shown that Rai1 can stabilize Rat1 secondary structure to facilitate Rat1 to degrade RNAs more effectively. They identified a large pocket in Rai1 and in Dom3z that contains highly conserved residues, including three acidic side chains that coordinates a divalent cation such as Mg^{+2} or Mn^{+2}. They demonstrated that Rai1 has pyrophosphohydrolase activity towards 5′ triphosphorylated RNA [86]. This team of investigators continued to demonstrate that Rai1 possesses decapping and endonuclease activities, which can remove the entire cap structure dinucleotide from an mRNA in vitro and in yeast cells, especially towards mRNAs with unmethylated, faulty or aberrant caps [87]. Thus, it is proposed that Rai1 is involved in quality control process to ensure mRNA 5′-end integrity.

The same research team also found a related protein for Rai1 in yeast known as Ydr370C, which is mainly expressed in the cytoplasm. The crystal structure of the Ydr370C protein from *Kluyveromyces lactis* (named as DXO1) was solved and its biochemical activity investigated [88]. Ydr370C has robust decapping activity for incompletely capped mRNA and 5′→3′ distributive exoribonuclease activity, but no activity for pyrophosphohydrolase. The cytoplasmic Ydr370C was renamed Dxo1 because of its decapping and exoribonuclease activities for mRNA. If both Dxo1 and Rai1 in yeast are disrupted, mRNAs with incomplete caps are produced and readily detectable under normal growth conditions. Therefore, yeast possesses two partially redundant proteins that detect and degrade incompletely capped mRNAs [88]. In mammals including humans, there is only one homolog for Rai1 and Dxo1, Dom3z. Dom3z possesses all three functional activities of Rai1 and Dxo1, which are pyrophosphohydrolase, decapping, and 5′ to 3′ exoribonuclease activities. Thus, Dom3z is given an alternative name as Dxo. Dom3z preferentially degrades defectively capped and inefficiently spliced pre-mRNAs (Figure 4A) [89,90]. Collectively, Rai1, Dxo1 and Dxo identify a novel class of proteins that functions in a quality control mechanism to ensure proper mRNA 5′-end fidelity or integrity [89].
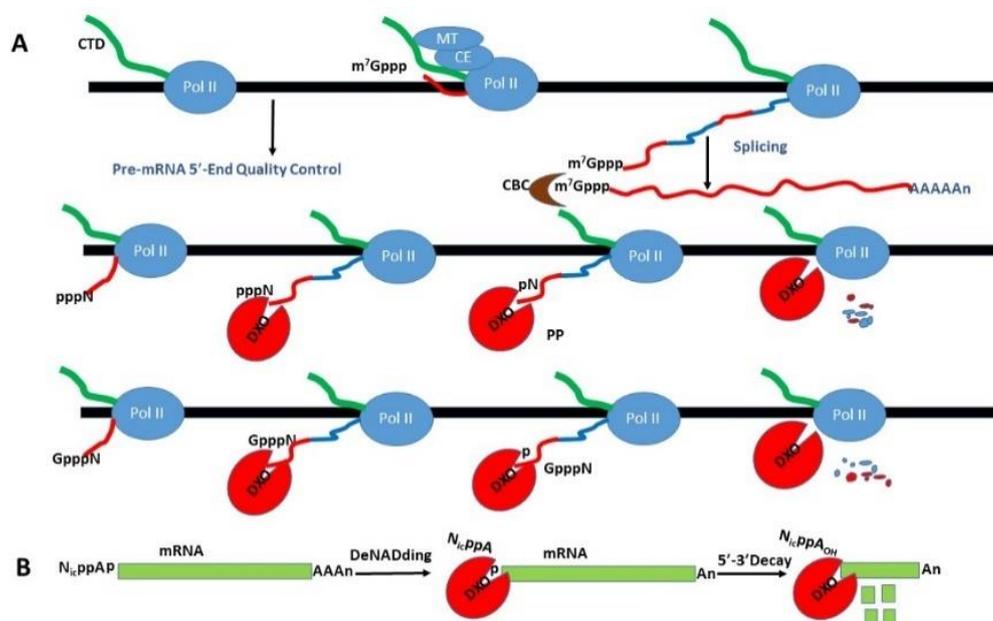

Figure 4. DXO decapping and DXO-mediated decay of NAD$^+$-capped mRNA. (**A**) DXO can remove the diphosphate from a triphosphorylated cap of mRNA, the entire cap structure of an unmethylated guanosine-capped mRNA, or aberrant pre-mRNAs. (**B**) DXO hydrolyzes the phosphodiester linkage 3′ to the adenosine to remove the entire NAD$^+$ moiety and generates an RNA molecule with 5′ monophosphate. Modified and redrawn from ref [89] for panel A, and ref [91] for panel B (with permission from Dr. Mike Kiledjian, Rutgers University, NJ).

If the first ($^m7$GpppN$_m$pRNA) or the second nucleotide ($^m7$GpppN$_m$pN$_m$pRNA) of the mRNA were methylated at the ribose 2′-O position, it would be protected from Dxo-mediated decapping and degradation [92].

Over the past forty years, it has been firmly established that mRNA in eukaryotes are predominantly modified by a 7-methyl-guanosine cap at the 5′ end during the initiation of gene transcription. This m^{7}G cap helps the processes for (a) RNA splicing to remove intronic sequences, (b) the nucleus-cytoplasm transport of mature mRNA for translation, and (c) protection of mature mRNA from degradation. It has been presumed that such capping process in eukaryotes was constitutive and efficient with little errors until yeast Rai1 is shown to play an important role on the quality control of mRNA capping and stability. Moreover, it has been found recently that some bacteria RNA can be capped at the 5′ end with nicotinamide adenine dinucleotide (NAD$^+$) to increase their stability. Unexpectedly, such NAD$^+$

caps are also present in 5–10% eukaryotic mRNA and some small nucleolar RNA (snoRNA). Notably, NAD$^+$-capped mRNAs are *not* translatable. They are also less stable and subject to Dxo-mediated decapping and 5′ →3′ decay, a process known as deNADding (Figure 4B) [91,93].

5. STK19 (RP1/G11)

An On-and-Off Switch—The Nuclear Kinase STK19 (RP1 or G11) Whose Malfunction Can Cause Melanoma

While the 5′ regulatory regions of the duplicated C4 genes were characterized, a canonical polyadenylation signal (AATAAA) was found in the promoter region of each C4 gene, suggesting the presence of a novel gene. Thus, cDNA clones were isolated and genomic DNA sequences for this novel gene were determined [10,11,94]. This gene was originally named RP to memorialize the late Professor Rodney Porter, in his research unit the human MHC complement genes were cloned and characterized.

In human, an intact RP1 gene is located 611 base pairs upstream of the first C4 gene, and a partially duplicated RP2 gene is present upstream of each of the succeeding C4 gene (Figure 1) [10,17]. Immediately upstream of RP1 is DOM3Z. RP1 is in head-to-head configuration with DOM3Z, and tail-to-head with C4 [17]. Upstream of RP2 is another partially duplicated gene fragment for tenascin X, TNXA. Human RP is part of a four-gene duplication complex in the HLA termed the RCCX module, which stands for RP, C4, CYP21 and TNX (Figure 1) [95,96].

RP1 is expressed in all tissues and its highest expression is in the adrenals and the testes of human and rodents. Homologs of RP1 are predominantly present among vertebrates but an exception is found in a sea anemone (*Exaiptasia pallida, LOC110250916/XP_020913230.1*). The RP1 protein contains a bipartite nuclear localization signal and was shown to be a nuclear protein [10]. Subsequently, recombinant protein of this gene, which was also named G11 by Duncan Campbell and colleagues [97], was shown to be a serine/threonine kinase [98]. This gene was renamed to STK19, which uncoupled publications for some of the original observations [10,95,99]. Other names shown at the NCBI website for this gene include D6S60; D6S60E; HLA-RP1.

STK19 gene contains 7 or 9 exons. The 9-exon STK19 gene is 11 kb in size and codes for 364 amino acids with a calculated molecular weight of 41.5 kDa. This long form of STK19 gene overlaps to the 5′ region of the DOM3Z gene. There is a variant coding for 368 amino acids that is due to an alternative splicing, which results in an extension of 12 nucleotides from the 3′ end of exon 5 to the intron 6, leading to the addition of four more amino acids after R221 (VCDC). The 7-exon version of STK19 is 9.1 kb in size that has a 267 bp intergenic distance from DOM3Z. It codes for 254 amino acids with a calculated molecular weight of 30 kDa. The first 110 amino acid residues encoded by the first two exons are absent in the shorter RP1 transcript. Murine STK19 has the small version with 254 amino acids.

Scattered in the introns of the human RP1 gene are 8 copies of short interspersed nuclear elements named *Alu*. One of them is an integral component of a mobile retroelement SVA that consists of a distinct SINE, 21 copies of GC-rich variable number tandem repeats (VNTRs), and an Alu element [10,15,17]. The SVA element in RP1 was flanked by short direct repeating sequences of 13 nucleotides, which was indicative of insertion by a discrete composite genetic unit (Figure 1B). The SVA in RP1 was the first identification and documentation of such distinct and discrete composite retroelements into a human genome [10]. The highly repetitive nature of the SVA sequences render specific biologic studies challenging. It was demonstrated certain member(s) of SVA has the capability to "retrotranspose" with help of the reverse transcriptase coded by active long interspersed element LINE-1 [100–102].

In terms of gene function, STK19 encodes a nuclear protein [10] that is a kind of serine/threonine kinase. The STK19 protein can bind ATP at its kinase domain, phosphorylate certain proteins at Ser/Thr residues, such as α-casein and histones [98]. The actual physiological properties and function, the substrates of the kinase activity for STK19, and the relationship with disease for STK19 are being gradually revealed [103]. Results of the interactome analyses revealed that nine proteins bind to STK19 (https://www.ncbi.nlm.nih.gov/gene/8859, [104]). These proteins are DNA polymerase subunit E or POLE, RNA polymerase II subunit G or POLR2G, transcription factor SP3, mRNA splicing

factor SF3BF4, basic helix-loop-helix protein e40 or BHLHE40 that controls circadian rhythm and cell differentiation, inosine monophosphate dehydrogenase 1 or IMPDH1 that is an enzyme essential for guanine nucleotide biosynthesis through catalysis of synthesis for xanthine monophosphate from inosine 5-monophosphate, membrane transporter for nucleosides or solute carrier 29 or SLC29A1, preferentially expressed antigen in melanoma PRAME that acts as a repressor of retinoic acid receptor and may promote growth of cancer cells, and tripartite motif protein TRIM23 that may be a ADP ribosylation factor for guanine nucleotide binding protein and plays a role in the formation of intracellular transport vesicles and for their movement from one compartment to another in the cell. These interacting proteins may be substrates or cofactors of STK19 kinase, implicating that STK19 could be engaged in DNA replication, RNA transcription, RNA splicing, nucleoside/nucleotide synthesis and transport, and control of gene expression for development and differential cellular functions.

While STK19 knockdown in cells revealed little effects on transcription, it is required for recovery of transcription after DNA damage and therefore important for the transcription-related DNA damage response [105]. STK19 interacts with Cockayne syndrome B protein (CSB, also named ERCC6) after DNA damage. Cells lacking STK19 were UV-sensitive.

By analysis of large-scale exome data from melanoma and uninvolved tissues, Eran Hodis and colleagues showed STK19 is a novel melanoma gene for its somatic hotspot of point mutations. While the D89N mutation of SKT19 has a 5% frequency in a melanoma [106], no SKT19 mutations were detectable in nevus associated-melanomas [107]. There are 10% gene mutations in STK19 in basal cell carcinoma (BCC). Strikingly, all of those mutations resulted in D89N substitution [108]. Recently, Yin and colleagues reported STK19 can phosphorylate NRAS at Ser-89 and activate oncogenic NRAS to promote melanoma genesis. When STK19 has the D89N change, it is a gain of function mutation that phosphorylates or interacts more efficiently with NRAS and enhances melanocyte transformation [109].

In addition, by next-generation sequencing and bioinformatics analysis, STK19 was found to have novel T/C missense single-nucleotide variants in Chr6:31947203 in induced pluripotent stem cell (iPSC) lines, which can be used to trace the genotype of original cells [110].

Through meta-analyses of published GWAS data, STK19, SKIV2L and NELF-E appeared to be potential pleiotropic genes for metabolic syndrome and inflammation [111]. Moreover, Amare et al. reported that genetic variant of STK19 gene correlated with schizophrenia by bivariate GWAS [112]. Other functions of STK19 remain to be demonstrated.

6. The Physical Linkage of NELF-E and SKIV2L (RD and SKI2W), DXO and STK19 (DOM3Z and RP1) and the Emergence of NSDK Quartet in the MHC

We ask when did the genes for NELF-E, SKIV2L, DXO and STK19 start to link together physically as a group in the animal kingdom, and their relationships with the class I and class II genes in the emergence of the MHC in the vertebrates. We attempted to obtain answers to these questions through analyzing data deposited in the National Center for Biotechnology Information (NCBI/PubMed: https://www.ncbi.nlm.nih.gov/home/ or https://www.ncbi.nlm.nih.gov/pubmed/).

The evolution history of NELF-E is intriguing because the presence of a leucine zipper and the RRM, which is a characteristic of a NELF-E like molecule, can be found in *Drosophila* but there are no RD repeats (https://www.ncbi.nlm.nih.gov/protein/NP_648241.1, [113]). The presence of the RD (Arg-Asp) and/or RE (Arg-Glu) repeats is definitive in zebrafish NELF-E (https://www.ncbi.nlm.nih.gov/protein/XP_005170084.1, [114]). Related gene or protein function for SKIV2L can be found in yeast, and for DXO in *C. elegans* [115]. STK19 homologs with high sequence identities are predominantly present among vertebrates. There is no evidence for a physical linkage for any two of the NSDK quartet prior to vertebrates.

For the data deposited in the NCBI, most chromosomal segments are relatively short, and their linkage status is not yet known. The neighboring gene linkages of NELF-E and SKIV2L in one chromosomal segment, and DXO and STK19 in another chromosomal segment are observable in shark, zebrafish, and coelacanth. Genes in each of these pairs are organized in "head-to-head"

configuration and therefore they share the same 5′ regulatory sequences in opposite orientations (Figure 1B). The physical linkage of all four genes (i.e., NELF-E, SKIV2L, DXO and STK19) together as a group seems to be definitive in frogs [116]. Again, we should emphasize that data available for analyses are limited.

In egg-laying mammal platypus, MHC class I and class II genes are present on chromosome X3, but most (if not all) of the class III genes form a linkage unit on chromosome X5. The assembly of the class III genes (including NSDK with complement convertases), class I genes and class II genes as the MHC locus appears intact in placental mammals.

While the NSDK quartet genes express ubiquitously, it is not known whether there are concerted controls of gene expression (such as genes in operons in prokaryotes), or differential responses to regulation during growth and development.

7. Conclusions

NELF-E-SKIV2L-DXO-STK19 is a complex genetic unit located between complement factor B and component C4 in the class III region of the major histocompatibility complex (MHC) [98,117]. NELF-E, SKIV2L and DXO make important roles in transcription, processing and degradation of a variety of RNAs and antiviral immunity [24,56]. As such, NELF-E or RD is a subunit of the negative elongation factor (NELF). It is an RNA-binding protein that imposes a promoter proximal pause in transcription by the RNA polymerase II. The purpose is to allow capping and quality checking prior to full strength elongation of RNA for gene expression [24,30]. SKI2W or SKIV2L RNA helicase, an engine to unwind helical structures of RNA molecules and propel them through the tunnel formed by cytoplasmic exosomes for 3′→5′ degradation. SKIV2L is an important negative regulator of the RIG-I-like receptors (RLR) mediated antiviral response and an important modulator of the type I interferon response in the cytoplasm [56]. The RLR [56,57], toll-like receptors, and complement systems are innate immune response effectors that engage in pattern recognitions of differentially modified biomolecules to differentiate self and non-self [118,119].

Genetic studies revealed that homozygous deficiency of human SKIV2L contributes to the pathogenesis of syndromic diarrhea or trichohepatoenteric syndrome (THES). Regulation of gene expression is critical in determining cell identity, development and responses to the environment. DXO or Dom3z possesses pyrophosphohydrolase, decapping and 5′-3′ exoribonuclease activities that can preferentially degrade incompletely 7-methylguanosine (m^7G) capped, or NAD^+ capped mRNAs from the 5′ end [89]. STK19 is a Serine/Threonine nuclear protein kinase. STK19 is important for the transcription-related DNA damage response [105,109]. STK19 phosphorylates N-RAS at Ser-89. Dysregulated phosphorylation of NRAS at S89 activates and promotes melanoma genesis [109]. It is necessary to point out that the biochemical and functional properties of STK19 are still largely unexplored. It is also not clear whether and how do the NSDK quartet proteins interact with each other, and the advantage for these four genes being tightly linked together.

Thus, NSDK seems to engage in the surveillance of host RNA integrity at the "head" and at the "tail" [120], in the destruction and turnover of faulty or expired RNA molecules or RNA viruses, and in the fine-tuning of innate immunity. The roles of NSDK in carcinogenesis, infectious and autoimmune diseases are only starting to emerge. We look forward to further studies on disease associations, functional mechanisms and therapeutics of these very old but non-classical MHC genes.

Abbreviations

ARS2	arsenite resistance protein 2;
CBC	cap-binding complex;
CDK9	cyclin dependent kinase 9;
CSB	Cockayne syndrome B protein (also named ERCC6);
DRB	5,6-*d*ichloro-1-β-ᴅ-*ri*bofuranosyl*b*enzimidazole;
DSIF	DRB sensitivity inducing factor;
DXO	decapping exoribonuclease;
EC *	activated elongation complex;
EMT	epithelial–mesenchymal transition;
HCC	hepatocellular carcinoma;
IRE-1	Inositol-requiring enzyme 1;
MDA5	melanoma differentiation-associated protein 5;
MK2	MAPK-activated protein kinase II;
NSDK	NELF-E (RD), SKIV2L (SKI2W), DXO (DOM3Z) and STK19;
NELF	negative elongation factor;
P-TEFb	positive transcription elongation factor b;
PAF	Pol II associated factor;
PARP-1	poly-ADP-ribose polymerase 1;
PEC	paused elongation complex;
PHAX	phosphorylated adaptor for RNA export;
RD	a gene with Arg-Asp repeats, also known as NELF-E;
RIG-I	retinoic acid-inducible gene 1;
RLRS	RIG-I-like receptors;
RNAPII	RNA polymerase II;
RP	the gene upstream of complement C4, also known as STK19;
RRM	RNA recognition and binding domain;
SKI	yeast genes by which mutations lead to superkilling phenotypes after viral infection;
snRNA	small nuclear RNA;
snoRNA	small nucleolar RNA;
SPT6	yeast homolog of suppressor for transposon Ty insertion, 6;
UPR	unfolded protein response;
TAR	transactivation response element in HIV;
ZEB1	zinc-finger E-box binding homeobox 1

References and Note

1. Carroll, M.C.; Campbell, R.D.; Bentley, D.R.; Porter, R.R. A molecular map of the human major histocompatibility complex class III region linking complement genes C4, C2 and factor B. *Nature* **1984**, *307*, 237–241. [CrossRef] [PubMed]

2. Horton, R.; Wilming, L.; Rand, V.; Lovering, R.C.; Bruford, E.A.; Khodiyar, V.K.; Lush, M.J.; Povey, S.; Talbot, C.C., Jr.; Wright, M.W.; et al. Gene map of the extended human MHC. *Nat. Rev. Genet.* **2004**, *5*, 889–899. [CrossRef] [PubMed]

3. Miller, F.W.; Cooper, R.G.; Vencovsky, J.; Rider, L.G.; Danko, K.; Wedderburn, L.R.; Lundberg, I.E.; Pachman, L.M.; Reed, A.M.; Ytterberg, S.R.; et al. Genome-wide association study of dermatomyositis reveals genetic overlap with other autoimmune disorders. *Arthritis Rheum.* **2013**, *65*, 3239–3247. [CrossRef] [PubMed]

4. Dawkins, R.; Leelayuwat, C.; Gaudieri, S.; Tay, G.; Hui, J.; Cattley, S.; Martinez, P.; Kulski, J. Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease. *Immunol. Rev.* **1999**, *167*, 275–304. [CrossRef] [PubMed]

5. Hauptmann, G.; Bahram, S. Genetics of the central MHC. *Curr. Opin. Immunol.* **2004**, *16*, 668–672. [CrossRef] [PubMed]

6. Lintner, K.E.; Wu, Y.L.; Yang, Y.; Spencer, C.H.; Hauptmann, G.; Hebert, L.A.; Atkinson, J.P.; Yu, C.Y. Early Components of the Complement Classical Activation Pathway in Human Systemic Autoimmune Diseases. *Front. Immunol.* **2016**, *7*, 36. [CrossRef] [PubMed]

7. Savelli, S.L.; Roubey, R.A.S.; Kitzmiller, K.J.; Zhou, D.; Nagaraja, H.N.; Mulvihill, E.; Barbar-Smiley, F.; Ardoin, S.P.; Wu, Y.L.; Yu, C.Y. Opposite Profiles of Complement in Antiphospholipid Syndrome (APS) and Systemic Lupus Erythematosus (SLE) Among Patients With Antiphospholipid Antibodies (aPL). *Front. Immunol.* **2019**, *10*, 885. [CrossRef]

8. Ricklin, D.; Hajishengallis, G.; Yang, K.; Lambris, J.D. Complement: a key system for immune surveillance and homeostasis. *Nat. Immunol.* **2010**, *11*, 785–797. [CrossRef]

9. Wu, L.C.; Morley, B.J.; Campbell, R.D. Cell-specific expression of the human complement protein factor B gene: evidence for the role of two distinct 5'-flanking elements. *Cell* **1987**, *48*, 331–342. [CrossRef]

10. Shen, L.; Wu, L.C.; Sanlioglu, S.; Chen, R.; Mendoza, A.R.; Dangel, A.W.; Carroll, M.C.; Zipf, W.B.; Yu, C.Y. Structure and genetics of the partially duplicated gene RP located immediately upstream of the complement C4A and the C4B genes in the HLA class III region. Molecular cloning, exon-intron structure, composite retroposon, and breakpoint of gene duplication. *J. Biol. Chem.* **1994**, *269*, 8466–8476.

11. Yu, C.Y. The complete exon-intron structure of a human complement component C4A gene. DNA sequences, polymorphism, and linkage to the 21-hydroxylase gene. *J. Immunol.* **1991**, *146*, 1057–1066. [PubMed]

12. Levi-Strauss, M.; Carroll, M.C.; Steinmetz, M.; Meo, T. A previously undetected MHC gene with an unusual periodic structure. *Science* **1988**, *240*, 201–204. [CrossRef] [PubMed]

13. Dangel, A.W.; Shen, L.; Mendoza, A.R.; Wu, L.C.; Yu, C.Y. Human helicase gene SKI2W in the HLA class III region exhibits striking structural similarities to the yeast antiviral gene SKI2 and to the human gene KIAA0052: emergence of a new gene family. *Nucleic Acids Res.* **1995**, *23*, 2120–2126. [CrossRef] [PubMed]

14. Yang, Z.; Shen, L.; Dangel, A.W.; Wu, L.C.; Yu, C.Y. Four ubiquitously expressed genes, RD (D6S45)-SKI2W (SKIV2L)-DOM3Z-RP1 (D6S60E), are present between complement component genes factor B and C4 in the class III region of the HLA. *Genomics* **1998**, *53*, 338–347. [CrossRef] [PubMed]

15. Yang, Z.; Yu, C.Y. Organizations and gene duplications of the human and mouse MHC complement gene clusters. *Exp. Clin Immunogenet.* **2000**, *17*, 1–17. [CrossRef]

16. Yu, C.Y.; Yang, Z.; Blanchong, C.A.; Miller, W. The human and mouse MHC class III region: a parade of 21 genes at the centromeric segment. *Immunol. Today* **2000**, *21*, 320–328.

17. Yang, Z.; Qu, X.; Yu, C.Y. Features of the two gene pairs RD-SKI2W and DOM3Z-RP1 located between complement component genes factor B and C4 at the MHC class III region. *Front. Biosci.* **2001**, *6*, D927–D935. [CrossRef]

18. Cheng, J.; Macon, K.J.; Volanakis, J.E. cDNA cloning and characterization of the protein encoded by RD, a gene located in the class III region of the human major histocompatibility complex. *Biochem. J.* **1993**, *294*, 589–593. [CrossRef]

19. Surowy, C.S.; Hoganson, G.; Gosink, J.; Strunk, K.; Spritz, R.A. The human RD protein is closely related to nuclear RNA-binding proteins and has been highly conserved. *Gene* **1990**, *90*, 299–302. [CrossRef]

20. Speiser, P.W.; White, P.C. Structure of the human RD gene: A highly conserved gene in the class III region of the major histocompatibility complex. *DNA* **1989**, *8*, 745–751. [CrossRef]

21. Yamaguchi, Y.; Takagi, T.; Wada, T.; Yano, K.; Furuya, A.; Sugimoto, S.; Hasegawa, J.; Handa, H. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* **1999**, *97*, 41–51. [CrossRef]

22. Fujinaga, K.; Irwin, D.; Huang, Y.; Taube, R.; Kurosu, T.; Peterlin, B.M. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol. Cell Biol.* **2004**, *24*, 787–795. [CrossRef] [PubMed]

23. Yamaguchi, Y.; Inukai, N.; Narita, T.; Wada, T.; Handa, H. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Mol. Cell Biol.* **2002**, *22*, 2918–2927. [CrossRef] [PubMed]

24. Narita, T.; Yamaguchi, Y.; Yano, K.; Sugimoto, S.; Chanarat, S.; Wada, T.; Kim, D.K.; Hasegawa, J.; Omori, M.; Inukai, N.; et al. Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. *Mol. Cell Biol.* **2003**, *23*, 1863–1873. [CrossRef] [PubMed]

25. Narita, T.; Yung, T.M.; Yamamoto, J.; Tsuboi, Y.; Tanabe, H.; Tanaka, K.; Yamaguchi, Y.; Handa, H. NELF interacts with CBC and participates in 3' end processing of replication-dependent histone mRNAs. *Mol. Cell* **2007**, *26*, 349–365. [CrossRef] [PubMed]

26. Schulze, W.M.; Cusack, S. Structural basis for mutually exclusive co-transcriptional nuclear cap-binding complexes with either NELF-E or ARS2. *Nat. Commun.* **2017**, *8*, 1302. [CrossRef] [PubMed]

27. Rao, J.N.; Neumann, L.; Wenzel, S.; Schweimer, K.; Rosch, P.; Wohrl, B.M. Structural studies on the RNA-recognition motif of NELF E, a cellular negative transcription elongation factor involved in the regulation of HIV transcription. *Biochem. J.* **2006**, *400*, 449–456. [CrossRef] [PubMed]

28. Pagano, J.M.; Kwak, H.; Waters, C.T.; Sprouse, R.O.; White, B.S.; Ozer, A.; Szeto, K.; Shalloway, D.; Craighead, H.G.; Lis, J.T. Defining NELF-E RNA binding in HIV-1 and promoter-proximal pause regions. *PLoS Genet.* **2014**, *10*, e1004090. [CrossRef] [PubMed]

29. Gibson, B.A.; Zhang, Y.; Jiang, H.; Hussey, K.M.; Shrimp, J.H.; Lin, H.; Schwede, F.; Yu, Y.; Kraus, W.L. Chemical genetic discovery of PARP targets reveals a role for PARP-1 in transcription elongation. *Science* **2016**, *353*, 45–50. [CrossRef]

30. Guo, J.; Price, D.H. RNA polymerase II transcription elongation control. *Chem. Rev.* **2013**, *113*, 8583–8603. [CrossRef]

31. Awwad, S.W.; Abu-Zhayia, E.R.; Guttmann-Raviv, N.; Ayoub, N. NELF-E is recruited to DNA double-strand break sites to promote transcriptional repression and repair. *EMBO Rep.* **2017**, *18*, 745–764. [CrossRef] [PubMed]

32. Borisova, M.E.; Voigt, A.; Tollenaere, M.A.X.; Sahu, S.K.; Juretschke, T.; Kreim, N.; Mailand, N.; Choudhary, C.; Bekker-Jensen, S.; Akutsu, M.; et al. p38-MK2 signaling axis regulates RNA metabolism after UV-light-induced DNA damage. *Nat. Commun.* **2018**, *9*, 1017. [CrossRef]

33. Vos, S.M.; Farnung, L.; Urlaub, H.; Cramer, P. Structure of paused transcription complex Pol II-DSIF-NELF. *Nature* **2018**, *560*, 601–606. [CrossRef] [PubMed]

34. Vos, S.M.; Farnung, L.; Boehning, M.; Wigge, C.; Linden, A.; Urlaub, H.; Cramer, P. Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature* **2018**, *560*, 607–612. [CrossRef] [PubMed]

35. Iida, M.; Iizuka, N.; Tsunedomi, R.; Tsutsui, M.; Yoshida, S.; Maeda, Y.; Tokuhisa, Y.; Sakamoto, K.; Yoshimura, K.; Tamesa, T.; et al. Overexpression of the RD RNA binding protein in hepatitis C virus-related hepatocellular carcinoma. *Oncol. Rep.* **2012**, *28*, 728–734. [CrossRef] [PubMed]

36. Dang, H.; Takai, A.; Forgues, M.; Pomyen, Y.; Mou, H.; Xue, W.; Ray, D.; Ha, K.C.H.; Morris, Q.D.; Hughes, T.R.; et al. Oncogenic Activation of the RNA Binding Protein NELFE and MYC Signaling in Hepatocellular Carcinoma. *Cancer Cell* **2017**, *32*, 101–114.e8. [CrossRef] [PubMed]

37. Chen, S.Y.; Teng, S.C.; Cheng, T.H.; Wu, K.J. miR-1236 regulates hypoxia-induced epithelial-mesenchymal transition and cell migration/invasion through repressing SENP1 and HDAC3. *Cancer Lett.* **2016**, *378*, 59–67. [CrossRef]

38. Butkyte, S.; Ciupas, L.; Jakubauskiene, E.; Vilys, L.; Mocevicius, P.; Kanopka, A.; Vilkaitis, G. Splicing-dependent expression of microRNAs of mirtron origin in human digestive and excretory system cancer cells. *Clin. Epigenetics* **2016**, *8*, 33. [CrossRef]

39. Wang, Y.; Yan, S.; Liu, X.; Zhang, W.; Li, Y.; Dong, R.; Zhang, Q.; Yang, Q.; Yuan, C.; Shen, K.; et al. miR-1236-3p represses the cell migration and invasion abilities by targeting ZEB1 in high-grade serous ovarian carcinoma. *Oncol. Rep.* **2014**, *31*, 1905–1910. [CrossRef]

40. An, J.X.; Ma, M.H.; Zhang, C.D.; Shao, S.; Zhou, N.M.; Dai, D.Q. miR-1236-3p inhibits invasion and metastasis in gastric cancer by targeting MTA2. *Cancer Cell Int.* **2018**, *18*, 66. [CrossRef]

41. Wang, C.; Tang, K.; Li, Z.; Chen, Z.; Xu, H.; Ye, Z. Targeted p21(WAF1/CIP1) activation by miR-1236 inhibits cell proliferation and correlates with favorable survival in renal cell carcinoma. *Urol. Oncol.* **2016**, *34*, 59.e23–59.e34. [CrossRef] [PubMed]

42. Synowsky, S.A.; Heck, A.J. The yeast Ski complex is a hetero-tetramer. *Protein Sci.* **2008**, *17*, 119–125. [CrossRef] [PubMed]

43. Widner, W.R.; Wickner, R.B. Evidence that the SKI antiviral system of Saccharomyces cerevisiae acts by blocking expression of viral mRNA. *Mol. Cell Biol.* **1993**, *13*, 4331–4341. [CrossRef] [PubMed]

44. Ridley, S.P.; Sommer, S.S.; Wickner, R.B. Superkiller mutations in Saccharomyces cerevisiae suppress exclusion of M2 double-stranded RNA by L-A-HN and confer cold sensitivity in the presence of M and L-A-HN. *Mol. Cell Biol.* **1984**, *4*, 761–770. [CrossRef] [PubMed]

45. Brown, J.T.; Bai, X.; Johnson, A.W. The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo. *RNA* **2000**, *6*, 449–457. [CrossRef]

46. Schneider, C.; Tollervey, D. Threading the barrel of the RNA exosome. *Trends Biochem. Sci.* **2013**, *38*, 485–493. [CrossRef]

47. Halbach, F.; Reichelt, P.; Rode, M.; Conti, E. The yeast ski complex: crystal structure and RNA channeling to the exosome complex. *Cell* **2013**, *154*, 814–826. [CrossRef]

48. Anderson, J.S.; Parker, R.P. The 3′ to 5′ degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3′ to 5′ exonucleases of the exosome complex. *EMBO J.* **1998**, *17*, 1497–1506. [CrossRef]

49. Weick, E.M.; Puno, M.R.; Januszyk, K.; Zinder, J.C.; DiMattia, M.A.; Lima, C.D. Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex. *Cell* **2018**, *173*, 1663–1677.e21. [CrossRef]

50. Thoms, M.; Thomson, E.; Bassler, J.; Gnadig, M.; Griesel, S.; Hurt, E. The Exosome Is Recruited to RNA Substrates through Specific Adaptor Proteins. *Cell* **2015**, *162*, 1029–1038. [CrossRef]

51. Lee, S.G.; Song, K. Identification and characterization of a bidirectional promoter from the intergenic region between the human DDX13 and RD genes. *Mol. Cells* **2000**, *10*, 47–53. [CrossRef] [PubMed]

52. Qu, X.; Yang, Z.; Zhang, S.; Shen, L.; Dangel, A.W.; Hughes, J.H.; Redman, K.L.; Wu, L.C.; Yu, C.Y. The human DEVH-box protein Ski2w from the HLA is localized in nucleoli and ribosomes. *Nucleic Acids Res.* **1998**, *26*, 4068–4077. [CrossRef] [PubMed]

53. Fuller-Pace, F.V. RNA helicases: modulators of RNA structure. *Trends Cell Biol.* **1994**, *4*, 271–274. [CrossRef]

54. Lee, S.G.; Song, K. Genomic organization of the human DDX13 gene located between RD and RP1 in the class III MHC complex. *Mol. Cells* **1997**, *7*, 414–418. [PubMed]

55. Yang, Z. Molecular genetic and biochemical studies of the human and mouse MHC complement gene clusters. PhD Dissertation, The Ohio State University, Columbus, OH, USA, 1999.

56. Eckard, S.C.; Rice, G.I.; Fabre, A.; Badens, C.; Gray, E.E.; Hartley, J.L.; Crow, Y.J.; Stetson, D.B. The SKIV2L RNA exosome limits activation of the RIG-I-like receptors. *Nat. Immunol.* **2014**, *15*, 839–845. [CrossRef] [PubMed]

57. Brisse, M.; Ly, H. Comparative Structure and Function Analysis of the RIG-I-Like Receptors: RIG-I and MDA5. *Front. Immunol.* **2019**, *10*, 1586. [CrossRef] [PubMed]

58. Fabre, A.; Charroux, B.; Martinez-Vinson, C.; Roquelaure, B.; Odul, E.; Sayar, E.; Smith, H.; Colomb, V.; Andre, N.; Hugot, J.P.; et al. SKIV2L mutations cause syndromic diarrhea, or trichohepatoenteric syndrome. *Am. J. Hum. Genet.* **2012**, *90*, 689–692. [CrossRef]

59. Hartley, J.L.; Zachos, N.C.; Dawood, B.; Donowitz, M.; Forman, J.; Pollitt, R.J.; Morgan, N.V.; Tee, L.; Gissen, P.; Kahr, W.H.; et al. Mutations in TTC37 cause trichohepatoenteric syndrome (phenotypic diarrhea of infancy). *Gastroenterology* **2010**, *138*, 2388–2398. [CrossRef]

60. Avery, G.B.; Villavicencio, O.; Lilly, J.R.; Randolph, J.G. Intractable diarrhea in early infancy. *Pediatrics* **1968**, *41*, 712–722.

61. Stankler, L.; Lloyd, D.; Pollitt, R.J.; Gray, E.S.; Thom, H.; Russell, G. Unexplained diarrhoea and failure to thrive in 2 siblings with unusual facies and abnormal scalp hair shafts: a new syndrome. *Arch. Dis. Child.* **1982**, *57*, 212–216. [CrossRef]

62. Girault, D.; Goulet, O.; Le Deist, F.; Brousse, N.; Colomb, V.; Cesarini, J.P.; de Potter, S.; Canioni, D.; Griscelli, C.; Fischer, A.; et al. Intractable infant diarrhea associated with phenotypic abnormalities and immunodeficiency. *J. Pediatr.* **1994**, *125*, 36–42. [CrossRef]

63. Dweikat, I.; Sultan, M.; Maraqa, N.; Hindi, T.; Abu-Rmeileh, S.; Abu-Libdeh, B. Tricho-hepato-enteric syndrome: a case of hemochromatosis with intractable diarrhea, dysmorphic features, and hair abnormality. *Am. J. Med. Genet. A* **2007**, *143A*, 581–583. [CrossRef] [PubMed]

64. Fernando, M.M.; Stevens, C.R.; Sabeti, P.C.; Walsh, E.C.; McWhinnie, A.J.; Shah, A.; Green, T.; Rioux, J.D.; Vyse, T.J. Identification of two independent risk factors for lupus within the MHC in United Kingdom families. *PLoS Genet.* **2007**, *3*, e192. [CrossRef] [PubMed]

65. McKay, G.J.; Silvestri, G.; Patterson, C.C.; Hogg, R.E.; Chakravarthy, U.; Hughes, A.E. Further assessment of the complement component 2 and factor B region associated with age-related macular degeneration. *Invest. Ophthalmol. Vis. Sci.* **2009**, *50*, 533–539. [CrossRef] [PubMed]

66. Kopplin, L.J.; Igo, R.P., Jr.; Wang, Y.; Sivakumaran, T.A.; Hagstrom, S.A.; Peachey, N.S.; Francis, P.J.; Klein, M.L.; SanGiovanni, J.P.; Chew, E.Y.; et al. Genome-wide association identifies SKIV2L and MYRIP as protective factors for age-related macular degeneration. *Genes Immun.* **2010**, *11*, 609–621. [CrossRef] [PubMed]

67. Lu, F.; Shi, Y.; Qu, C.; Zhao, P.; Liu, X.; Gong, B.; Ma, S.; Zhou, Y.; Zhang, Q.; Fei, P.; et al. A genetic variant in the SKIV2L gene is significantly associated with age-related macular degeneration in a Han Chinese population. *Invest. Ophthalmol. Vis. Sci.* **2013**, *54*, 2911–2917. [CrossRef] [PubMed]

68. Yoneyama, S.; Sakurada, Y.; Mabuchi, F.; Sugiyama, A.; Kubota, T.; Iijima, H. Genetic variants in the SKIV2L gene in exudative age-related macular degeneration in the Japanese population. *Ophthalmic Genet.* **2014**, *35*, 151–155. [CrossRef] [PubMed]

69. Liu, K.; Chen, L.J.; Tam, P.O.; Shi, Y.; Lai, T.Y.; Liu, D.T.; Chiang, S.W.; Yang, M.; Yang, Z.; Pang, C.P. Associations of the C2-CFB-RDBP-SKIV2L locus with age-related macular degeneration and polypoidal choroidal vasculopathy. *Ophthalmology* **2013**, *120*, 837–843. [CrossRef] [PubMed]

70. Kondo, N.; Honda, S.; Kuno, S.; Negi, A. Role of RDBP and SKIV2L variants in the major histocompatibility complex class III region in polypoidal choroidal vasculopathy etiology. *Ophthalmology* **2009**, *116*, 1502–1509. [CrossRef] [PubMed]

71. Fabre, A.; Martinez-Vinson, C.; Goulet, O.; Badens, C. Syndromic diarrhea/Tricho-hepato-enteric syndrome. *Orphanet. J. Rare Dis.* **2013**, *8*, 5. [CrossRef] [PubMed]

72. Lee, W.S.; Teo, K.M.; Ng, R.T.; Chong, S.Y.; Kee, B.P.; Chua, K.H. Novel mutations in SKIV2L and TTC37 genes in Malaysian children with trichohepatoenteric syndrome. *Gene* **2016**, *586*, 1–6. [CrossRef] [PubMed]

73. Zheng, B.; Pan, J.; Jin, Y.; Wang, C.; Liu, Z. Targeted next-generation sequencing identification of a novel missense mutation of the SKIV2L gene in a patient with trichohepatoenteric syndrome. *Mol. Med. Rep.* **2016**, *14*, 2107–2110. [CrossRef] [PubMed]

74. Hiejima, E.; Yasumi, T.; Nakase, H.; Matsuura, M.; Honzawa, Y.; Higuchi, H.; Okafuji, I.; Yorifuji, T.; Tanaka, T.; Izawa, K.; et al. Tricho-hepato-enteric syndrome with novel SKIV2L gene mutations: A case report. *Medicine (Baltimore)* **2017**, *96*, e8601. [CrossRef] [PubMed]

75. Xinias, I.; Mavroudi, A.; Mouselimis, D.; Tsarouchas, A.; Vasilaki, K.; Roilides, I.; Lacaille, F.; Giouleme, O. Trichohepatoenteric syndrome: A rare mutation in SKIV2L gene in the first Balkan reported case. *SAGE Open Med. Case Rep.* **2018**, *6*, 2050313X18807795. [CrossRef] [PubMed]

76. Ashton, J.J.; Andreoletti, G.; Coelho, T.; Haggarty, R.; Batra, A.; Afzal, N.A.; Beattie, R.M.; Ennis, S. Identification of Variants in Genes Associated with Single-gene Inflammatory Bowel Disease by Whole-exome Sequencing. *Inflamm. Bowel Dis.* **2016**, *22*, 2317–2327. [CrossRef]

77. Vely, F.; Barlogis, V.; Marinier, E.; Coste, M.E.; Dubern, B.; Dugelay, E.; Lemale, J.; Martinez-Vinson, C.; Peretti, N.; Perry, A.; et al. Combined Immunodeficiency in Patients With Trichohepatoenteric Syndrome. *Front. Immunol.* **2018**, *9*, 1036. [CrossRef] [PubMed]

78. Bourgeois, P.; Esteve, C.; Chaix, C.; Beroud, C.; Levy, N.; consortium, T.c.; Fabre, A.; Badens, C. Tricho-Hepato-Enteric Syndrome mutation update: Mutations spectrum of TTC37 and SKIV2L, clinical analysis and future prospects. *Hum. Mutat.* **2018**, *39*, 774–789. [CrossRef]

79. Verloes, A.; Lombet, J.; Lambert, Y.; Hubert, A.F.; Deprez, M.; Fridman, V.; Gosseye, S.; Rigo, J.; Sokal, E. Tricho-hepato-enteric syndrome: Further delineation of a distinct syndrome with neonatal hemochromatosis phenotype, intractable diarrhea, and hair anomalies. *Am. J. Med. Genet.* **1997**, *68*, 391–395. [CrossRef]

80. Poulton, C.; Pathak, G.; Mina, K.; Lassman, T.; Azmanov, D.N.; McCormack, E.; Broley, S.; Dreyer, L.; Gration, D.; Taylor, E.; et al. Tricho-hepatic-enteric syndrome (THES) without intractable diarrhoea. *Gene* **2019**, *699*, 110–114. [CrossRef]

81. Monies, D.M.; Rahbeeni, Z.; Abouelhoda, M.; Naim, E.A.; Al-Younes, B.; Meyer, B.F.; Al-Mehaidib, A. Expanding phenotypic and allelic heterogeneity of tricho-hepato-enteric syndrome. *J. Pediatr. Gastroenterol. Nutr.* **2015**, *60*, 352–356. [CrossRef]

82. Lintner, K.E.; Patwardhan, A.; Rider, L.G.; Abdul-Aziz, R.; Wu, Y.L.; Lundstrom, E.; Padyukov, L.; Zhou, B.; Alhomosh, A.; Newsom, D.; et al. Gene copy-number variations (CNVs) of complement C4 and C4A deficiency in genetic risk and pathogenesis of juvenile dermatomyositis. *Ann. Rheum. Dis.* **2016**, *75*, 1599–1606. [CrossRef] [PubMed]

83. Baechler, E.C.; Batliwalla, F.M.; Karypis, G.; Gaffney, P.M.; Ortmann, W.A.; Espe, K.J.; Shark, K.B.; Grande, W.J.; Hughes, K.M.; Kapur, V.; et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2610–2615. [CrossRef] [PubMed]

84. Ronnblom, L.; Pascual, V. The innate immune system in SLE: type I interferons and dendritic cells. *Lupus* **2008**, *17*, 394–399. [CrossRef] [PubMed]

85. Xue, Y.; Bai, X.; Lee, I.; Kallstrom, G.; Ho, J.; Brown, J.; Stevens, A.; Johnson, A.W. Saccharomyces cerevisiae RAI (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p. *Mol. Cell Biol.* **2000**, *20*, 4006–4015. [CrossRef] [PubMed]

86. Xiang, S.; Cooper-Morgan, A.; Jiao, X.; Kiledjian, M.; Manley, J.L.; Tong, L. Structure and function of the 5′–>3′ exoribonuclease Rat1 and its activating partner Rai1. *Nature* **2009**, *458*, 784–788. [CrossRef] [PubMed]

87. Jiao, X.; Xiang, S.; Oh, C.; Martin, C.E.; Tong, L.; Kiledjian, M. Identification of a quality-control mechanism for mRNA 5′-end capping. *Nature* **2010**, *467*, 608–611. [CrossRef] [PubMed]

88. Chang, J.H.; Jiao, X.; Chiba, K.; Oh, C.; Martin, C.E.; Kiledjian, M.; Tong, L. Dxo1 is a new type of eukaryotic enzyme with both decapping and 5′-3′ exoribonuclease activity. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1011–1017. [CrossRef] [PubMed]

89. Jiao, X.; Chang, J.H.; Kilic, T.; Tong, L.; Kiledjian, M. A mammalian pre-mRNA 5′ end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol. Cell* **2013**, *50*, 104–115. [CrossRef] [PubMed]

90. Zhai, L.T.; Xiang, S. mRNA quality control at the 5′ end. *J. Zhejiang Univ. Sci. B* **2014**, *15*, 438–443. [CrossRef]

91. Kiledjian, M. Eukaryotic RNA 5′-End NAD(+) Capping and DeNADding. *Trends Cell Biol* **2018**, *28*, 454–464. [CrossRef]

92. Picard-Jean, F.; Brand, C.; Tremblay-Letourneau, M.; Allaire, A.; Beaudoin, M.C.; Boudreault, S.; Duval, C.; Rainville-Sirois, J.; Robert, F.; Pelletier, J.; et al. Correction: 2′-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. *PLoS ONE* **2018**, *13*, e0202308. [CrossRef] [PubMed]

93. Jiao, X.; Doamekpor, S.K.; Bird, J.G.; Nickels, B.E.; Tong, L.; Hart, R.P.; Kiledjian, M. 5′ End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. *Cell* **2017**, *168*, 1015–1027.e10. [CrossRef] [PubMed]

94. Belt, K.T.; Yu, C.Y.; Carroll, M.C.; Porter, R.R. Polymorphism of human complement component C4. *Immunogenetics* **1985**, *21*, 173–180. [CrossRef] [PubMed]

95. Yang, Z.; Mendoza, A.R.; Welch, T.R.; Zipf, W.B.; Yu, C.Y. Modular variations of the human major histocompatibility complex class III genes for serine/threonine kinase RP, complement component C4, steroid 21-hydroxylase CYP21, and tenascin TNX (the RCCX module). A mechanism for gene deletions and disease associations. *J. Biol. Chem.* **1999**, *274*, 12147–12156. [PubMed]

96. Yu, C.Y.; Chung, E.K.; Yang, Y.; Blanchong, C.A.; Jacobsen, N.; Saxena, K.; Yang, Z.; Miller, W.; Varga, L.; Fust, G. Dancing with complement C4 and the RP-C4-CYP21-TNX (RCCX) modules of the major histocompatibility complex. *Prog. Nucleic Acid Res. Mol. Biol.* **2003**, *75*, 217–292. [PubMed]

97. Sargent, C.A.; Anderson, M.J.; Hsieh, S.L.; Kendall, E.; Gomez-Escobar, N.; Campbell, R.D. Characterisation of the novel gene G11 lying adjacent to the complement C4A gene in the human major histocompatibility complex. *Hum. Mol. Genet.* **1994**, *3*, 481–488. [CrossRef] [PubMed]

98. Gomez-Escobar, N.; Chou, C.F.; Lin, W.W.; Hsieh, S.L.; Campbell, R.D. The G11 gene located in the major histocompatibility complex encodes a novel nuclear serine/threonine protein kinase. *J. Biol. Chem.* **1998**, *273*, 30954–30960. [CrossRef] [PubMed]

99. Blanchong, C.A.; Zhou, B.; Rupert, K.L.; Chung, E.K.; Jones, K.N.; Sotos, J.F.; Zipf, W.B.; Rennebohm, R.M.; Yu, C.Y. Deficiencies of human complement component C4A and C4B and heterozygosity in length variants of RP-C4-CYP21-TNX (RCCX) modules in caucasians. The load of RCCX genetic diversity on major histocompatibility complex-associated disease. *J. Exp. Med.* **2000**, *191*, 2183–2196. [CrossRef] [PubMed]

100. Ostertag, E.M.; Goodier, J.L.; Zhang, Y.; Kazazian, H.H., Jr. SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am. J. Hum. Genet.* **2003**, *73*, 1444–1451. [CrossRef]

101. Hancks, D.C.; Kazazian, H.H., Jr. SVA retrotransposons: Evolution and genetic instability. *Semin Cancer Biol* **2010**, *20*, 234–245. [CrossRef]

102. Savage, A.L.; Bubb, V.J.; Breen, G.; Quinn, J.P. Characterisation of the potential function of SVA retrotransposons to modulate gene expression patterns. *BMC Evol. Biol.* **2013**, *13*, 101. [CrossRef] [PubMed]

103. Lehner, B.; Semple, J.I.; Brown, S.E.; Counsell, D.; Campbell, R.D.; Sanderson, C.M. Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region. *Genomics* **2004**, *83*, 153–167. [CrossRef]

104. STK19 serine/threonine kinase 19 [Homo sapiens (human)]. Available online: https://www.ncbi.nlm.nih.gov/gene/8859 (accessed on 1 June 2019).

105. Boeing, S.; Williamson, L.; Encheva, V.; Gori, I.; Saunders, R.E.; Instrell, R.; Aygun, O.; Rodriguez-Martinez, M.; Weems, J.C.; Kelly, G.P.; et al. Multiomic Analysis of the UV-Induced DNA Damage Response. *Cell Rep.* **2016**, *15*, 1597–1610. [CrossRef] [PubMed]

106. Hodis, E.; Watson, I.R.; Kryukov, G.V.; Arold, S.T.; Imielinski, M.; Theurillat, J.P.; Nickerson, E.; Auclair, D.; Li, L.; Place, C.; et al. A landscape of driver mutations in melanoma. *Cell* **2012**, *150*, 251–263. [CrossRef] [PubMed]

107. Shitara, D.; Tell-Marti, G.; Badenas, C.; Enokihara, M.M.; Alos, L.; Larque, A.B.; Michalany, N.; Puig-Butille, J.A.; Carrera, C.; Malvehy, J.; et al. Mutational status of naevus-associated melanomas. *Br. J. Dermatol.* **2015**, *173*, 671–680. [CrossRef] [PubMed]

108. Bonilla, X.; Parmentier, L.; King, B.; Bezrukov, F.; Kaya, G.; Zoete, V.; Seplyarskiy, V.B.; Sharpe, H.J.; McKee, T.; Letourneau, A.; et al. Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma. *Nat. Genet.* **2016**, *48*, 398–406. [CrossRef] [PubMed]

109. Yin, C.; Zhu, B.; Zhang, T.; Liu, T.; Chen, S.; Liu, Y.; Li, X.; Miao, X.; Li, S.; Mi, X.; et al. Pharmacological Targeting of STK19 Inhibits Oncogenic NRAS-Driven Melanomagenesis. *Cell* **2019**, *176*, 1113–1127.e16. [CrossRef] [PubMed]

110. Ishikawa, T. Next-generation sequencing traces human induced pluripotent stem cell lines clonally generated from heterogeneous cancer tissue. *World J. Stem Cells* **2017**, *9*, 77–88. [CrossRef]

111. Kraja, A.T.; Chasman, D.I.; North, K.E.; Reiner, A.P.; Yanek, L.R.; Kilpelainen, T.O.; Smith, J.A.; Dehghan, A.; Dupuis, J.; Johnson, A.D.; et al. Pleiotropic genes for metabolic syndrome and inflammation. *Mol. Genet. Metab.* **2014**, *112*, 317–338. [CrossRef]

112. Amare, A.T.; Vaez, A.; Hsu, Y.H.; Direk, N.; Kamali, Z.; Howard, D.M.; McIntosh, A.M.; Tiemeier, H.; Bultmann, U.; Snieder, H.; et al. Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. *Mol. Psychiatry* **2019**. [CrossRef]

113. Negative elongation factor E [Drosophila melanogaster]. Available online: https://www.ncbi.nlm.nih.gov/protein/NP_648241.1 (accessed on 1 June 2019).

114. Negative elongation factor E isoform X2 [Danio rerio]. Available online: https://www.ncbi.nlm.nih.gov/protein/XP_005170084.1 (accessed on 1 June 2019).

115. Paulsen, J.E.; Capowski, E.E.; Strome, S. Phenotypic and molecular analysis of mes-3, a maternal-effect gene required for proliferation and viability of the germ line in C. elegans. *Genetics* **1995**, *141*, 1383–1398. [PubMed]

116. Session, A.M.; Uno, Y.; Kwon, T.; Chapman, J.A.; Toyoda, A.; Takahashi, S.; Fukui, A.; Hikosaka, A.; Suzuki, A.; Kondo, M.; et al. Genome evolution in the allotetraploid frog Xenopus laevis. *Nature* **2016**, *538*, 336–343. [CrossRef] [PubMed]

117. Milner, C.M.; Campbell, R.D. Genetic organization of the human MHC class III region. *Front. Biosci.* **2001**, *6*, D914–D926. [CrossRef] [PubMed]

118. Janeway, C.A., Jr.; Medzhitov, R. Innate immune recognition. *Annu. Rev. Immunol.* **2002**, *20*, 197–216. [CrossRef] [PubMed]

119. Medzhitov, R.; Janeway, C., Jr. Innate immunity. *N. Engl. J. Med.* **2000**, *343*, 338–344. [CrossRef] [PubMed]

120. Note added at update: A recent review that is highly relevant and provides comprehensive background knowledge to this article: Cramer, P. Organization and regulation of gene transcription. *Nature* **2019**, *573*, 45–54.