Genes & Diseases (2022) 9, 941–958

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://ees.elsevier.com/gendis/default.asp

REVIEW ARTICLE

The pseudogenes of eukaryotic translation elongation factors (EEFs): Role in cancer and other human diseases

Luigi Cristiano

R&D Division, Prestige, 18 via Vecchia, Terranuova Bracciolini, AR 52028, Italy

Received 11 November 2020; accepted 29 March 2021
Available online 16 April 2021

KEYWORDS
Cancer;
EEFs;
Non-coding RNA;
Pseudogene;
Translation;
Translation elongation factor

Abstract The eukaryotic translation elongation factors (EEFs), i.e. *EEF1A1*, *EEF1A2*, *EEF1B2*, *EEF1D*, *EEF1G*, *EEF1E1* and *EEF2*, are coding-genes that play a central role in the elongation step of translation but are often altered in cancer. Less investigated are their pseudogenes. Recently, it was demonstrated that pseudogenes have a key regulatory role in the cell, especially via non-coding RNAs, and that the aberrant expression of ncRNAs has an important role in cancer development and progression. The present review paper, for the first time, collects all that published about the EEFs pseudogenes to create a base for future investigations. For most of them, the studies are in their infancy, while for others the studies suggest their involvement in normal cell physiology but also in various human diseases. However, more investigations are needed to understand their functions in both normal and cancer cells and to define which can be useful biomarkers or therapeutic targets.

Introduction

The eukaryotic translation elongation factors (EEFs) play a central role in the proteins biosynthesis during the elongation step of translation (Fig. 1). They include the eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*), eukaryotic translation elongation factor 1 alpha 2 (*EEF1A2*), eukaryotic translation elongation factor 1 beta 2 (*EEF1B2*), eukaryotic translation elongation factor 1 delta (*EEF1D*), eukaryotic translation elongation factor 1 gamma (*EEF1G*), eukaryotic translation elongation factor 1 epsilon 1 (*EEF1E1*), and eukaryotic translation elongation factor-2 (*EEF2*). These genes, and related proteins, can be grouped into two large subfamilies, namely non-alpha EEFs and alpha EEFs. Many published studies reported their biological significance as well as their involvement in cancer and

E-mail address: luigicristiano@libero.it.
Peer review under responsibility of Chongqing Medical University.

https://doi.org/10.1016/j.gendis.2021.03.009

Abbreviations

CCS-3	cervical cancer suppressor 3
ceRNA	Competitive endogenous RNA
EEF1A1	Eukaryotic translation elongation factor 1 alpha 1
EEF1A1L14	Eukaryotic translation elongation factor 1-alpha 1-like 14
EEF1A2	Eukaryotic translation elongation factor 1 alpha 2
EEF1B2	Eukaryotic translation elongation factor 1 beta 2
EEF1D	Eukaryotic translation elongation factor 1 delta
EEF1E1	Eukaryotic translation elongation factor 1 epsilon 1
EEF1G	Eukaryotic Translation Elongation Factor 1 gamma
eEF1H	eukaryotic translation elongation factor-1 macromolecular complex
EEF2	eukaryotic translation elongation factor-2
EEFs	Eukaryotic translation elongation factors
lncRNAs	Long non-coding RNAs
MARS	Multiaminoacyl-tRNA synthetase macromolecular complex
ncRNAs	Non-coding RNAs
PTI-1	Prostate tumor-inducinge gene-1 (alias, EEF1A1L14)
sncRNAs	Short non-coding RNAs

other human diseases. Nevertheless, the role and biological function of their pseudogenes in normal and pathological states are still poorly studied.

Until recently, pseudogenes were believed to be junk DNA, i.e. relics, non-functional versions, of parental protein-coding genes no longer able to encode a protein and devoid of any biological significance or usefulness. Recent transcriptomic and proteomic analyses have shown that both pseudogene-derived transcripts and pseudogene-derived proteins or pseudogene-derived short-peptides can be found,[2–5] thus demonstrating that pseudogenes play a biological role in the cell. In fact, they can be positive or negative regulators of the genome, transcriptome and proteome.

At the DNA level, a pseudogene can affect its parental gene in many ways, including homologous recombination, transfer of small DNA sequences (gene conversion) and enhance or inhibit its transcription.[6]

At the RNA level, a pseudogene can affect the expression of its parental gene with different mechanisms, involving one or more of its own transcripts.[6] These, often called pseudo-mRNAs (or ψmRNAs), are frequently non-coding RNAs (ncRNAs)[7] and include long ncRNAs (lncRNAs) and short ncRNAs (sncRNAs). Thus, transcribed pseudogenes, via their transcripts, may act as positive or negative regulators of parental gene expression[8] in many ways, including the recently discovered mechanism of the competitive endogenous RNA (ceRNA) network.[9] Furthermore, it is known as the aberrant expression of ncRNAs plays a key role in the development and progression of cancer.[2,9,10] The alteration in the expression levels of the pseudogenes, both transcribed pseudogenes and untranscribed pseudogenes, especially in cancer, could in fact have direct or indirect consequences on the cell.

At the protein level, pseudogene-derived proteins or pseudogene-derived short-peptides can positively or negatively affect the activity of the parental protein. Furthermore, a protein product of a pseudogene may have biological activity in tissues where the parental gene is not expressed or in other cellular compartments. The role of a protein produced by a pseudogene can also be revealed only in a pathologic condition such as cancer.[6]

The expression profile of pseudogenes has been reported to vary in different tissues, under different conditions, both physiological than pathological,[11] but it cannot be excluded that it varies over time, i.e. during embryogenesis[12] and/or childhood or adulthood, as well as it can be acquired somatically, as shown during cancer development.[13]

Pseudogenes are classified into three main categories: processed pseudogenes, unprocessed pseudogenes, and unitary pseudogenes.[14,15] Processed pseudogenes (PPs) are pseudogenes devoid of introns and other regulatory elements (such as enhancers and promoter) and derive from the reverse transcription of mRNA followed by the re-insertion of respective DNA (cDNA) into the genome (retrotransposition) and therefore are often also called retro-pseudogenes. In this regard, the copy number of a retropseudogene could be related to the expression level of the gene from which it derives. Furthermore, they can be found in new locations on different chromosomes than their parental coding-gene and many of them have been reported to be actively transcribed.[16,17]

The unprocessed pseudogenes, on the other hand, can contain introns and regulatory sequences. They result from gene duplication during unequal crossing-over and are generally found on the same chromosome of the parental protein-coding gene. The subcategory of transcribed pseudogenes, both unprocessed and processed, shows one or more transcripts. Finally, the unitary pseudogenes (orphans) are considered to be previously active genes that become inactive due to mutations and genomic alterations and have no homologous active gene in the genome.

This review paper, for the first time, collects and summarizes all that are known and currently published on EEFs pseudogenes to create a state of the art from which to build further research and insights.

Materials and methods

A list of annotated pseudogenes by each EEF gene was obtained from NCBI:Gene (https://www.ncbi.nlm.nih.gov/gene/) by typing the official symbol of the parental gene and then searching for annotated pseudogenes on its profile on the "General gene information" sub-tab.

Each pseudogene is searched for published papers by typing its official symbol on Pubmed (https://pubmed.ncbi.nlm.nih.gov/), Academia (https://www.academia.edu/),

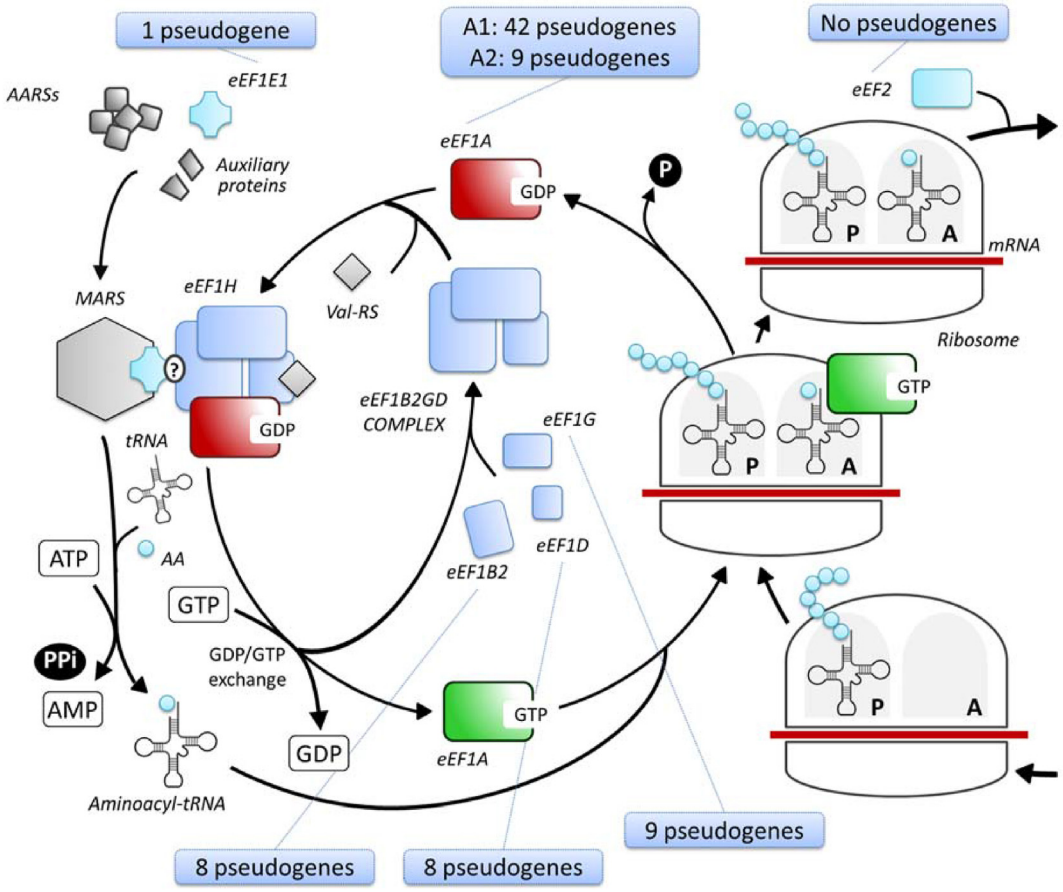

Figure 1 The elongation step of translation. The active form of eEF1A (eEF1A-GTP), delivers an aminoacylated tRNA to the A site of the ribosome. Following the proper codon-anticodon recognition the GTP is hydrolyzed and the inactive eEF1A-GDP is released from the ribosome and then it is bound by eEF1H protein complex. eEF1H is formed previously by the binding of eEF1B2, eEF1G, eEF1D and Val-RS. This complex promotes the exchange between GDP and GTP to regenerate the active form of eEF1A. eEF1E1 collaborates to anchor MARS complex to eEF1H. eEF2 is subsequently involved for ribosome translocation. A box is added to each EEFs indicating the number of pseudogenes known so far.[1]

ResearchGate (https://www.researchgate.net/) and Google Scholar (https://scholar.google.it/).

To complete the data collection, the search is extended to other datasets and databases where many different types of information on sequences, transcripts, levels of expression and related characteristics are reported. In particular, the data used for this review were obtained from NCBI:Geo Profiles (https://www.ncbi.nlm.nih.gov/geoprofiles/), Open Targets Platform (https://www.targetvalidation.org/), Ensembl (http://www.ensembl.org/index.html), Oasis/Pfizer (http://www.oasis-genomics.org/), FusionHub (https://fusionhub.persistent.co.in/), GenAtlas/Paris (http://genatlas.medecine.univ-paris5.fr/), Atlas of Genetics and Cytogenetics in Oncology and Haematology (http://atlasgeneticsoncology.org/index.html),[18] GeneCards/Weizmann (https://www.genecards.org/), Source/Princeton (https://source-search.princeton.edu/cgi-bin/source/sourceSearch), Gwas Catalog (https://www.ebi.ac.uk/gwas/), HGNC (https://www.genenames.org/), GTEx Portal (https://www.gtexportal.org/home/), cBioPortal (https://www.cbioportal.org/), OncoMX (https://oncomx.org/searchview/) and FireBrowse (http://www.firebrowse.org/).

Pseudogenes of non-alpha eukaryotic translation elongation factors

Non-alpha EEFs collect nearly all components of the eukaryotic translation elongation factor-1 macromolecular complex (eEF1H), namely eEF1B2, eEF1D and eEF1G, as well as a component of multiaminoacyl-tRNA synthetase macromolecular complex (MARS), that is eEF1E1, and eEF2. All of these genes encode at least one protein, but more frequently several protein isoforms, which play a central role in peptide elongation during protein biosynthesis. eEF1B2, eEF1D, and eEF1G join the valyl t-RNA synthetase (valRS) to form the macromolecular complex eEF1BGD which is involved in the regeneration of the active form of eEF1A, i.e. converts the inactive GDP-bound form of eEF1A (eEF1A-GDP) into its active GTP-bound form (eEF1A-GTP).[19] eEF1E1 interacts with different aminoacyl-tRNA synthetases[20] and could contribute to the anchoring of the macromolecular aminoacyl-tRNA synthetases complex (MARS) to the EF1H complex in the translation elongation step.[21] Finally, eEF2 is required for translocation of the peptidyl-tRNA from A-site to P-site of the ribosome.

All these factors exhibit canonical functions and multiple non-canonical roles (moonlight roles) within the cell[22] and are frequently altered in expression, gene amplification and genomic rearrangements in many cancers and other diseases.[23] All have at least one pseudogene, but more frequently more than one, dispersed in the human genome (Fig. 2) with the exception of *EEF2* for which no pseudogenes in humans are known.

The pseudogenes reported for non-alpha EEFs are classified into processed pseudogenes, unprocessed pseudogenes and transcribed unprocessed pseudogenes. These pseudogenes are listed in Table 1A (more detail in the T1ASuppl supplementary material). All non-alpha EEF coding genes, briefly, and their pseudogenes, more extensively, will be treated individually.

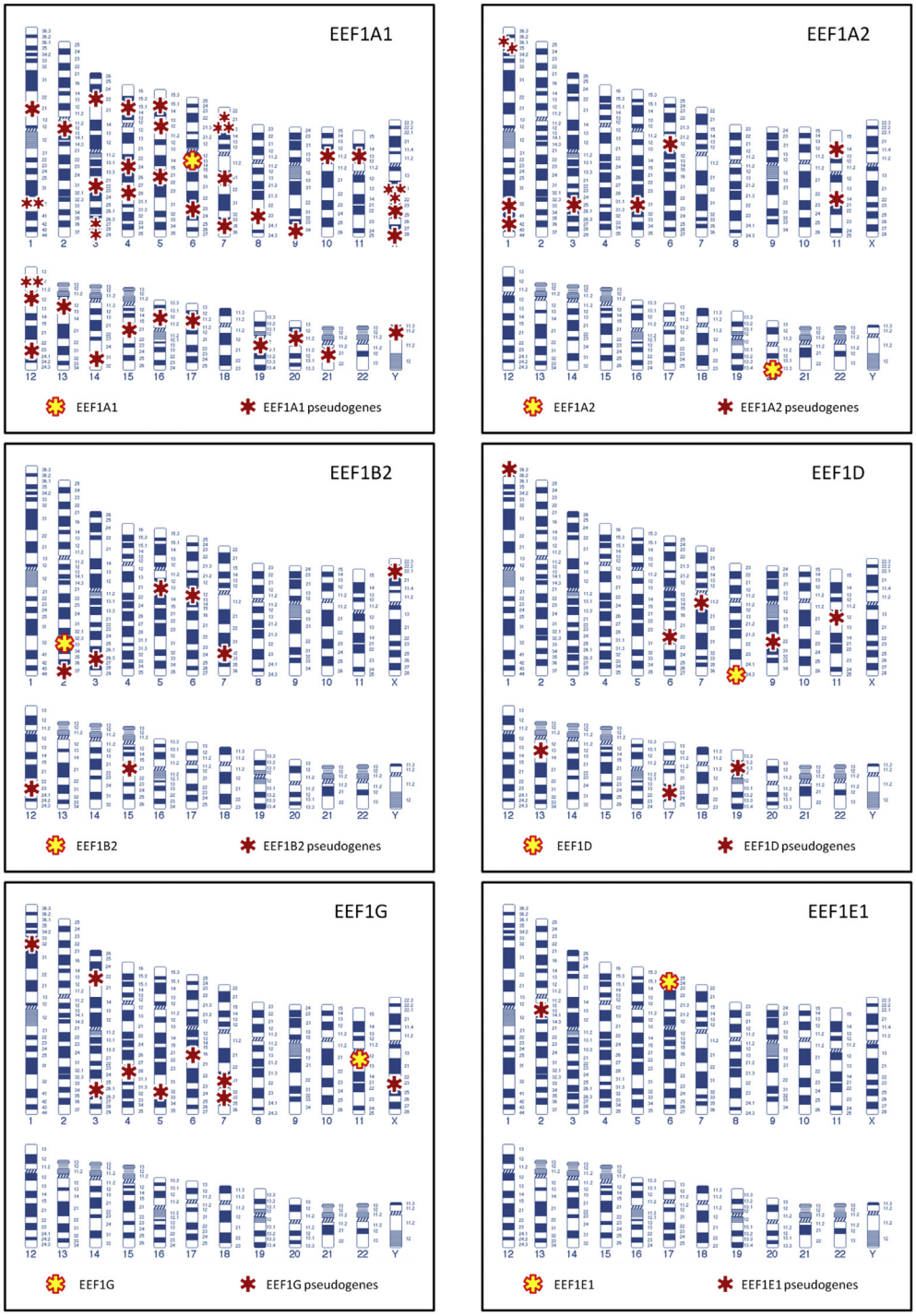

Figure 2 Localization of EEFs pseudogenes. The figure shows the locations of each pseudogene and its respective parental gene in the human genome. The data have been extracted from Gene (NCBI).

Table 1A Pseudogenes of non-alpha EEFs. List of all non-alpha EEFs pseudogenes so far discovered and the correlation with diseases where they are reported or there is evidence about them (see also supplementary table TA1SUPPL).

RFG	PS	Description	Status	CHR	Location	Length (nt)	Main diseases
EEF1B2	*EEF1B2P1* [24,25]	*EEF1B2* pseudogene 1	Processed pseudogene	15	15q21.2	880	Non-squamous non-small cell lung cancer (NSCLC) (?)[26]
	EEF1B2P2 [27,28]	*EEF1B2* pseudogene 2		5	5q13.1	803	—
	EEF1B2P3 [28]	*EEF1B2* pseudogene 3		X	Xp22.11	764	Human bone osteosarcoma epithelial cell line) (U2OS) (?), acute myeloid leukemia (AML) cell lines (KG-1, MOLM-14) (?),Hepatocellular carcinoma (?), HIV-1 reverse transcription cofactor (?)[29–32]
	EEF1B2P4	*EEF1B2* pseudogene 4		12	12q23.3	1161	—
	EEF1B2P5	*EEF1B2* pseudogene 5	Unprocessed pseudogene	6	6q12	1877	—
	EEF1B2P6 [33]	*EEF1B2* pseudogene 6	Processed pseudogene	7	7q32.3	766	—
	EEF1B2P7	*EEF1B2* pseudogene 7		2	2q37.1	799	—
	EEF1B2P8	*EEF1B2* pseudogene 8		3	3q26.31	796	—
EEF1D	*EEF1DP1* [34]	*EEF1D* pseudogene 1	Processed pseudogene	19	19p13.12	980	Acute myeloid leukemia cell lines (HL-60, MOLM-14, THP-1, U937) (?), diffuse large B-cell lymphoma cell lines (DHL4, DHL6) (?), hepatocellular carcinoma cell line (Huh-7) (?), Human bone osteosarcoma epithelial cell line) (U2OS) (?), melanoma (?)[29]
	EEF1DP2	*EEF1D* pseudogene 2		9	9q22.31	976	Melanoma (?)
	EEF1DP3 [35,36]	*EEF1D* pseudogene 3	Transcribed unprocessed pseudogene	13	13q13.1	575	Prostate carcinoma, breast carcinoma, ankylosing spondylitis, adrenocortical carcinoma (ACC), pheochromocytoma and Paraganglioma (PCPG), brain lower-grade glioma (LGG), rectum adenocarcinoma (READ), cervical squamous cell carcinoma, endocervical adenocarcinoma (CESC), uterine carcinosarcoma (UCS), head and neck squamous cell carcinoma (HNSC), hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), mesothelioma, acute myeloid leukemia (AML), lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), skin cutaneous melanoma (SKCM), pancreatic adenocarcinoma (PAAD), sarcoma (SARC), bladder urothelial carcinoma (BLCA), chromophobe renal cell carcinoma (KICH), kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), synucleinopathy and Parkinson's disease (?), non-small cell lung cancer (?), multiple sclerosis (?), large B-cell lymphoma cell lines (SUDHL4, Toledo, OCI-Ly3) (?), epidermolysis bullosa simplex (?)[37-45]

(*continued on next page*)

Table 1A (*continued*)

RFG	PS	Description	Status	CHR	Location	Length (nt)	Main diseases
	EEF1DP4 [46]	*EEF1D* pseudogene 4	Processed pseudogene	7	7q11.21	1500	Breast carcinoma (?), colon cancer (?), glioma (?), osteosarcoma (?), primary myelofibrosis (?)
	EEF1DP5	*EEF1D* pseudogene 5		6	6q22.33	888	Breast carcinoma, Human bone osteosarcoma epithelial cell line) (U2OS) (?)[29,47]
	EEF1DP6	*EEF1D* pseudogene 6		1	1p36.32	437	Acute myeloid leukemia (?), systemic juvenile idiopathic arthritis (?), neuropathy in Charcot-Marie-Tooth disease type 1A (?)[48–50]
	EEF1DP7	*EEF1D* pseudogene 7	Transcribed unprocessed pseudogene	17	17q23.3	510	—
	EEF1DP8	*EEF1D* pseudogene 8	Processed pseudogene	11	11q12.3	609	—
EEF1G	*EEF1GP1* [46]	*EEF1G* Pseudogene 1	Processed pseudogene	7	7q31.33	2151	Human bone osteosarcoma epithelial cell line) (U2OS) (?)[29]
	EEF1GP2	*EEF1G* Pseudogene 2		5	5q32	835	—
	EEF1GP3	*EEF1G* Pseudogene 3		3	3p22.1	1391	—
	EEF1GP4 [51–53]	*EEF1G* Pseudogene 4		3	3q26.1	1402	—
	EEF1GP5	*EEF1G* Pseudogene 5		X	Xq23	1405	Prostate carcinoma (?), Duchenne muscular dystrophy (?), Human bone osteosarcoma epithelial cell line) (U2OS) (?)[29]
	EEF1GP6	*EEF1G* Pseudogene 6		6	6q16.1	929	—
	EEF1GP7	*EEF1G* Pseudogene 7		1	1p32.3	793	—
	EEF1GP8	*EEF1G* Pseudogene 8		4	4q28.2	1241	—
	LOC729998	*EEF1G* Pseudogene 9		7	7q33	1412	Acute myeloid leukemia cell lines (HL-60, MOLM-14, THP-1, U937) (?), diffuse large B-cell lymphoma cell lines (DHL4, DHL6) (?), hepatocellular carcinoma cell line (Huh-7) (?)
EEF1E1	*EEF1E1P1*	*EEF1E1* pseudogene 1	Processed pseudogene	2	2q13	1596	Coronary artery disease (?)[54]
EEF2	—	—	—	—	—	—	—

Abbreviations: RFG, related functional gene; PS, pseudogene; CHR, Chromosome; [(?)], uncertain; [-], unknown.

Pseudogenes of *EEF1B2*

EEF1B2, also known as *eEF1β* or *eEF1Bα*, is a coding-gene located on Chromosome 2 (2q33.3). Several alternative splicing transcript variants have been observed[55] but to date only one protein has been detected.[56] Like the other members of the eEF1H complex, it is involved in the elongation step of translation and collaborates closely with eEF1D and eEF1G in the conversion of eEF1A from its inactive GDP-bound form to its active GTP-bound form.[56,57]

Analysis of the sequences reported in the human genome revealed the presence of eight pseudogenes for *EEF1B2* which are mostly classified as processed pseudogenes[55] and probably related to recent retrotransposition events.[24] The alternative forms *EEF1B3* and *EEF1B4*, previously designated for *EEF1B2*,[28] instead have shown to be pseudogenes namely *EEF1B2P2* and *EEF1B2P3* respectively. However, the pseudogenes of *EEF1B2* are poorly studied and publications have been made only for some of them.

The *EEF1B2* pseudogene 1, alias *EEF1B2P1*, was first reported in 1991.[25] It was first referred to as a gene paralogue

of *EEF1B2*, named *EEF1B1*, but was latter better described as a processed pseudogene.[24] It has been studied as a baseline putative marker for the prediction of overall patient survival in advanced non-squamous non-small cell lung cancer (NSCLC)[26] but its biological significance in this cancer is unknown.

The *EEF1B2* pseudogene 2, alias *EEF1B2P2* was first reported in 1993[27] as an isoform of *EEF1B2* called EF-1β5a[28] but it has subsequently been classified as a processed pseudogene. A transcript of this pseudogene was found in the human brain and muscle where this isoform replaces the transcription of *EEF1B2*.[24]

The *EEF1B2* pseudogene 3, alias *EEF1B2P3*, was first reported in 1993[28] and later in an analysis of gene cluster in the human bone osteosarcoma epithelial cell line (*U2OS*)[29] and in hepatocellular carcinoma,[31] but its significance in these diseases is unknown. Some studies report differences in its expression levels: in particular, it has been found to be upregulated during *HIV-1* infection, so it may be a critical reverse transcription cofactor of *HIV-1*.[32] However, it is not clear why. Furthermore, the expression levels of *EEF1B2P3* decrease after the use of a dihydroorotate dehydrogenase inhibitor in KG-1 and *MOLM-14* acute myeloid leukaemia (AML) cell lines.[30] The *EEF1B2* pseudogene 6, alias *EEF1B2P6*, was first reported in 2007[33] but no other studies have been conducted.

The others, i.e. the *EEF1B2* pseudogene 4 (alias *EEF1B2P4*) the *EEF1B2* pseudogene 5 (alias *EEF1B2P5*), the *EEF1B2* pseudogene 7 (alias *EEF1B2P7*) and the *EEF1B2* pseudogene 8 (alias *EEF1B2P8*), are predicted by genome sequence analysis but are not yet supported by experimental evidence.

Pseudogenes of *EEF1D*

EEF1D, alias *eEF1δ* or *eEF1Bδ*, is a coding gene with several alternative splicing transcript variants that encode several protein isoforms.[58] Like the other members of the eEF1H complex, it is involved in the elongation step of translation and closely collaborates with eEF1B2 in the conversion of eEF1A from its inactive GDP-bound form to its active GTP-bound form.[56,57] Analysis of the human genome revealed the presence of eight pseudogenes. Some are poorly characterized while others are better known, especially *EEF1DP3*.

The *EEF1D* pseudogene 1, alias *EEF1DP1*, was first reported in 2001[34] and in datasets on some cancer cell lines of hepatocellular carcinoma, acute myeloid leukemia, diffuse large B-cell lymphoma, human bone osteosarcoma (*U2OS*)[29] and melanoma without specific information, so its significance in these diseases is unclear. The same happens for *EEF1D* pseudogene 2, alias *EEF1DP2*, that is reported in datasets on melanoma. It is not entirely clear whether it is expressed or expressed even at a low level.

The *EEF1D* pseudogene 3, alias *EEF1DP3*, is the most studied pseudogene compared to the others of *EEF1D* and, in general, of all non-alpha EEFs pseudogenes. First reported in 2005[36], but also later by other authors,[35] it is found on chromosome 13. To note that chromosome 13 is known to carry some putative oncogenes involved in cancer, including breast cancer type 2 (*BRCA2*) and retinoblastoma (*RB1*) genes.[59]

EEF1DP3 is classified as a transcribed unprocessed pseudogene and the genomic sequence contains four non-coding exons. It is not yet known it undergoes post-transcriptional modifications, however it is transcribed and produces a long non-coding RNA (lncRNA) of 575 nt.[35] It is known that lncRNAs, like other ncRNAs, can modulate gene expression both at the transcriptional level, interacting with the parental gene promoter, and at the post-transcriptional level, acting as microRNA decoys and thus may play key roles in cellular biological processes.[2,8,14] Nowadays, the exact role of *EEF1DP3* in healthy tissues is still unknown, however, it has been reported to be over-expressed in the heart, particularly in the left ventricle and is also expressed in the normal trachea, liver, testis, kidney, bladder and brain. Conversely, a low expression is found in the adrenal gland, colon and pituitary gland.[60]

Numerous mutations and alterations in the genomic sequence for *EEF1DP3* have been discovered which include copy number variations, translocations and interchromosomal translocations with the formation of novel fusion genes. These have been found in many kinds of cancer, such as in breast cancer[39–42] and Burkitt's lymphoma, but also non-neoplastic disorders.[60] The most reported of these alterations is the *EEF1DP3/FRY* fusion originating from the read-through transcription between *EEF1DP3* and *FRY* gene. *EEF1DP3/FRY* is a recurrent read-through fusion transcript that was first detected *in vitro* in *KPL4* breast carcinoma cell-line[61] and then was also detected *in vivo* in breast cancer samples but cannot be detected in breast normal tissues counterparts or blood samples from *EEF1DP3/FRY* positive patients.[42] It has been detected in some types of non-neoplastic disorders and in some cancers such as malignant melanoma, Burkitt's lymphoma, lung cancer and breast cancer.[41,42,61]

EEF1DP3 is abnormally expressed in a very large list of cancers and diseases.[60] It has been reported to be highly expressed in adrenal carcinomas i.e. adrenocortical carcinoma (ACC) and pheochromocytoma and paraganglioma (PCPG), brain lower-grade glioma (LGG), rectum adenocarcinoma (READ), gynaecological cancers such as cervical squamous cell carcinoma, endocervical adenocarcinoma (CESC) and uterine carcinosarcoma (UCS), head and neck squamous cell carcinoma (HNSC), hepatocellular carcinoma samples (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), mesothelioma, acute myeloid leukemia (AML) and lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), skin cutaneous melanoma (SKCM), pancreatic adenocarcinoma (PAAD), sarcoma (SARC) and urinary tract cancers such as urothelial bladder carcinoma (BLCA), chromophobe renal cell carcinoma (KICH), kidney renal clear cell carcinoma (KIRC) and kidney renal papillary cell carcinoma (KIRP).[37] Furthermore, it is also overexpressed in prostate adenocarcinoma (PRAD)[37,38] while its loss by deletion has been associated with an increased risk and predisposition for ankylosing spondylitis (AS).[43–45] It was also reported in datasets on various neurodegenerative disorders such as synucleinopathy and Parkinson's disease, but also in non-small cell lung cancer, multiple sclerosis and epidermolysis bullosa simplex. However, its role in these diseases is unclear.

The *EEF1D* pseudogene 4, alias *EEF1DP4*, was first described in 1998.[46] It was reported in datasets on glioma, breast cancer, primary myelofibrosis, colon cancer and osteosarcoma. However, its significance in these diseases is still unknown. A similar situation is also for the *EEF1D* pseudogene 5, alias *EEF1DP5*, which is not clearly reported in breast cancer[47] and in a gene cluster analysis in the human bone osteosarcoma epithelial cell line (U2OS).[29] Furthermore, this pseudogene exhibits frequent genomic deletions whose role is completely unknown.

EEF1D pseudogene 6 (*EEF1DP6*) is reported in datasets on acute myeloid leukemia,[48] systemic juvenile idiopathic arthritis[49] and neuropathy in Charcot-Marie-Tooth disease type 1A[50] but like other pseudogenes, its significance is unknown. The last ones foreseen by the analysis of the genome are *EEF1D* pseudogene 7 (*EEF1DP7*) and *EEEF1D* pseudogene 8 (*EEF1DP8*). However, they are not yet supported by any experimental evidence.

Pseudogenes of *EEF1G*

EEF1G, alias EEF1γ or EEF1Bγ, is a coding gene located on Chromosome 11 (11q12.3). At least five alternative splicing variants have been observed, of which two are protein-coding while the others are ncRNA sequences. Like the other components of the eEF1H complex, it is involved in the elongation step of translation and most likely stimulates the activity of eEF1B2 and guarantees stability to the entire eEF1H complex.[22,62] Analysis of the human genome revealed the presence of nine pseudogenes for *EEF1G* classified as processed pseudogenes.[63] These pseudogenes are studied very marginally.

The *EEF1G* pseudogene 1, alias *EEF1GP1*, was first reported in 1998[46] and later in a gene cluster analysis dataset on human bone osteosarcoma epithelial cell line (*U2OS*)[29] but its involvement is unclear.

EEF1G pseudogene 4, alias *EEF1GP4*, was reported by some studies on the sequencing of the human genome.[51–53] *EEF1G* pseudogene 5, alias *EEF1GP5*, is not clearly reported with regard to prostate cancer and Duchenne muscular dystrophy thus its significance in these diseases is unknown. It is also reported in a gene cluster analysis dataset on human bone osteosarcoma epithelial cell line (U2OS).[29] Similar considerations can be made for *EEF1G* pseudogene 9, alias *LOC729998*, which appears in datasets on some cancer cell lines of hepatocellular carcinoma, acute myeloid leukaemia, and diffuse large B-cell lymphoma.

The others, i.e. the *EEF1G* pseudogene 2 (alias *EEF1GP2*), the *EEF1G* pseudogene 3 (alias *EEF1GP3*), the *EEF1G* pseudogene 6 (alias *EEF1GP6*), the *EEF1G* pseudogene 7 (alias *EEF1GP7*) and the *EEF1G* pseudogene 8 (alias *EEF1GP8*), are predicted by genome sequence analysis but are not yet supported by any experimental evidence.

Pseudogenes of *EEF1E1*

EEF1E1 is known under some names such as aminoacyl tRNA synthetase complex-interacting multifunctional protein 3 (AIMP3) and P18 and was first identified by Mao and colleagues in 1998.[64] eEF1E1 plays a role as an auxiliary component of the macromolecular aminoacyl-tRNA synthetases complex (MARS) in the elongation step of translation, in particular, it interacts with several aminoacyl-tRNA synthetases[20] and could contribute to the anchoring of the MARS complex to the EF1H complex.[21,65] Its expression is frequently found altered in human cancer cells[66,67] and is considered a putative tumor suppressor gene.[68,69]

Sequence analysis of the human genome revealed the presence of only one pseudogene related to *EEF1E1* on chromosome 2, precisely in the location 2q13. This pseudogene has been named eukaryotic translation elongation factor 1 epsilon 1 pseudogene 1, alias *EEF1E1P1*, and is classified as a processed pseudogene. It shows 93.47% identity with the alternative splicing transcript variant 1 mRNA of EEF1E1 (*RefSeq* NM_004280.5) but no sequence identity or homology was found with the transcript variant 2 mRNA of *EEF1E1*, so it can be assumed that the origin of *EEF1E1P1* is due to a probable retrotransposition event from the *EEF1E1* variant 1 mRNA alone. It is reported in a study on genetic loci related to coronary artery disease[54] but its significance in this disease is unclear. Until now, no one has studied this pseudogene on cancers.

Pseudogenes of alpha eukaryotic translation elongation factors

Alpha EEFs collect the remaining components of the eEF1H complex, i.e. eEF1A1 and its isoform eEF1A2. These genes are found in different locations in the human genome and encode at least one protein that plays a central role in peptide elongation during protein biosynthesis, like the other members of eEF1H. In particular, eEF1A allows the delivery of aminoacyl-tRNAs to the ribosome mediated by the hydrolysis of GTP. Indeed, during the translation elongation step, the inactive GDP-bound form of eEF1A (eEF1A-GDP) is converted to its active GTP-bound form (eEF1A-GTP) by eEF1BGD complex by GTP hydrolysis, thus acting as a guanine nucleotide exchange factor (GEF), regenerating eEF1A-GTP for the successive elongation cycle. Both eEF1A1 and eEF1A2 exhibit canonical functions and multiple non-canonical roles (moonlight roles) within the cell and, like other EEFs, are often altered in expression, gene amplification and genomic rearrangements in many types of cancers and other diseases.

The pseudogenes reported for alpha EEFs, in particular *EEF1A1*, are very numerous[70] and are mostly considered retropseudogenes.[71] They are classified into processed pseudogenes, unprocessed pseudogenes and transcribed unprocessed pseudogenes. These pseudogenes are listed in Table 1B (more detail in the T1ASuppl supplementary material). Below are described in detail, one by one, the pseudogenes of the alpha EEFs (see also Fig. 2).

Pseudogenes of *EEF1A1*

EEF1A1 is a coding gene of 5283 nt long located on Chromosome 6 (6q13) with several alternative splicing transcript variants and protein isoforms of which most studied are the prostate tumor-inducing gene-1, alias *PTI-1* or EEF 1-alpha 1-like 14 (*EEF1A1L14*),[113] and cervical cancer suppressor 3 (*CCS-3*).[114] Today it is one of the most studied proteins both

Table 1B Pseudogenes of alpha EEFs. List of all alpha EEFs pseudogenes so far discovered and the correlation with diseases where they are reported or there is evidence about them (see also supplementary table TA1SUPPL).

RFG	PS	Description	Status	CHR	Location	Length (nt)	Main diseases
EEF1A1	*EEF1A1P1* [72]	*EEF1A1* pseudogene 1	Processed pseudogene	21	21q21.2	1034	Celiac disease (?), oral squamous cell carcinoma, osteosarcoma (?)[73–75]
	EEF1A1P2	*EEF1A1* pseudogene 2		14	14q31.1	1912	Bladder cancer (?), Uterine cancer (?), Colorectal cancer (?)
	EEF1A1P3 [70,76]	*EEF1A1* pseudogene 3		13	13q12.2	1623	–
	EEF1A1P4 [70]	*EEF1A1* pseudogene 4		12	12p12.3	1659	–
	EEF1A1P5 [70,77–81]	*EEF1A1* pseudogene 5		9	9q34.13	1747	Hepatocellular carcinoma (?), nasopharyngeal carcinoma (?), oral squamous cell carcinoma, hepatitis E virus cofactor [31,74,82,83]
	EEF1A1P6 [70,78]	*EEF1A1* pseudogene 6		7	7p15.3	1746	Rectum cancer (?), schizophrenia (?), multiple myeloma (?), hepatocellular carcinoma (?)[31,84,85]
	EEF1A1P7 [70]	*EEF1A1* pseudogene 7		19	19q13.12	2142	Hepatocellular carcinoma (?), breast cancer (?)[31,86,87]
	EEF1A1P8 [70]	*EEF1A1* pseudogene 8		3	3q27.1	1644	–
	EEF1A1P9 [70]	*EEF1A1* pseudogene 9		4	4q24	1751	Duchenne muscular dystrophy (DMD) (?), acute lymphoblastic leukemia (?), metastatic prostate cancer (?), prostate adenocarcinoma cell line (LNCaP) (?), melanoma (?), kidney cancer (?), osteosarcoma (?), hepatocellular carcinoma (?), glioma, cervical cancer (?), autism spectrum disorders (?)[31,75,88–92]
	EEF1A1P10 [70]	*EEF1A1* pseudogene 10		7	7q35	1650	–
	EEF1A1P11 [70,93]	*EEF1A1* pseudogene 11		1	1p21.3	1748	Osteosarcoma (?), lung cancer (?), colon cancer, type 2 diabetes mellitus [75,94,95]
	EEF1A1P12 [70]	*EEF1A1* pseudogene 12		2	2q12.2	1698	Hepatocellular carcinoma (?), osteosarcoma (?), multiple myeloma (?), oral squamous cell carcinoma, epilepsy (?)[31,74,75,85,96]
	EEF1A1P13 [70,97,98]	*EEF1A1* pseudogene 13		5	5p15.2	1747	–
	EEF1A1P14 [70,99]	*EEF1A1* pseudogene 14		1	1q31.3	1666	Liver cancer (?), rectum cancer (?), ovarian cancer (?), oral squamous cell carcinoma, breast cancer (?) [74,100]
	EEF1A1P15 [70]	*EEF1A1* pseudogene 15		X	Xq21.33	1689	–
	EEF1A1P16	*EEF1A1* pseudogene 16		12	12p12.3	1635	Gastric cancer, Glioma (?)[101,102]
	EEF1A1P17	*EEF1A1* pseudogene 17		12	12q12	1413	–
	EEF1A1P18	*EEF1A1* pseudogene 18		11	11q13.1	467	–

(*continued on next page*)

Table 1B (*continued*)

RFG	PS	Description	Status	CHR	Location	Length (nt)	Main diseases
	EEF1A1P19 [103]	*EEF1A1* pseudogene 19		5	5p12	1645	Hepatocellular carcinoma (?)[31]
	EEF1A1P20	*EEF1A1* pseudogene 20		5	5q21.1	1644	Nonalcoholic fatty liver disease (?) [104]
	EEF1A1P21	*EEF1A1* pseudogene 21		4	4p15.1	1339	Oral squamous cell carcinoma[74]
	EEF1A1P22 [99]	*EEF1A1* pseudogene 22		15	15q21.3	1639	Multiple myeloma (?)[85]
	EEF1A1P23	*EEF1A1* pseudogene 23	Transcribed processed pseudogene	3	3q29	658	—
	EEF1A1P24	*EEF1A1* pseudogene 24	Processed pseudogene	3	3p22.1	1638	Acute lymphoblastic leukemia (?)
	EEF1A1P25	*EEF1A1* pseudogene 25		3	3q22.3	1471	—
	EEF1A1P26	*EEF1A1* pseudogene 26		7	7p21.2	1383	Oral squamous cell carcinoma, type 2 diabetes mellitus (?)[74,105]
	EEF1A1P27	*EEF1A1* pseudogene 27		7	7p21.1	1151	Oral squamous cell carcinoma[74]
	EEF1A1P28	*EEF1A1* pseudogene 28		7	7q21.13	1671	EBV-positive T/NK-cell lymphoma (?)[106]
	EEF1A1P29	*EEF1A1* pseudogene 29		X	Xq21.2	1443	Breast cancer (?), lung cancer (?), prostate cancer (?), colorectal cancer (?), leukemia (?)[87,107],
	EEF1A1P30	*EEF1A1* pseudogene 30		X	Xq24	2354	—
	EEF1A1P31 [108]	*EEF1A1* pseudogene 31	Unprocessed pseudogene	X	Xq28	10,389	—
	EEF1A1P32	*EEF1A1* pseudogene 32		1	1q31.3	2156	Oral squamous cell carcinoma[74]
	EEF1A1P33	*EEF1A1* pseudogene 33	Processed pseudogene	12	12q23.1	1662	—
	EEF1A1P34	*EEF1A1* pseudogene 34		20	20p11.23	1464	—
	EEF1A1P35	*EEF1A1* pseudogene 35		4	4q28.3	1646	—
	EEF1A1P36	*EEF1A1* pseudogene 36		6	6q23.2	1431	—
	EEF1A1P37	*EEF1A1* pseudogene 37		8	8q23.3	557	Oral squamous cell carcinoma[74]
EEF1A1	*EEF1A1P38*	*EEF1A1* pseudogene 38	Processed pseudogene	16	16p12.1	1937	Gastric cancer, oral squamous cell carcinoma[74,101]
	EEF1A1P39	*EEF1A1* pseudogene 39		10	10p11.23	250	Oral squamous cell carcinoma[74]
	EEF1A1P40	*EEF1A1* pseudogene 40		X	Xq22.3	573	—
	EEF1A1P41	*EEF1A1* pseudogene 41		Y	Yp11.2	374	—
	EEF1A1P43 [46]	*EEF1A1* pseudogene 43	Unprocessed pseudogene	17	17p11.2	3871	Smith-Magenis syndrome (?)[109]
EEF1A2	*EEF1A1P42*	*EEF1A1* pseudogene 42	Processed pseudogene	6	6p12.3	2214	Hepatocellular carcinoma cell line (Huh-7) (?), Diffuse large B-cell lymphoma cell lines (DHL4, DHL6) (?), Acute myeloid leukemia cell lines (HL-60, MOLM-14, THP-1, U937) (?)
	LOC401677 [52,53,110]	*EEF1A2* pseudogene	Transcribed unprocessed	11	11p14.1	931	Melanoma cell line (FEMX-I) (?)

Table 1B (*continued*)

RFG	PS	Description	Status	CHR	Location	Length (nt)	Main diseases
	LOC441880	*EEF1A2* pseudogene	pseudogene	1	1p35.2	1383	–
	LOC642791 [99,111]	*EEF1A2* pseudogene		11	11q14.3	2001	–
	LOC729856 (112)	Elongation factor 1-alpha-like	Unprocessed pseudogene	1	1p36.11	468	–
	LOC100421798	*EEF1A2* pseudogene	Processed pseudogene	5	5q31.1	1148	–
	LOC100421817	*EEF1A2* pseudogene		3	3q25.1	1344	–
	LOC100421840	*EEF1A2* pseudogene		1	1q32.1	1329	–
	LOC100421842	*EEF1A2* pseudogene		1	1q42.13	554	–

Abbreviations: RFG, related functional gene; PS, pseudogene; CHR, Chromosome; [(?)], uncertain; [-], unknown.

for its fundamental role in the cell and for its involvement in many human diseases, especially cancer.[23] In fact, it plays a key role in the elongation step of translation in which it is responsible for the enzymatic delivery of aminoacyl tRNAs to the ribosome. Furthermore, it is expressed in all tissues except the brain, heart and skeletal muscles where it is replaced by the isoform *EEF1A2*.[115]

Analysis of the human genome has revealed a high number of pseudogenes for *EEF1A1*, currently 42, dispersed throughout the human genome in almost all chromosomes (except Chromosome 18 and 22) and sometimes even present several times on the same chromosome (for example Chromosome 3 and X). They are still little studied; however, some have been linked to cancer and other human diseases.

The *EEF1A1* pseudogene 1, alias *EEF1A1P1*, was first reported in 2000[72] and more recently in a study on celiac disease (CeD) where it is found strongly upregulated in the first-degree relatives of patients with CeD.[73] The significance of this discovery is unclear so the role of *EEF1AP1* in CeD needs to be better explained. Furthermore, it is also reported in a study on osteosarcoma[75] and oral squamous cell carcinoma (with copy loss).[74] The *EEF1A1* pseudogene 2 (*EEF1A1P2*)[70,76] is reported in one dataset to be significantly downregulated in bladder and uterine cancer while it is upregulated in colorectal cancer. No other studies have been conducted on it.

The *EEF1A1* pseudogene 5, alias *EEF1A1P5*, was first mentioned in 1996[70] and subsequently in other studies.[79–81] Interestingly, it is overexpressed in the fetal trabecular meshwork and fetal cornea tissues but not in the respective tissues in adults.[77] The role of *EEF1A1P5* in these tissues is unknown. It is also reported in hepatocellular carcinoma,[31] oral squamous cell carcinoma (with copy gain),[74] and as a co-interacting protein with Hepatitis E virus (*HEV*) viral proteins, in particular with HEV-Macro domain.[83] Furthermore, it interacts with the long non-coding RNA LINC01133, which is downregulated in nasopharyngeal carcinoma (NPC).[82] The significance of this interaction, however, is unknown.

The *EEF1A1* pseudogene 6, alias *EEF1A1P6*, was first reported in 1996[70] and is highly expressed in human primary monocytes.[78] Subsequently, it is reported in a study on schizophrenia,[84] hepatocellular carcinoma,[31] multiple myeloma,[85] and in a dataset on rectum cancer. The contribution of *EEF1A1P6* in these diseases is unknown.

The *EEF1A1* pseudogene 7 (*EEF1A1P7*), the EEF1A1 pseudogene 9 (*EEF1A1P9*) and the EEF1A1 pseudogene 12 (*EEF1A1P12*) were first reported in 1996[70] and subsequently in a study on hepatocellular carcinoma.[31] Furthermore, *EEF1A1P7* was also reported in two studies on breast cancer both as such[86] and as aberrant transcript fused with *EEF1A1P29*.[87]

EEF1A1P9 is reported in familiar melanoma (FM) where it was found downregulated after UV-exposure of FM cultured fibroblasts.[88] It is also reported in kidney cancer,[89] osteosarcoma[75] and autism spectrum disorders, in which copy gain of a genomic region that includes *EEF1A1P9*[92] is shown. In glioma it has been reported that it is a protective factor: in fact, patients with a high expression for *EEF1A1P9* had a favorable prognosis so it can play an important role in the onset and progression of glioma.[90] *EEF1A1P9* is shown to be downregulated in cervical cancer[91] and is reported in datasets on Duchenne muscular dystrophy (DMD), acute lymphoblastic leukaemia, metastatic prostate cancer and in the LNCaP prostate adenocarcinoma cell line. *EEF1A1P12* is reported in osteosarcoma,[75] multiple myeloma,[85] oral squamous cell carcinoma (with copy gain)[74] and epilepsy.[96]

The *EEF1A1* pseudogene 11, alias *EEF1A1P11*, was first reported in 1996[70] and subsequently by other authors.[93] It has been shown to be upregulated in colon cancer compared to normal tissue while it is downregulated in lung cancer.[94] Furthermore, it is reported in datasets on osteosarcoma[75] and type 2 diabetes mellitus.[95]

The *EEF1A1* pseudogene 13, alias *EEF1A1P13* was reported in various studies starting in 1996[70,97,98] but very little is known about it except that it is downregulated in patients with Chikungunya virus infection.[116]

The *EEF1A1* pseudogene 14 (*EEF1A1P14*), was first reported in 1996[70] and later by other authors.[99] It is reported on breast cancer,[100] oral squamous cell carcinoma (with copy gain)[74] and also appears in datasets on liver cancer,

Table 2 Pseudogenes and human diseases. List of human diseases in which the EEFs pseudogenes are suspected to be involved or there are evidences about their implication. Cancers are grouped according to the International Classification of Diseases for Oncology (ICD-O-3) and TCGA abbreviations are reported in brackets while the other human diseases are grouped according to International Statistical Classification of Diseases and Related Health Problems (ICD-11). For references see the supplementary table TA1SUPPL (where reference is missing the data are extracted mainly from GEO Profiles/NCBI and Open Targets Platform but also from other datasets; see paragraph "Material and methods").

Disease list				Pseudogenes
Cancer (included tissues and cell cultures)	Solid tumors	Lip, oral cavity and pharynx	Nasopharyngeal carcinoma	*EEF1A1P5*
			Oral squamous cell carcinoma	*EEF1A1P1, EEF1A1P5, EEF1A1P12, EEF1A1P14, EEF1A1P21, EEF1A1P26, EEF1A1P27, EEF1A1P32, EEF1A1P37, EEF1A1P38, EEF1A1P39*
		Digestive organs	Gastric cancer/stomach adenocarcinoma (STAD)	*EEF1A1P16, EEF1A1P38*
			Rectum adenocarcinoma (READ)	*EEF1DP3, EEF1A1P6, EEF1A1P14*
			Colon adenocarcinoma (COAD)	*EEF1DP4, EEF1A1P2, EEF1A1P11, EEF1A1P29*
			Liver hepatocellular carcinoma (LIHC)	*EEF1B2P2, EEF1DP3, LOC729998, EEF1A1P5, EEF1A1P6, EEF1A1P7, EEF1A1P9, EEF1A1P12, EEF1A1P14, EEF1A1P19, EEF1A1P42*
			Pancreatic adenocarcinoma (PAAD)	*EEF1DP3*
		Respiratory system and intrathoracic organs	Lung adenocarcinoma (LUAD)	*EEF1DP3, EEF1A1P11, EEF1A1P29*
			Lung squamous cell carcinoma (LUSC)	*EEF1DP3*
			Mesothelioma (MESO)	*EEF1DP3*
			Non-squamous non-small cell lung cancer (NSCLC)	*EEF1B2P1, EEF1DP3*
		Skin	Skin cutaneous melanoma (SKCM)	*EEF1DP1, EEF1DP2, EEF1DP3, EEF1A1P9, LOC401677*
		Bones, joints and articular cartilage	Bone osteosarcoma	*EEF1B2P2, EEF1DP1, EEF1DP4, EEF1DP5, EEF1GP1, EEF1GP5, EEF1A1P1, EEF1A1P9, EEF1A1P11, EEF1A1P12*
		Connective, subcutaneous and other soft tissues	Sarcoma (SARC)	*EEF1DP3*
		Eye, brain and other parts of central nervous system	Glioma	*EEF1DP3, EEF1DP4, EEF1A1P9, EEF1A1P16*
		Peripheral nerves and autonomic nervous system	Primary myelofibrosis	*EEF1DP4*
		Breast	Breast carcinoma (BRCA)	*EEF1DP3, EEF1DP4, EEF1DP5, EEF1A1P7, EEF1A1P14, EEF1A1P29*
		Female genital organs	Ovarian cancer	*EEF1A1P14*
			Uterine cancer/carcinosarcoma (UCS)/	*EEF1A1P2, EEF1DP3*

			Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC)	*EEF1DP3, EEF1A1P9*
		Male genital organs	Prostate carcinoma (PRAD)	*EEF1DP3, EEF1GP5, EEF1A1P9, EEF1A1P29*
		Urinary tract	Adrenocortical carcinoma (ACC)	*EEF1DP3*
			Bladder cancer/urothelial carcinoma (BLCA)	*EEF1A1P2, EEF1DP3*
			Chromophobe renal cell carcinoma (KICH)	*EEF1DP3*
			Kidney renal clear cell carcinoma (KIRC)	*EEF1DP3, EEF1A1P9*
			Kidney renal papillary cell carcinoma (KIRP)	*EEF1DP3*
		Thyroid and other endocrine glands	Pheochromocytoma and Paraganglioma (PCPG)	*EEF1DP3*
		Other and ill-defined sites	Head and neck squamous cell carcinoma (HNSC)	*EEF1DP3*
	Hematological malignancies		Lymphoid neoplasm diffuse large B-cell lymphoma (DLBC)	*EEF1DP1, EEF1DP3, LOC729998, EEF1A1P42*
			Acute lymphoblastic leukemia	*EEF1A1P9, EEF1A1P24,*
			Leukemia	*EEF1A1P29*
			Multiple myeloma	*EEF1A1P6, EEF1A1P12, EEF1A1P22*
			Acute myeloid leukemia (LAML)	*EEF1B2P2, EEF1DP1, EEF1DP3, EEF1DP6, LOC729998, EEF1A1P42*
			EBV-positive T/NK-cell lymphoma	*EEF1A1P28*
Other human diseases	Infectious Agents		HIV-1 reverse transcription cofactor	*EEF1B2P2*
			Hepatitis E virus cofactor	*EEF1A1P5*
	Mental, behavioural or neurodevelopmental disorders		Autism spectrum disorders	*EEF1A1P9*
			Schizophrenia	*EEF1A1P6*
	Developmental anomalies		Neuropathy in Charcot-Marie-Tooth disease type 1A	*EEF1DP6*
			Smith-Magenis syndrome	*EEF1A1P43*
	Diseases of the musculoskeletal system or connective tissue		Ankylosing spondylitis	*EEF1DP3*
			Systemic juvenile idiopathic arthritis	*EEF1DP6*
	Diseases of the nervous system		Synucleinopathy and Parkinson's disease	*EEF1DP3*
			Multiple sclerosis	*EEF1DP3*
			Duchenne muscular dystrophy	*EEF1GP5, EEF1A1P9*
			Epilepsy	*EEF1A1P12*
	Diseases of the skin		Epidermolysis bullosa simplex	*EEF1DP3*
	Diseases of the digestive system		Celiac disease	*EEF1A1P1*
			Nonalcoholic fatty liver disease	*EEF1A1P20*
	Endocrine, nutritional or metabolic diseases		Type 2 diabetes mellitus	*EEF1A1P11, EEF1A1P26*
	Diseases of the circulatory system		Coronary artery disease	*EEF1E1P1*

ovarian cancer and rectal cancer, but no other studies have been conducted.

EEF1A1 pseudogene 16 (*EEF1A1P16*) and *EEF1A1* pseudogene 38 (*EEF1A1P38*) are both upregulated in gastric cancer patient samples.[101] Furthermore, *EEF1A1P16* is also reported on glioma[102] while *EEF1A1P38* is reported on oral squamous cell carcinoma (with copy gain).[74]

EEF1A1 pseudogene 19 (*EEF1A1P19*) is reported in hepatocellular carcinoma[31,103] while *EEF1A1* pseudogene 20, alias *EEF1A1P20*, is reported in a study concerning single nucleotide polymorphisms associated with the pathology of non-alcoholic fatty liver disease.[104]

EEF1A1 pseudogene 21 (*EEF1A1P21*) is reported in oral squamous cell carcinoma (with copy gain).[74] *EEF1A1* pseudogene 22 (*EEF1A1P22*) is reported in multiple myeloma[85,99] while *EEF1A1* pseudogene 24, alias *EEF1A1P24*, is reported in datasets on acute lymphoblastic leukaemia.

EEF1A1 pseudogene 26, alias *EEF1A1P26*, is reported in type 2 diabetes mellitus[105] and oral squamous cell carcinoma (with copy gain)[74] while *EEF1A1* pseudogene 27 (*EEF1A1P27*) is reported in oral squamous cell carcinoma (with copy gain).[74]

EEF1A1 pseudogene 28 (*EEF1A1P2*) shows copy number gain in EBV-positive T/NK-cell lymphoma[106] and *EEF1A1* pseudogene 29 (*EEF1A1P29*) is reported in breast cancer,[87,107] lung cancer, prostate cancer, colorectal cancer and leukaemia.[107]

EEF1A1 pseudogene 31 (*EEF1A1P31*) was first reported in 2018[108] while *EEF1A1* pseudogene 32 (*EEF1A1P32*) is reported in oral squamous cell carcinoma (with copy gain).[74]

EEF1A1 pseudogene 37 (*EEF1A1P37*) and EEF1A1 pseudogene 39 (*EEF1A1P39*) are reported together on oral squamous cell carcinoma with copy loss for the former and a copy gain for the latter.[74]

EEF1A1 pseudogene 43, alias *EEF1A1P43* or formerly known as *EEF1A3*, was first reported in 1998[46] and later in Smith-Magenis syndrome where, however, it is not considered important because it does not show significant physiological effects.[109]

The *EEF1A1* pseudogene 4 (*EEF1A1P4*), the *EEF1A1* pseudogene 8 (*EEF1A1P8*), the *EEF1A1* pseudogene 10 (*EEF1A1P10*) and the *EEF1A1* pseudogene 15 (*EEF1A1P15*) are first described in 1996[70] but no other studies have been done. Lastly, the remaining ones, i.e. *EEF1A1* pseudogene 2 (alias *EEF1A1P2*), *EEF1A1* pseudogene 17 (*EEF1A1P17*), *EEF1A1* pseudogene 18 (*EEF1A1P18*), *EEF1A1* pseudogene 23 (*EEF1A1P23*), *EEF1A1* pseudogene 25 (*EEF1A1P25*), *EEF1A1* pseudogene 30 (*EEF1A1P30*), *EEF1A1* pseudogene 33 (*EEF1A1P33*), *EEF1A1* pseudogene 34 (*EEF1A1P34*), *EEF1A1* pseudogene 35 (*EEF1A1P35*), *EEF1A1* pseudogene 36 (*EEF1A1P36*), *EEF1A1* pseudogene 40 (*EEF1A1P40*) and *EEF1A1* pseudogene 41 (*EEF1A1P41*), are predicted by genome sequence analysis but are not yet supported by experimental evidence, so they are very little known.

Pseudogenes of *EEF1A2*

EEF1A2 is a coding gene located on Chromosome 20 (20q13.33) and at the same time it is an isoform of *EEF1A1* that performs the same function in the translation elongation step. The switch between the two isoforms occurs only in the brain, heart and skeletal muscle. *EEF1A2* shows expression alterations and various genomic anomalies in many cancers.[23,117] Analysis of the human genome revealed nine poorly studied pseudogenes for *EEF1A2* listed below.

EEF1A1 pseudogene 42, alias *EEF1A1P42*, is associated with *EEF1A2* pseudogenes instead with *EEF1A1* pseudogenes and is reported in datasets on some cancer cell lines of hepatocellular carcinoma, acute myeloid leukaemia and diffuse large B-cell lymphoma without any other type of study.

LOC401677 is reported in some papers[52,53,110] and in datasets on melanoma, in particular in the *FEMX-I* melanoma cell line.

LOC642791 and *LOC729856* are reported in some studies[99,111,112] but not much more is known about them.

The other *EEF1A2* pseudogenes, namely *LOC441880*, *LOC100421798*, *LOC100421817*, *LOC100421840* and *LOC100421842*, are predicted by genome sequence analysis but are not yet supported by experimental evidence, so they are unknown.

Conclusion and perspective

All the coding genes belonging to EEFs play an important role in the cell and undergo important alterations in cancer. Similarly, even if still in its infancy, the studies available so far on the respective pseudogenes highlight at least two important aspects: first, they certainly have one or more roles in the cell, most likely via ncRNAs, but the possibility of other forms of regulation is not excluded, including through proteins or peptides still unknown, and second that they certainly have a role in human pathologies, first of all in cancer.

EEFs pseudogenes discovered to date are very numerous, especially for *EEF1A1*, and this could not only be a simple result of chance, a consequence of errors or evolution, but could reflect a complex system of genomic regulation that is still poorly understood today. *EEF1A1*, for example, is very conserved in the evolution of the species, so much so that its counterpart is also known in bacteria with the name of EF-Tu. Therefore, it is the oldest gene in the EEFs and has certainly been the subject of many events during the evolution of the species. However, it may be equally true that the abundance of its pseudogenes in the human genome is not only entirely linked to evolution but could also be related to other factors, including the high transcription of its parental gene. Indeed, in some cases there is a positive correlation between high levels of gene expression, especially for housekeeping genes, and the increase in the number of related pseudogenes in the human genome.[118] This is true, apart for *EEF1A1*, also for *GAPDH* and *RPL21* (for more details see supplementary table TA1SUPPL), both of which are highly transcribed.

The other members of the EEFs, among them, have a similar number of pseudogenes and this is less than ten except for *EEF1E1* which has only one pseudogene. On an evolutionary level, the latter could certainly be the most recent, but it is also significant that its parental gene is considered a putative tumor suppressor gene[68] that is often downregulated in cancer.[69] In fact, *EEF1E1* also has the least number of genomic rearrangements.

It is currently not known whether the pseudogenes of EEFs have a regulatory role in the expression of the

respective parental gene as described for others[119] and for many there is still no evidence of their involvement in the development and/or progression of human cancers or other human diseases because there is no sufficient knowledge about them to understand their repercussions on cellular behavior. Furthermore, EEFs pseudogenes could theoretically produce non-coding transcripts, but there is currently no firm evidence for this.

The studies in which EEFs pseudogenes have most appeared concern oral squamous cell carcinoma, hepatocellular carcinoma, osteosarcoma, breast cancer and acute myeloid leukaemia (Table 2). However, their exact role in these cancers is not yet defined while the most studied pseudogenes are *EEF1DP3* and *EEF1A1P9*, although they must be well characterized and understood. More work is needed for all these pseudogenes, especially for those to date less known, to achieve two very important goals, in addition to general knowledge about them, which are their role as possible biomarkers, both diagnostic and prognostic, as determined for others,[3] and their possible role as therapeutic targets.

In conclusion, EEFs pseudogenes may play a role in the cell, probably in gene regulation, and are involved in many human diseases, including cancer. In the future, it will be important to characterize them and explore their ability to modulate parental gene expression under different cellular conditions, their precise mechanisms of function and the possibility of using them as new biomarkers or therapeutic targets for cancer management and treatment or other human diseases.

Conflict of interests

The author has no conflict of interests to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gendis.2021.03.009.

References

1. Cristiano L. Translation elongation factors: are useful biomarkers in cancer? *Open Access J Biog Sci Res*. 2020;6(1):1—7.
2. Chan JJ, Tay Y. Noncoding RNA:RNA regulatory networks in cancer. *Int J Mol Sci*. 2018;19(5):1310.
3. Poliseno L, Marranci A, Pandolfi PP. Pseudogenes in human cancer. *Front Med*. 2015:2:68.
4. Kim MS, Pinto SM, Getnet D, et al. A draft map of the human proteome. *Nature*. 2014;509(7502):575—581.
5. Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature*. 2012;489(7414):101—108.
6. Poliseno L. Pseudogenes: newly discovered players in human cancer. *Sci Signal*. 2012;5(242):re5.
7. Chan WL, Chang JG. Pseudogene-derived endogenous siRNAs and their function. *Methods Mol Biol*. 2014;1167:227—239.
8. Hu X, Yang L, Mo YY. Role of pseudogenes in tumorigenesis. *Cancers (Basel)*. 2018;10(8):256.
9. An Y, Furber KL, Ji S. Pseudogenes regulate parental gene expression via ceRNA network. *J Cell Mol Med*. 2018;21(1):185—192.
10. Xiao-Jie L, Ai-Mei G, Li-Juan J, Jiang X. Pseudogene in cancer: real functions and promising signature. *J Med Genet*. 2015;52(1):17—24.
11. Kalyana-Sundaram S, Kumar-Sinha C, Shankar S, et al. Expressed pseudogenes in the transcriptional landscape of human cancers. *Cell*. 2012;149(7):1622—1634.
12. Savtchenko ES, Schiff TA, Jiang CK, Freedberg IM, Blumenberg M. Embryonic expression of the human 40-kD keratin: evidence from a processed pseudogene sequence. *Am J Hum Genet*. 1988;43(5):630—637.
13. Cooke S, Shlien A, Marshall J, et al. Processed pseudogenes acquired somatically during cancer development. *Nat Commun*. 2014;5:3644.
14. Kovalenko TF, Patrushev LI. Pseudogenes as functionally significant elements of the genome. *Biochemistry (Mosc)*. 2018;83(11):1332—1349.
15. Grandér D, Johnsson P. Pseudogene-expressed RNAs: emerging roles in gene regulation and disease. *Curr Top Microbiol Immunol*. 2016;394:111—126.
16. Johnsson P, Morris KV, Grandér D. Pseudogenes: a novel source of trans-acting antisense RNAs. *Methods Mol Biol*. 2014;1167:213—226.
17. Pei BK, Sisu C, Frankish A, et al. The GENCODE pseudogene resource. *Genome Biol*. 2012;13(9):R51.
18. Huret JL, Ahmad M, Arsaban M, et al. Atlas of genetics and cytogenetics in oncology and haematology in 2013. *Nucleic Acids Res*. 2013;41(D1):D920—D924.
19. Sasikumar AN, Perez WB, Kinzy TG. The many roles of the eukaryotic elongation factor 1 complex. *Wiley Interdiscipl Rev RNA*. 2012;3(4):543—555.
20. Tao Y, Fang P, Kim S, Guo M, Young NL, Marshall AG. Mapping the contact surfaces in the Lamin A: AIMP3 complex by hydrogen/deuterium exchange FT-ICR mass spectrometry. *PLoS One*. 2017;12(8):e0181869.
21. Quevillon S, Mirande M. The p18 component of the multisynthetase complex shares a protein motif with the beta and gamma subunits of eukaryotic elongation factor 1. *FEBS Lett*. 1996;395(1):63—67.
22. Ejiri S. Moonlighting functions of polypeptide elongation factor 1: from actin bundling to zinc finger protein R1-associated nuclear localization. *Biosci Biotechnol Biochem*. 2002;66(1):1—21.
23. Hassan MK, Kumar D, Naik M, Dixit M. The expression profile and prognostic significance of eukaryotic translation elongation factors in different cancers. *PLoS One*. 2018;13(1):e0191377.
24. Chambers DM, Rouleau GA, Abbott CM. Comparative genomic analysis of genes encoding translation elongation factor 1B(alpha) in human and mouse shows EEF1B1 to be a recent retrotransposition event. *Genomics*. 2001;77(3):145—148.
25. von der Kammer H, Klaudiny J, Zimmer M, Scheit KH. Human elongation factor 1 beta: cDNA and derived amino acid sequence. *Biochem Biophys Res Commun*. 1991;177(1):312—317.
26. Baty F, Joerger M, Früh M, Klingbiel D, Zappa F, Brutsche M. 24h-gene variation effect of combined bevacizumab/erlotinib in advanced non-squamous non-small cell lung cancer using exon array blood profiling. *J Transl Med*. 2017;15(1):66.
27. Vanwetswinkel S, Kriek J, Andersen GR, et al. Solution structure of the 162 residue C-terminal domain of human elongation factor 1Bgamma. *J Biol Chem*. 2003;278(44):43443—43451.
28. Pizzuti A, Gennarelli M, Novelli G, et al. Human elongation factor EF-1 beta: cloning and characterization of the EF1 beta 5a gene and assignment of EF-1 beta isoforms to chromosomes 2,5,15 and X. *Biochem Biophys Res Commun*. 1993;197(1):154—162.

29. Chapman AR, Lee DF, Cai W, et al. Correlated gene modules uncovered by single-cell transcriptomics with high detectability and accuracy. *BioRxiv*. 2020:e892190.
30. Zhou J, Quah JY, Ng Y, et al. ASLAN003, a potent dihydroorotate dehydrogenase inhibitor for differentiation of acute myeloid leukemia. *Haematologica*. 2020;105(9):2286—2297.
31. Süt BB. Data article on genes that share similar expression patterns with EEF1 complex proteins in hepatocellular carcinoma. *Data Brief*. 2020;29:105162.
32. Díez-Fuertes F, De La Torre-Tarazona HE, Calonge E, et al. Transcriptome sequencing of peripheral blood mononuclear cells from elite controller-long term non progressors. *Sci Rep*. 2019;9(1):14265.
33. Burton PR, Clayton DG, Cardon LR, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145):661—678.
34. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304—1351.
35. Kimura K, Wakamatsu A, Suzuki Y, et al. Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res*. 2006;16(1):55—65.
36. Rual JF, Venkatesan K, Hao T, et al. Towards a proteome-scale map of the human protein-protein interaction network. *Nature*. 2005;437(7062):1173—1178.
37. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, et al. The cancer genome atlas pan-cancer analysis project. *Nat Genet*. 2013;45(10):1113—1120.
38. Erho N, Buerki C, Triche TJ, Davicioni E, Vergara IA. Transcriptome-wide detection of differentially expressed coding and non-coding transcripts and their clinical significance in prostate cancer. *J Oncol*. 2012:2012:541353.
39. Fimereli D, Fumagalli D, Brown D, et al. Genomic hotspots but few recurrent fusion genes in breast cancer. *Genes Chromosomes Cancer*. 2018;57(7):331—338.
40. Alaei-Mahabadi B, Bhadury J, Karlsson JW, Nilsson JA, Larsson E. Global analysis of somatic structural genomic alterations and their impact on gene expression in diverse human cancers. *Proc Natl Acad Sci U S A*. 2016;113(48):13768—13773.
41. Babiceanu M, Qin F, Xie Z, et al. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res*. 2016;44(6):2859—2872.
42. Kim J, Kim S, Ko S, et al. Recurrent fusion transcripts detected by whole-transcriptome sequencing of 120 primary breast cancer samples. *Genes Chromosomes Cancer*. 2015;54(11):681—691.
43. Shahba S, Jafari Shakib R, Jamshidi A, et al. Association study of copy number variation in BMP8A gene with the risk of ankylosing spondylitis in Iranian population. *J Cell Biochem*. 2018;120(5):8359—8365.
44. Yim SH, Jung SH, Chung B, Chung YJ. Clinical implications of copy number variations in autoimmune disorders. *Kor J Intern Med*. 2015;30(3):294—304.
45. Jung SH, Yim SH, Hu HJ, et al. Genome-wide copy number variation analysis identifies deletion variants associated with ankylosing spondylitis. *Arthritis Rheum*. 2014;66(8):2103—2112.
46. Sanger Center. Genome Sequencing Center. Toward a complete human genome sequence. *Genome Res*. 1998;8(11):1097—1108.
47. Stefansson OA, Jonasson JG, Olafsdottir K, et al. Genomic and phenotypic analysis of BRCA2 mutated breast cancers reveals co-occurring changes linked to progression. *Breast Cancer Res*. 2011;13(5):R95.
48. Lv H, Zhang M, Shang Z, et al. Genome-wide haplotype association study identify the FGFR2 gene as a risk gene for acute myeloid leukemia. *Oncotarget*. 2017;8(5):7891—7899.
49. Ombrello MJ, Arthur VL, Remmers EF, et al. Genetic architecture distinguishes systemic juvenile idiopathic arthritis from other forms of juvenile idiopathic arthritis: clinical and therapeutic implications. *Ann Rheum Dis*. 2017;76(5):906—913.
50. Tao F, Beecham GW, Rebelo AP, et al. Modifier gene candidates in charcot-marie-tooth disease type 1a: a case-only genome-wide association study. *J Neuromuscul Dis*. 2019;6(2):201—211.
51. Muzny DM, Scherer SE, Kaul R, et al. The DNA sequence, annotation and analysis of human chromosome 3. *Nature*. 2006;440(7088):1194—1198.
52. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431(7011):931—945.
53. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860—921.
54. van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ Res*. 2018;122(3):433—443.
55. Cristiano L. EEF1B2 (eukaryotic translation elongation factor 1 beta 2). *Genet Cytogenet Oncol Haematol*. 2020;24(9):338—345.
56. Le Sourd F, Boulben S, Le Bouffant R, et al. eEF1B: at the dawn of the 21st century. *Biochim Biophys Acta*. 2006;1759(1—2):13—31.
57. Browne GJ, Proud CG. Regulation of peptide-chain elongation in mammalian cells. *Eur J Biochem*. 2002;269(22):5360—5368.
58. Cristiano L. EEF1D (eukaryotic translation elongation factor 1 delta). *Genet Cytogenet Oncol Haematol*. 2020;24(3):117—135.
59. Dunham A, Matthews LH, Burton J, et al. The DNA sequence and analysis of human chromosome 13. *Nature*. 2004;428(6982):522—528.
60. Cristiano L. EEF1DP3 (eukaryotic translation elongation factor 1 delta pseudogene 3). *Genet Cytogenet Oncol Haematol*. 2020;24(4):164—169.
61. Kim D, Salzberg SL. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biol*. 2011;12(8), R72.
62. Mansilla F, Friis I, Jadidi M, Nielsen KM, Clark BF, Knudsen CR. Mapping the human translation elongation factor eEF1H complex using the yeast two-hybrid system. *Biochem J*. 2002;365(Pt 3):669—676.
63. Cristiano L. EEF1G (eukaryotic translation elongation factor 1 gamma). *Atlas Genet Cytogenet Oncol Haematol*. 2020;24(2):58—68.
64. Mao M, Fu G, Wu JS, et al. Identification of genes expressed in human CD34(+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning. *Proc Natl Acad Sci U S A*. 1998;95(14):8175—8180.
65. Deineko V. On ARS-interacting multifunctional protein p18. *Nat Prec*. 2008.
66. Kim SM, Jeon Y, Kim D, et al. AIMP3 depletion causes genome instability and loss of stemness in mouse embryonic stem cells. *Cell Death Dis*. 2018;9(10):972.
67. Park BJ, Oh YS, Park SY, et al. AIMP3 haploinsufficiency disrupts oncogene-induced p53 activation and genomic stability. *Cancer Res*. 2006;66(14):6913—6918.
68. Park BJ, Kang JW, Lee SW, et al. The haploinsufficient tumor suppressor p18 upregulates p53 via interactions with ATM/ATR. *Cell*. 2005;120(2):209—221.
69. Kim SS, Hur SY, Kim YR, Yoo NJ, Lee SH. Expression of AIMP1, 2 and 3, the scaffolds for the multi-tRNA synthetase complex, is downregulated in gastric and colorectal cancer. *Tumori*. 2011;97(3):380—385.
70. Lund A, Knudsen SM, Vissing H, Clark B, Tommerup N. Assignment of human elongation factor 1 alpha genes: EEF1A

maps to chromosome 6q14 and EEF1A2 to 20q13.3. *Genomics*. 1996;36(2):359—361.
71. Madsen HO, Poulsen K, Dahl O, Clark BF, Hjorth JP. Retropseudogenes constitute the major part of the human elongation factor 1 alpha gene family. *Nucleic Acids Res*. 1990; 18(6):1513—1516.
72. Hattori M, Fujiyama A, Taylor TD, et al. The DNA sequence of human chromosome 21. *Nature*. 2000;405(6784):311—319.
73. Acharya P, Kutum R, Pandey R, et al. First degree relatives of patients with celiac disease harbour an intestinal transcriptomic signature that might protect them from enterocyte damage. *Clin Transl Gastroenterol*. 2018;9(10):195.
74. Vincent-Chong VK, Salahshourifar I, Razali R, Anwar A, Zain RB. Immortalization of epithelial cells in oral carcinogenesis as revealed by genome-wide array comparative genomic hybridization: a meta-analysis. *Head Neck*. 2015; 38(Suppl 1):E783—E797.
75. Liu F, Xing L, Zhang X, Zhang X. A four-pseudogene classifier identified by machine learning serves as a novel prognostic marker for survival of osteosarcoma. *Genes (Basel)*. 2019; 10(6):414.
76. Bonaldo MF, Yu MT, Jelenc P, et al. Selection of cDNAs using chromosome-specific genomic clones: application to human chromosome 13. *Hum Mol Genet*. 1994;3(9):1663—1673.
77. Carnes MU, Allingham RR, Ashley-Koch A, Hauser MA. Transcriptome analysis of adult and fetal trabecular meshwork, cornea, and ciliary body tissues by RNA sequencing. *Exp Eye Res*. 2018;167:91—99.
78. Mirsafian H, Ripen AM, Manaharan T, Mohamad SB, Merican AF. Toward a reference gene catalog of human primary monocytes. *OMICS*. 2016;20(11):627—634.
79. Pieragostino D, Agnifili L, Fasanella V, et al. Shotgun proteomics reveals specific modulated protein patterns in tears of patients with primary open angle glaucoma naïve to therapy. *Mol Biosyst*. 2013;9(6):1108—1116.
80. de Mateo S, Castillo J, Estanyol JM, Ballescà JL, Oliva R. Proteomic characterization of the human sperm nucleus. *Proteomics*. 2011;11(13):2714—2726.
81. Buschow SI, van Balkom BW, Aalberts M, Heck AJ, Wauben M, Stoorvogel W. MHC class II-associated proteins in B-cell exosomes and potential functional implications for exosome biogenesis. *Immunol Cell Biol*. 2010;88(8):851—856.
82. Zhang W, Du M, Wang T, et al. Long non-coding RNA LINC01133 mediates nasopharyngeal carcinoma tumorigenesis by binding to YBX1. *Am J Cancer Res*. 2019;9(4):779—790.
83. Ojha NK, Lole KS. Hepatitis E virus ORF1 encoded non structural protein-host protein interaction network. *Virus Res*. 2016;213:195—204.
84. Evgrafov OV, Armoskus C, Wrobel BB, et al. Gene expression in patient-derived neural progenitors implicates WNT5A signaling in the etiology of schizophrenia. *Biol Psychiatr*. 2020;88(3):236—247.
85. Gruber F, Keats JJ, McBride K, et al. Bayesian network models of multiple myeloma: drivers of high risk and durable response. *Blood*. 2016;128(22):4406.
86. Elaine Hardman W, Primerano DA, Legenza MT, Morgan J, Fan J, Denvir J. mRNA expression data in breast cancers before and after consumption of walnut by women. *Data Brief*. 2019;25:104050.
87. Lin Y-Y, Gawronski A, Hach F, et al. Computational identification of micro-structural variations and their proteogenomic consequences in cancer. *Bioinformatics*. 2018;34(10): 1672—1681.
88. Fan M, Pfeffer SR, Lynch HT, et al. Altered transcriptome signature of phenotypically normal skin fibroblasts heterozygous for CDKN2A in familial melanoma: relevance to early intervention. *Oncotarget*. 2013;4(1):128—141.
89. Ha MJ, Baladandayuthapani V, Do KA. Prognostic gene signature identification using causal structure learning: applications in kidney cancer. *Canc Inf*. 2015;14(Suppl 1):23—35.
90. Wang Y, Liu X, Guan G, Xiao Z, Zhao W, Zhuang M. Identification of a five-pseudogene signature for predicting survival and its ceRNA network in glioma. *Front Oncol*. 2019;9:1059.
91. Wu W, Sui J, Liu T, et al. Integrated analysis of two-lncRNA signature as a potential prognostic biomarker in cervical cancer: a study based on public database. *Peer J*. 2019;7: e6761.
92. Stobbe G, Liu Y, Wu R, Hudgings LH, Thompson O, Hisama FM. Diagnostic yield of array comparative genomic hybridization in adults with autism spectrum disorders. *Genet Med*. 2014; 16(1):70—77.
93. Ran D, Daye ZJ. Gene expression variability and the analysis of large-scale RNA-seq studies with the MDSeq. *Nucleic Acids Res*. 2017;45(13):e127.
94. Cakir MV, Wirth H, Hopp L, Binder H. MicroRNA expression landscapes in stem cells, tissues, and cancer. *Methods Mol Biol*. 2014;1107:279—302.
95. Xiao D, Zhang SM, Li X, et al. IL-1B rs1143623 and EEF1A1P11-RPL7P9 rs10783050 polymorphisms affect the glucose-lowing efficacy of metformin in Chinese overweight or obese Type 2 diabetes mellitus patients. *Pharmacogenomics*. 2015; 16(14):1621—1629.
96. Wittkowski KM, Sonakya V, Song T, Seybold MP, Keddache M, Durner M. From single-SNP to wide-locus: genome-wide association studies identifying functionally related genes and intragenic regions in small sample studies. *Pharmacogenomics*. 2013;14(4):391—401.
97. Audere M, Rutka K, Inaskina I, et al. Genetic linkage studies of a North Carolina macular dystrophy family. *Medicina*. 2016; 52(3):180—186.
98. van Soest S, van Rossem MJ, Heckenlively JR, et al. Integrated genetic and physical map of the 1q31->q32.1 region, encompassing the RP12 locus, the F13B and HF1 genes, and the EEF1AL11 and RPL30 pseudogenes. *Cytogenet Cell Genet*. 1999;84(1—2):22—27.
99. Comuzzie AG, Cole SA, Laston SL, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLoS One*. 2012;7(12):e51954.
100. Rodrigues-Peres RM, de S Carvalho B, Anurag M, et al. Copy number alterations associated with clinical features in an underrepresented population with breast cancer. *Mol Genet Genom Med*. 2019;7(7):e00750.
101. Abed S, Baghaei K, Pakzad P, Hashemi M, Zali MR. Evaluation of epithelial-mesenchymal transition genes involved in iranian gastric cancer patients via transcriptome analysis. *Int J Canc Manag*. 2019;12(12):e94924.
102. Tian H, Cong P, Qi R, et al. Decreased invasion ability of hypotaurine synthesis deficient glioma cells was partially due to hypomethylation of Wnt5a promoter. *Biocell*. 2017;41: 27—32.
103. Lai KP, Li JW, Chan TF, et al. Transcriptomic and methylomic analysis reveal the toxicological effect of 2,3,7,8-Tetrachlorodibenzodioxin on human embryonic stem cell. *Chemosphere*. 2018;206:663—673.
104. Balsano C, Porcu C, Sideri S, Tavolaro S. Fat and hepatocellular carcinoma. *Hepatoma Res*. 2018;4:38.
105. Kato N. Insights into the genetic basis of type 2 diabetes. *J Diabetes Invest*. 2013;4(3):233—244.
106. Ng SB, Chung TH, Kato S, et al. Epstein-Barr virus-associated primary nodal T/NK-cell lymphoma shows a distinct molecular signature and copy number changes. *Haematol*. 2018;103(2): 278—287.
107. Tutar L, Özgür A, Tutar Y. Involvement of miRNAs and pseudogenes in cancer. *Methods Mol Biol*. 2018;1699:45—66.

108. Jorge P, Garcia E, Gonçalves A, et al. Classical fragile-X phenotype in a female infant disclosed by comprehensive genomic studies. *BMC Med Genet*. 2018;19(1):74.
109. Lucas RE, Vlangos CN, Das P, Patel PI, Elsea SH. Genomic organisation of the approximately 1.5 Mb Smith-Magenis syndrome critical interval: transcription map, genomic contig, and candidate gene analysis. *Eur J Hum Genet*. 2001;9(12):892—902.
110. Taylor TD, Noguchi H, Totoki Y, et al. Human chromosome 11 DNA sequence and analysis including novel gene identification. *Nature*. 2006;440(7083):497—500.
111. Dick DM, Aliev F, Krueger RF, et al. Genome-wide association study of conduct disorder symptomatology. *Mol Psychiatr*. 2011;16(8):800—808.
112. Takahashi K, Tatsumi N, Fukami T, Yokoi T, Nakajima M. Integrated analysis of rifampicin-induced microRNA and gene expression changes in human hepatocytes. *Drug Metabol Pharmacokinet*. 2014;29(4):333—340.
113. Mansilla F, Hansen LL, Jakobsen H, Kjeldgaard NO, Clark BF, Knudsen CR. Deconstructing PTI-1: PTI-1 is a truncated, but not mutated, form of translation elongation factor 1A1, eEF1A1. *Biochim Biophys Acta*. 2005;727(2):116—124.
114. Choi WI, Kim Y, Kim Y, et al. Eukaryotic translation initiator protein 1A isoform, CCS-3, enhances the transcriptional repression of p21CIP1 by proto-oncogene FBI-1 (Pokemon/ZBTB7A). *Cell Physiol Biochem*. 2009;23(4—6):359—370.
115. Abbott CM, Proud CG. Translation factors: in sickness and in health. *Trends Biochem Sci*. 2004;29(1):25—31.
116. Soares-Schanoski A, Baptista Cruz N, de Castro-Jorge LA, et al. Systems analysis of subjects acutely infected with the Chikungunya virus. *PLoS Pathog*. 2019;15(6):e1007880.
117. Lee M, Surh YJ. eEF1A2 as a putative oncogene. *Ann N Y Acad Sci*. 2009;1171:87—93.
118. McDonell L, Drouin G. The abundance of processed pseudogenes derived from glycolytic genes is correlated with their expression level. *Genome*. 2012;55(2):147—151.
119. Hirotsune S, Yoshida N, Chen A, et al. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature*. 2003;423(6935):91—96.