

## Review

# Differential expression of the three independent CaM genes coding for an identical protein: Potential relevance of distinct mRNA stability by different codon usage

Mads Munk<sup>a</sup>, Eduardo Villalobo<sup>b</sup>, Antonio Villalobo<sup>c,\*</sup>, Martin W. Berchtold<sup>a,\*</sup>

<sup>a</sup> Department of Biology, University of Copenhagen, 13 Universitetsparken, Copenhagen Ø DK-2100, Denmark

<sup>b</sup> Facultad de Biología, Departamento de Microbiología, Universidad de Sevilla, Avenida Reina Mercedes 6, Sevilla E-41012, Spain

<sup>c</sup> Cancer and Human Molecular Genetics Area – Oto-Neurosurgery Research Group, University Hospital La Paz Research Institute (IdiPAZ), Paseo de la Castellana 261, Madrid E-28046, Spain



## ARTICLE INFO

## Keywords:

Calmodulin genes  
Codon usage  
Epigenetic control  
miRNA control  
Mutation-mediated diseases  
Phylogeny  
Pseudogenes

## ABSTRACT

The Ca<sup>2+</sup>-sensor protein calmodulin (CaM) is a major regulator of multiple cell functions. A unique and puzzling feature of human, and all so far investigated mammals, is the presence of three distinct CaM genes on different chromosomes, which code for identical proteins. How this case of apparent genetic redundancy evolved and why it could be to the advantage of the mammalian organisms is not well established. With a main focus on humans, this article aims to review existing literature addressing how the genes nonetheless differ in function. Clearly, the three CaM genes are differentially expressed in different tissues, during development, in response to different stimuli, and other factors including environmental conditions. As shown in hippocampal neurons, different mRNAs from the CaM genes may even localize differently within the same cell. Regulation of CaM gene expression is achieved by a variety of regulatory elements present in the three genes, including different promoter/insulator elements and 3'- and 5'-noncoding regions differing in length and sequence, as well as regulation by epigenetic factors and miRNAs. Here, we hypothesize that predicted differences in mRNA stability and translational efficiency due to divergent codon usage could play an additional regulatory role as the three genes differ markedly in their use of synonymous codons. *CALM3*, predicted to produce a relatively stable mRNA may be important where the transcription level is low or transiently absent, e.g. during spermatogenesis. In contrast, *CALM2* with a predicted much shorter mRNA half-life, may provide better temporal control of CaM levels. Deciphering the underlying mechanisms responsible for all this complexity may help to understand why this unique multigenic arrangement may be an advantage for the optimal spatio-temporal expression of CaM in higher eukaryotes. Finally, we discuss the expression of the CaM genes in selected human pathologies, and how mutations in these genes are responsible for the appearance of serious congenital syndromes, mainly affecting the heart, and although less known, possibly also affecting the functionality of the central nervous system and other organs.

## Abbreviations

3'/5'-UTR, 3'- and 5'-untranslated regions;  
ALS, amyotrophic lateral sclerosis;  
AP1, activator protein 1;  
AP2, activator protein 2;  
bp, base pairs;  
CALM1,2,3/Calm1,2,3, calmodulin-1,2,3 genes;  
CaM, calmodulin;

CaML, calmodulin-like protein;  
CaMK-II, calmodulin-dependent protein kinase-II;  
Ca<sub>v</sub>1.2, L-type voltage-gated Ca<sup>2+</sup> channel 1.2;  
CCR4-NOT, carbon catabolite repression negative on TATA-less;  
cDNA, copy DNA;  
CDS, coding sequence;  
CPEB, cytoplasmic polyadenylation element binding protein;  
CPVT, catecholaminergic polymorphic ventricular tachycardia;  
CRE, cAMP-response element;

\* Corresponding authors.

E-mail addresses: [antonio.villalobo@idipaz.es](mailto:antonio.villalobo@idipaz.es) (A. Villalobo), [mabe@bio.ku.dk](mailto:mabe@bio.ku.dk) (M.W. Berchtold).

<https://doi.org/10.1016/j.ceca.2022.102656>

Received 1 July 2022; Received in revised form 1 September 2022; Accepted 25 September 2022

Available online 8 October 2022

0143-4160/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

CREB, cAMP-response element binding protein;  
 CSC, codon stabilization coefficient;  
 CTCF, CCCTC-binding protein;  
 DACT3, disheveled binding antagonist of  $\beta$ -catenin 3 gene;  
 DACT3-AS1, DACT antisense RNA 1;  
 EPCAM-DT, epithelial cell adhesion molecule divergent transcript lncRNA;  
 GAPDHP38, glyceraldehyde 3-phosphate dehydrogenase pseudogene 38;  
 GNG8/Gng8, G protein subunit  $\gamma$ 8 gene;  
 HEK, human embryonic kidney;  
 IP<sub>3</sub>, inositol-1,4,5-trisphosphate;  
 IVF, idiopathic ventricular tachycardia;  
 LINC00642, lncRNA-642;  
 LINC02317, lncRNA-2317;  
 lncRNA, long non-coding RNA;  
 LQTS, long-QT syndrome;  
 Mef2a, myocytes enhancer factor 2a;  
 Mir-3558, microRNA-3558;  
 MLCK, myosin light-chain kinase;  
 NFAT, nuclear factor of activated T cells;  
 NMDA, N-methyl-D-aspartate;  
 NRDE2/Nrde2, nuclear RNA interference defective-2 gene;  
 Pnma8a/b, paraneoplastic Ma antigen protein 8a/8b gene;  
 Pnma2, paraneoplastic Ma antigen-like protein 2 gene;  
 PPP5D1P, Ser/Thr protein phosphatase tetratricopeptide repeat domain containing 1 pseudogene;  
 PSMC1/Psmc1, 26S proteasome subunit - ATPase 1 gene;  
 PTGIR/Ptgir, prostaglandin I<sub>2</sub> receptor gene;  
 qPCR, quantitative polymerase chain reaction;  
 RPL21P11, ribosomal protein L21 pseudogene 11;  
 Rps6ka5, ribosomal protein S6 kinase a5 gene;  
 RyR2, ryanodine receptor 2;  
 siRNA, small interfering RNA;  
 SNP, single nucleotide polymorphism;  
 SP1, specificity protein 1;  
 STPG4/Stpg4, sperm-tail PG-rich repeat containing 4 gene;  
 TPM, transcripts per million;  
 TTC7A/Ttc7b, tetratricopeptide repeat domain 7A/7b gene.  
 nTPM, normalized transcripts per million

## 1. Introduction

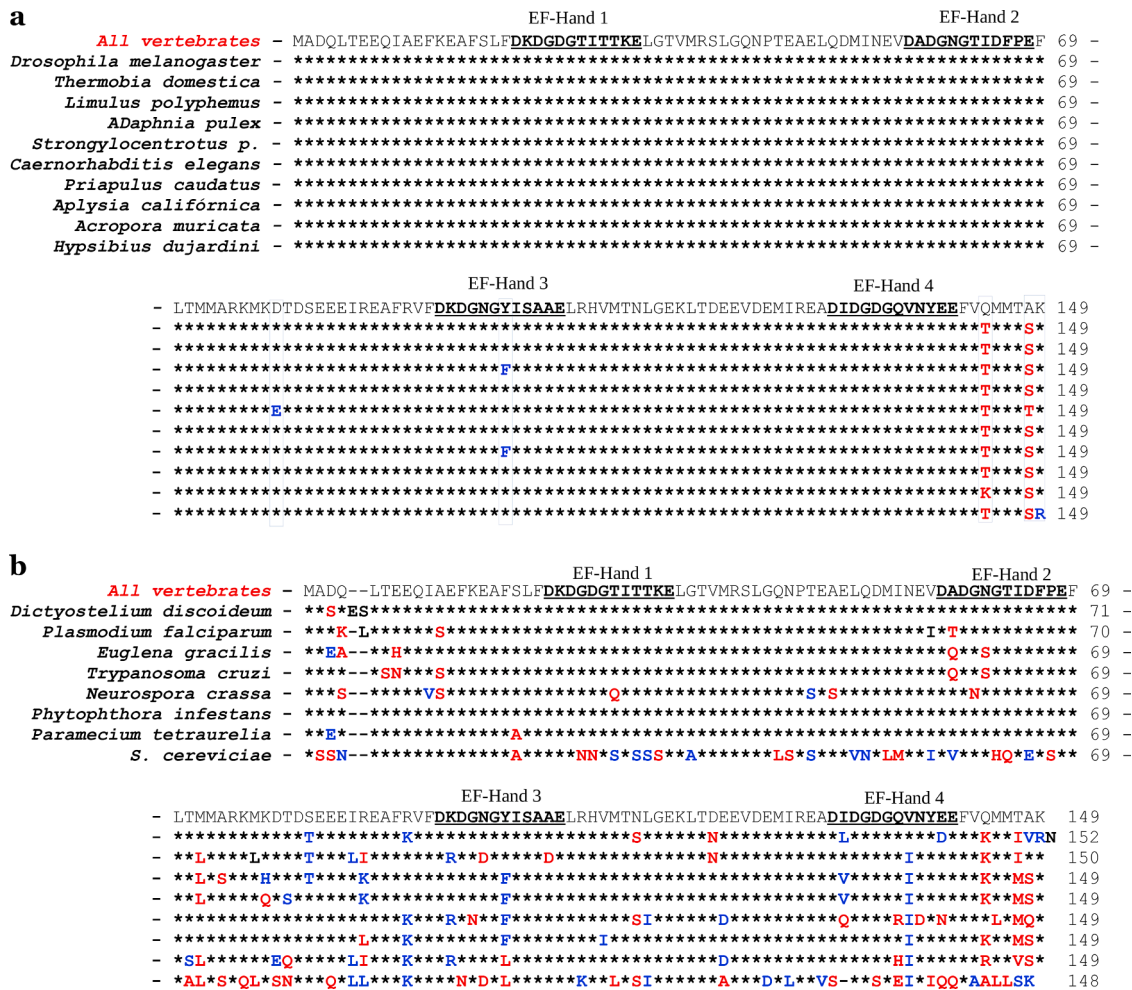
The evolutionary selection of  $\text{Ca}^{2+}$  ions as a key signaling element, based on its charge, ionic radius, coordinating and chemical properties, was accompanied by the appearance of proteins holding distinct  $\text{Ca}^{2+}$ -binding motifs. This made it possible to transduce changes in the concentration of free  $\text{Ca}^{2+}$ , first in the interior of protocells and early prokaryotes, and thereafter in the cytosol and lumen of the different intracellular compartments of more complex eukaryotic cells (reviewed in [1]). Among the proteins containing  $\text{Ca}^{2+}$ -binding EF-hand motifs [2] in eukaryotes, calmodulin (CaM) stands out by its ubiquity, high phylogenetic conservation, a high number of targeted proteins under its control and the versatility of cellular functions, in which this protein is involved (reviewed in [3–7]). CaM has two pairs of EF-hands located at each of the N- and C-lobes connected by a flexible linker (reviewed in [8]). Furthermore, due to the ability of each lobe to act independently, CaM has the capacity to bridge different or identical target proteins, leading to dimerization, or interact with distinct regions of the same protein forming new functional domains [9] (reviewed in [8]). Upon  $\text{Ca}^{2+}$  binding to CaM, a series of hydrophobic patches containing methionine residues are exposed and its 3D-structure undergoes modifications allowing interaction with target proteins, thereby regulating their functions. Nevertheless, CaM is also able to bind to and control some proteins in its  $\text{Ca}^{2+}$ -free form (reviewed in [10,11]).

CaM gene(s) are present in all so far examined eukaryotes, from

human to single cell organisms such as yeast. CaM is not only a ubiquitous protein among eukaryotes, but is also extraordinary well conserved among them [12]. This is exemplified by the fact that the protein is identical in all studied vertebrates, even though the genes show remarkable diversity at the nucleotide level [12]. An alignment of CaM sequences from different invertebrates shows only few changes (2 or 3 amino acids), mostly located in the C-terminus, with respect to vertebrate CaM (Fig. 1a). Of interest, the presence, in most cases, of a threonine at position 144 and serine at position 148 in CaM of invertebrates, in contrast to vertebrates, having a glutamine and an alanine, respectively, at these positions, may suggest that these positions in invertebrate CaM could serve as specific phosphorylation targets for Ser/Thr-protein kinases, as it occurs in different Ser/Thr residues of CaM in many organisms (reviewed in [13,14]).

A single CaM gene is present in most non-metazoan eukaryotes, including yeast/fungi [15–20], protozoans [21,22], and green algae [23]. However, there are some exceptions, as for example in *Trypanosoma cruzi* and *T. brucei*, which have eight and three tandemly repeated CaM genes, respectively, either in two different loci or in a single one [24,25]. Fig. 1b shows an alignment of the sequences of CaM from different protists and yeast/fungi, as compared with vertebrate CaM. Interestingly, the highest number of substitutions is found in the C-terminal helix of EF-hand IV. Even though *Sacharomyces cerevisiae* shows the lowest identity (59% with vertebrate CaM), it can functionally be replaced by vertebrate calmodulin [26]. In plants, several CaM isoforms and multiple CaM-like proteins (CaML) encoded by distinct genes are expressed. For example, in rice 5 CaM genes and 32 CaML genes and in *Arabidopsis* 7 CaM genes and 50 CaML genes are present with differential structural features and specific target recognition properties. We are not considering protist, invertebrate or plant CaMs and CaMLs in this review, as extensive information is available on the properties, functions and evolution of both protein families, particularly in plants (reviewed in [27–29]).

It is well known that many organisms harbor multiple and distinct CaM genes giving rise to identical proteins [30], as for example the three CaM genes present in all so far investigated mammals [31]. The coding sequences for the three human CaM genes (*CALM1,2,3*) share only about 80% identity at the nucleotide level [12]. Comparing nucleotide coding sequences of *CALM1*, *CALM2* and *CALM3* of rat and human reveals similarities of 88, 90 and 90%, respectively. In chicken, 2 CaM genes (*CALM1,2*) are present with an identity of about 85%, and identities of about 92% with their human counterparts [32,33]. This redundancy of identical proteins being expressed by multiple independent genes is puzzling. Seemingly, this phenomenon may represent, at least in part, a safeguard for cell function and viability in case of mutational mishaps. However, something more is perhaps hidden behind this phenomenon. Evidence is mounting that there is differential expression of the three CaM genes in a given tissue, as for example in the white matter of the rat spinal cord [34], among many other tissues. Furthermore, it has been shown that natural occurring monogenic CaM mutations [35] provoke organ-specific dysfunctions, as shown for the heart. This underscores either that the expression of the single affected CaM gene is the only one expressed in specific regions of the myocardium (see Section 4.1), or alternatively, that the expression of the other CaM genes is not apt to rescue the dysfunction. This may occur either because: (i) the three CaM genes are not expressed in sufficient quantities; (ii) they present extemporal expression timing; (iii) the other expressed wild-type CaM is targeted to a different subcellular location; (iv) and/or probably most important, because the arrhythmogenic mutant CaM proteins may work in a dominant negative fashion on target proteins (e.g. RyR2, CaMK-II, SK2 channel,  $\text{Ca}_v$  1.2 channel) and therefore their function cannot be compensated by wild-type CaM [36–38].



**Fig. 1.** a. Comparison of CaM sequences in invertebrates. Multisequence alignment of CaM from the indicated invertebrates, as compared with vertebrate CaM (highlighted in red), using the CLUSTALW multi-alignment tool (<https://www.genome.jp/tools-bin/clustalw>). Identical residues are marked (\*). Residues differing from the ones in vertebrate CaM are highlighted in red (non-conserved substitutions) and blue (conserved substitutions). b. Comparison of CaM sequences in lower eukaryotes. Multisequence alignment of CaM from the indicated lower eukaryotes species, as compared with vertebrate CaM (highlighted in red), using the CLUSTALW tool (<https://www.genome.jp/tools-bin/clustalw>). Identical residues are marked (\*). Residues differing from the ones in vertebrate CaM are highlighted in red (non-conserved substitutions) and blue (conserved substitutions).

**2. From a single EF-hand Ca<sup>2+</sup>-binding primordial protein to CaM with four EF-hands**

EF-hand motifs, first described in the Ca<sup>2+</sup> buffering protein parvalbumin by Kretsinger and Nockolds in 1973 [39], are structural elements found in a large number of proteins of roughly 30 amino acids, containing a loop of 12 residues, providing the Ca<sup>2+</sup> coordinating amino residues, between two perpendicularly arranged alpha helices. Most EF-hands are able to bind Ca<sup>2+</sup>, though with greatly variable affinities. Calcium is bound in a pentagonal bipyramid configuration [40] with oxygen atoms of acidic amino acids and water. As EF-hand proteins have been found in prokaryotes (reviewed in [41]) it is possible that EF-hands have arisen from a single non-eukaryotic organism that gave rise to eukaryotes or that EF-hands have been evolved in early eukaryotes and then were transmitted to prokaryotes [42–45]. Alternatively, it cannot be excluded that EF-hands may have evolved independently in prokaryotes and eukaryotes.

In most cases, two EF-hands, generated by duplication, are placed together in a characteristic way providing enhanced stability and leading to increased Ca<sup>2+</sup> binding affinity. A single protein may have several mostly paired EF-hands. Calmodulin, as well as some other proteins including troponin C and myosin light chains, contains 2 pairs of EF-hands distributed in 2 lobes. The fact that domains 1 and 3, as well as

2 and 4, are most similar points to duplication of a single domain with 2 EF-hands, which themselves have evolved from a single one EF-hand domain. Interestingly, the 4 EF-hands in CaM have different structural and functional properties, which is believed to contribute to the huge flexibility of CaM to interact with and regulate a large number (> 300) of target proteins.

**3. The three mammalian CaM genes**

Mammalian CaM is special in that it is encoded by 3 different nonallelic genes, placed on three different chromosomes, all producing an identical protein sequence [30]. In human, for instance, extensive use of synonymous codons means only about 80% identity at the nucleotide level [12]. On the other hand, the individual gene orthologs among different mammals have higher identities. The three CaM genes have the same intron-exon boundaries, which are also conserved in some non-mammalian species. Chicken and frog lack the CALM3 gene, and only have two CALM genes, which are orthologs to mammalian CALM1 and CALM2, indicating that CALM3 was either developed at a later evolutionary step, or that it was present but lost again in these species. Having three different promoters and 5'/3'-noncoding RNA sequences, the three CaM genes of mammals are differentially regulated during development (reviewed in [46]), and their mRNAs were shown to be

differentially localized in neuronal cells.

### 3.1. CaM genes and pseudogenes. Comparison of their genomic organization: intron-exon organization, chromosomal localization

#### 3.1.1. CaM genes

The chromosomal loci of the three human CaM genes were first reported in 1993 [31] to be on chromosomes 14, 2 and 19 for *CALM1*, *CALM2* and *CALM3*, respectively (Table 1), as analyzed by *in situ* hybridization of metaphase spreads of human lymphocytes, and confirmed with the precise physical localization after the human genome was sequenced [47–49]. *CALM1* and *CALM3* are located on the forward DNA strand while *CALM2* is placed on the reverse strand [47–50] (see Fig. 2). The *CALM2* gene is the longest of the three CaM genes [51], and all human *CALM* genes (as well as all known mammalian CaM genes) have 6 exons and 5 introns (see Table 1). Of interest and supporting a common evolutionary history, the intron/exon boundaries with respect to the coding sequences are conserved, not only in human *CALM1*, *CALM2*, and *CALM3* [52], but also in other mammals, as well as in the two chicken CaM genes. Table 2 summarizes the sequences of the 10 intron/exon boundaries in the CaM genes of human, rat, mouse and chicken, showing that all of them are conserved. The internal exons 2–5, containing only coding sequences, are identical in length, while exons 1 and 6, with mostly noncoding sequences, are different in length (see Table 1). Exon 1 includes the 5'-UTR as well as the ATG start codon at its 3'-end, while exon 6, includes the 10 C-terminal codons as well as the 3'-UTR sequence (see Table 1).

The location of the orthologous genes upstream and downstream of

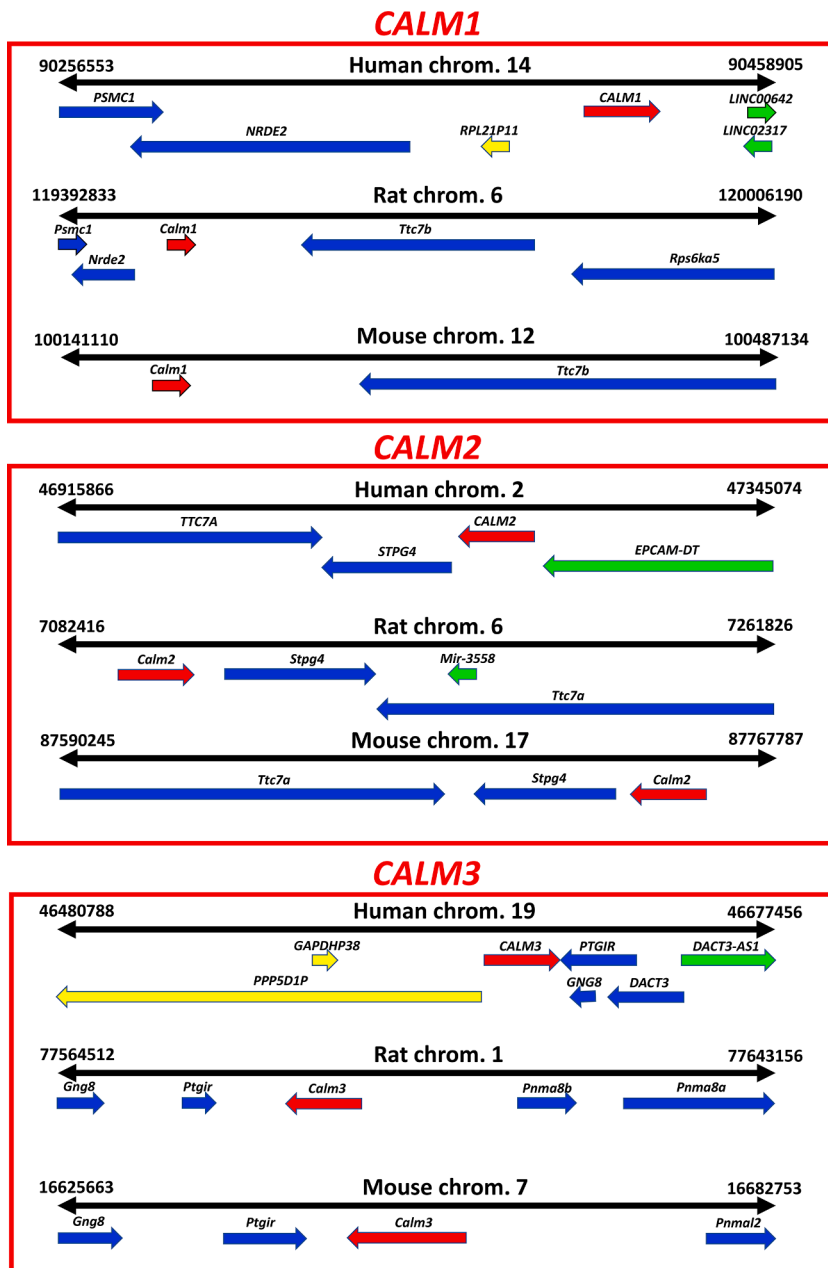
the three CaM genes in human, rat and mouse is similar (see Fig. 2). The genes *PSMC1/Psmc1* and *NRDE2/Nrde2*, encoding the 26S proteasome subunit-ATPase 1 and the nuclear exosome regulator NRDE2 (necessary for RNA interference-2), are located upstream of *CALM1/Calm1* in human and rat, respectively. The genes *STPG4/Stpg4* and *Ttc7b*, coding for the sperm-tail PG-rich repeat containing 4 protein and the tetratricopeptide repeat domain 7b protein are located downstream of *CALM1* in human rat and mouse, respectively. The gene encoding the 7a isoform of the latter protein is located upstream of *CALM2/Calm2* in human and mouse, while it is placed downstream in rat. Moreover, the genes *PTGIR/Ptgir* and *GNG8/Gng8*, coding the prostaglandin I2 receptor and the G protein subunit g8, respectively, are located downstream of *CALM3* in human, and upstream of *Calm3* in rat and mouse. The occurrence of synteny, in the chromosome regions where the CaM genes in the three species are located, reinforces the idea of the common evolutionary history, as stated above. Whether regulatory elements in the mentioned genes located in the genomic regions close to the CaM genes have an influence on the transcriptional activities of the 3 CaM genes is unknown.

Even though mammalian CaM genes encode identical proteins, as described above, the nucleotide sequence varies quite markedly [53]. As an example, Fig. 3a shows extensive differences in the nucleotide coding sequence of the three CaM genes in human, presenting more than one hundred nucleotide differences, almost exclusively from differences in the third codon position. Similar numbers of variations are also found in rat (Fig. 3a) and mouse (*not shown*) and mostly in the same positions. In contrast, only less than half of the variations are found when comparing the nucleotide coding sequences of the three CaM genes in human versus

**Table 1**  
Human calmodulin genes and mRNAs.

Properties	<i>CALM1</i>	<i>CALM2</i>	<i>CALM3</i>	Refs.
Chromosome locus	14q24-q31 (14q32.11)	2p21.1-p21.3 (2p21)	19q13.2-q13.3 (19q13.32)	[31,47–49]
Gene length (bp)	11,240	16,885	9453	[47–49]
Exons length (bp)	Exon 1 (251) Exon 2 (31) Exon 3 (144) Exon 4 (107) Exon 5 (136) Exon 6 (3573)	Exon 1 (71) Exon 2 (31) Exon 3 (144) Exon 4 (107) Exon 5 (136) Exon 6 (721)	Exon 1 (108) Exon 2 (31) Exon 3 (144) Exon 4 (107) Exon 5 (136) Exon 6 (1658)	[50,52]
Introns length (bp)	Intron 1 (2831) Intron 2 (1163) Intron 3 (2459) Intron 4 (410) Intron 5 (174)	Intron 1 (5676) Intron 2 (8071) Intron 3 (126) Intron 4 (427) Intron 5 (918)	Intron 1 (4389) Intron 2 (2339) Intron 3 (141) Intron 4 (257) Intron 5 (143)	[50,52]
Promoter				[47,51,176]
TATA box	4	2	1	
CAAT box	4	4	3	
GC box	1	1	-	
AGGGA motif	-	6	3	
CRE bind. site	-	2	3	
SP1 bind. site	5	1	8	
AP1 bind. site	-	1	1	
AP2 bind. site (number)	-	2	5	
Poly-A signals (number, sequences)	(5) 5'-AATAA-3' 5'-AATAAA-3' 5'-TTTTAT-3'	(3) 5'-AATAAA-3' 5'-TTTTAT-3'	(3) 5'-ATAAA-3' 5'-AATAAA-3' 5'-AATAAAAT-3'	[94,177]
CPEB-binding sites (number, sequence)	1 5'-UUUUUAU-3'	2 5'-UUUUUAU-3'	None	[98]
Potential CTCF-binding sites (sequence) <sup>1</sup>	5'-CACACCCTGGTGGC-3' Score 10.9943	5'-TTGCACCTTGTGGA-3' Score 7.33699	5'-GGAACTGCT-3' Score 12.137	[178]
miRNA Let7a binding to 3'-UTR (sequence)	5'-UACCUC-3'			[103]
Staufen2 binding			Located in a retained intron within the 3'-UTR	[99]

<sup>1</sup> The potential CTCF-binding sites shown were determined using the CTCFBSDB 2.0 prediction tool. For each gene, the sequence shown corresponds to that with the highest score in the coding orientation from the output reported. bp, base pair; CPEB, cytoplasmic polyadenylation element binding protein; CTCF, CCCTC-binding protein.



**Fig. 2.** Genes located upstream and downstream of the three CaM gene loci in human, rat and mouse. The drawing shows the location of protein-coding genes (blue arrows), pseudogenes (yellow arrows) and some RNA-coding genes (green arrows) upstream and downstream of the three CaM genes (red arrows) in human, rat and mouse. Other predicted genes in the indicated DNA segments are not shown. The length of the genes and inter-gene segments are not drawn to scale. *CALM1/Calm1*, calmodulin-1 gene; *CALM2/Calm2*, calmodulin-2 gene; *CALM3/Calm3*, calmodulin-3 gene; DACT3, disheveled binding antagonist of  $\beta$ -catenin 3; DACT3-AS1, disheveled binding antagonist of  $\beta$ -catenin 3 antisense RNA 1; EPCAM-DT, epithelial cell adhesion molecule divergent transcript lncRNA; GAPDHP38, glyceraldehyde 3-phosphate dehydrogenase pseudogene 38; GNG8/Gng8, G protein subunit g8; LINC00642, lncRNA-642; LINC02317, lncRNA-2317; Mir-3558, microRNA-3558; NRDE2/Nrde2, nuclear exosome regulator NRDE2 (necessary for RNA interference-2); Pnma8a/b, paraneoplastic Ma antigen protein 8a/8b; Pnma12, paraneoplastic Ma antigen-like protein 2; PPP5D1P, Ser/Thr protein phosphatase tetratricopeptide repeat domain containing 1 pseudogene; PSMC1/Psmc1, 26S proteasome subunit - ATPase 1; PTGIR/Ptgir, prostaglandin I2 receptor; RPL21P11, ribosomal protein L21 pseudogene 11; Rps6ka5, ribosomal protein S6 kinase a5; STPG4/Stpg4, sperm-tail PG-rich repeat containing 4; TTC7A/Ttc7b, tetratricopeptide repeat domain 7A/7b. Adapted from [47–50].

rat, mouse or cow (see Fig. 3b). This shows that there are more differences in the three CaM genes among them in a given species than in a given CaM gene among different vertebrates.

### 3.1.2. CaM pseudogenes

The lack of introns, truncation and mutational disruption of the CDS are the main characteristics of eukaryote pseudogenes (e.g. [52,54,55]), although these characteristics do not always apply. Pseudogenes are widely distributed in the genome of many living organisms, and despite early assumptions of lack of function, a growing number of evidences support the functionality of some of them (reviewed in [56–58]). Some of the mechanisms by which pseudogenes could have a functional role include: (i) they could encode a sequence-altered or truncated functional protein reminiscent of the parental one; (ii) they could be transcribed into siRNAs or antisense RNAs, preventing translation of the parental gene, or yielding lncRNAs, which directly interact and regulate the expression of certain proteins; and (iii) they could favor DNA folding allowing interaction of different chromatin regions, and this could

positively or negatively affect the expression of nearby genes, and to allow the conversion of their parental genes into dysfunctional ones by transferring segments via non-allelic recombination (reviewed in [57, 58]).

Pseudogenes derived from the CaM genes have been identified in multiple vertebrates, including human [59–61]. Among them, a couple of presumably non-functional intronless pseudogenes, derived from the genuine human *CALM1*, denoted *CALMIP1* and *CALMIP2* presenting disruptive mutations in their CDS were first described by Rhyner et al. [52], as well as three retro-pseudogenes derived from the human *CALM2* gene [62]. *Alu* elements, repetitive sequences that are cut by the restriction endonuclease *Alu*, were identified upstream of the 5'-UTR and downstream of the 3'-UTR in one of these pseudogenes [62]. *Alu* elements are dispersed by retro-transposition. Transcription of *Alu* elements, reverse transcription of their RNA and insertion of the resulting cDNA into the targeted genes are the three steps required in the dispersion process, usually leading to multiple effects by modifying functions of other genes [63]. The retro-pseudogenes derived from

**Table 2**  
The intron/exon boundaries of the CaM genes in human, rat, mouse and chicken.

Species	Intron/ exon boundary	CALM1	CALM2	CALM3
Human	1	ATGgttag...(+3 M1)	ATGgttag...(+3 M1)	ATGgttag...(+3 M1)
	2	...tatagGCTGA	... aacagGCTGA	...tacagGCTGA
	3	TGCTGgtaag...(+1 E12)	TGCAGgtgag...(+1 E12)	TGCAGgtgag...(+1 E12)
	4	...tcagAATTC	...aacagAATTC	...cccagAGTTC
	5	TGATGgtaag...(+1 G60)	TGATGgtaag...(+1 G60)	AGATGgtgag...(+1 G60)
	6	...tcagGTAAT	...tcagGTAAT	...tcagGGAAC
	7	ACAAGgtaat...(+3 K95)	ATAAGgtacg...(+3 K95)	ACAAGgtaag...(+3 K95)
	8	...cacagGATGG	...gtagGATGG	... cccagGATGG
	9	TGAAGgtaaa...(+1 E141)	TGAAGgtaaa...(+1 E141)	TGAAGgtgag...(+1 E141)
	10	...tcagAATTC	...tcagAGTTT	...cacagAGTTT
Rat	1	ATGgttag...(+3 M1)	ATGgttag...(+3 M1)	ATGgttag...(+3 M1)
	2	...ttagGCTGA	...tcagGCTGA	...tcagGCTGA
	3	TGCTGgtaag...(+1 E12)	TGCAGgtaag...(+1 E12)	TGCAGgtgag...(+1 E12)
	4	...tcagAATTC	...tcagAATTC	...cccagAGTTC
	5	CGACGgtgag...(+1 G60)	CGACGgtgag...(+1 G60)	TGATGgtgag...(+1 G60)
	6	...taaagGGAAT	...taaagGGAAT	...gtagGCAAT
	7	ACAAGgtaat...(+3 K95)	ACAAGgtaat...(+3 K95)	ACAAGgtaag...(+3 K95)
	8	...cacagGATGG	... cacagGATGG	... cccagGATGG
	9	TGAAGgtaaa...(+1 E141)	TGAAGgtaaa...(+1 E141)	TGAAGgtgag...(+1 E141)
	10	...tcagAATTC	...tcagAATTC	...ggcagAGTTT
Mouse	1	ATGgttag...(+3 M1)	ATGgttag...(+3 M1)	ATGgttag...(+3 M1)
	2	...ttagGCTGA	...tcagGCTGA	...tatagGCTGA
	3	TGCTGgtaag...(+1 E12)	TGCAGgtgag...(+1 E12)	TGCTGgtaag...(+1 E12)
	4	...tcagAATTC	...cccagAGTTC	...tcagAATTC
	5	TGATGgtaag...(+1 G60)	TGATGgtgag...(+1 G60)	TGATGgtaag...(+1 G60)
	6	...taaagGCAAT	...gtagGCAAT	...tcagGTAAT
	7	ACAAGgtaat...(+3 K95)	ACAAGgtaag...(+3 K95)	ACAAGgtaat...(+3 K95)
	8	...cacagGATGG	... cccagGATGG	... cacagGATGG
	9	TGAAGgtaaa...(+1 E141)	TGAAGgtgag...(+1 E141)	TGAAGgtaaa...(+1 E141)
	10	...tcagAATTC	...ggcagAGTTT	...tcagAATTC
Chicken	1	ATGgttag...(+3 M1)	ATGgttag...(+3 M1)	-
	2	...ttagGCTGA	...tcagGCTGA	-
	3	TGCTGgtaag...(+1 E12)	TGCAGgtgag...(+1 E12)	-
	4	...ttagAATTC	...tcagAATTC	-
	5	TGATGgtaag...(+1 G60)	TGATGgtaag...(+1 G60)	-
	6	...taaagGTAAT	...tcagGCAAT	-
	7	ACAAGgtatt...(+3 K95)	ACAAGgtact...(+3 K95)	-
	8	...cacagGATGG	... cccagGATGG	-
	9	TGAAGgtaat...(+1 E141)	TGAAGgtaag...(+1 E141)	-
	10	...tcagAATTC	...tcagAGTTT	-

The transcripts in the databases Ensemble (<https://www.ensembl.org/index.html>) or NCBI (<https://www.ncbi.nlm.nih.gov/>) used in the analysis were as follows: From human: *CALM1*, ENST00000356978.9; *CALM2*, ENST00000272298.12; and *CALM3*, ENST00000291295.14; from rat: *Calm1*, ENSRNOT0000064679.2; *Calm2*, ENSRNOT0000022603; and *Calm3*,

GenBank X14265.1; from mouse: *Calm1*, ENSMUST00000110082.11; *Calm2*, ENSMUST00000040440.7; and *Calm3*, ENSMUST00000019514.10; and from chicken: *Calm1*, ENSGALT00000045488.3; *Calm2*, ENSGALT00000016279.5. Exon sequence segments are highlighted in capital letters. The site of splicing in the codon and the amino acid involved are shown in parenthesis.

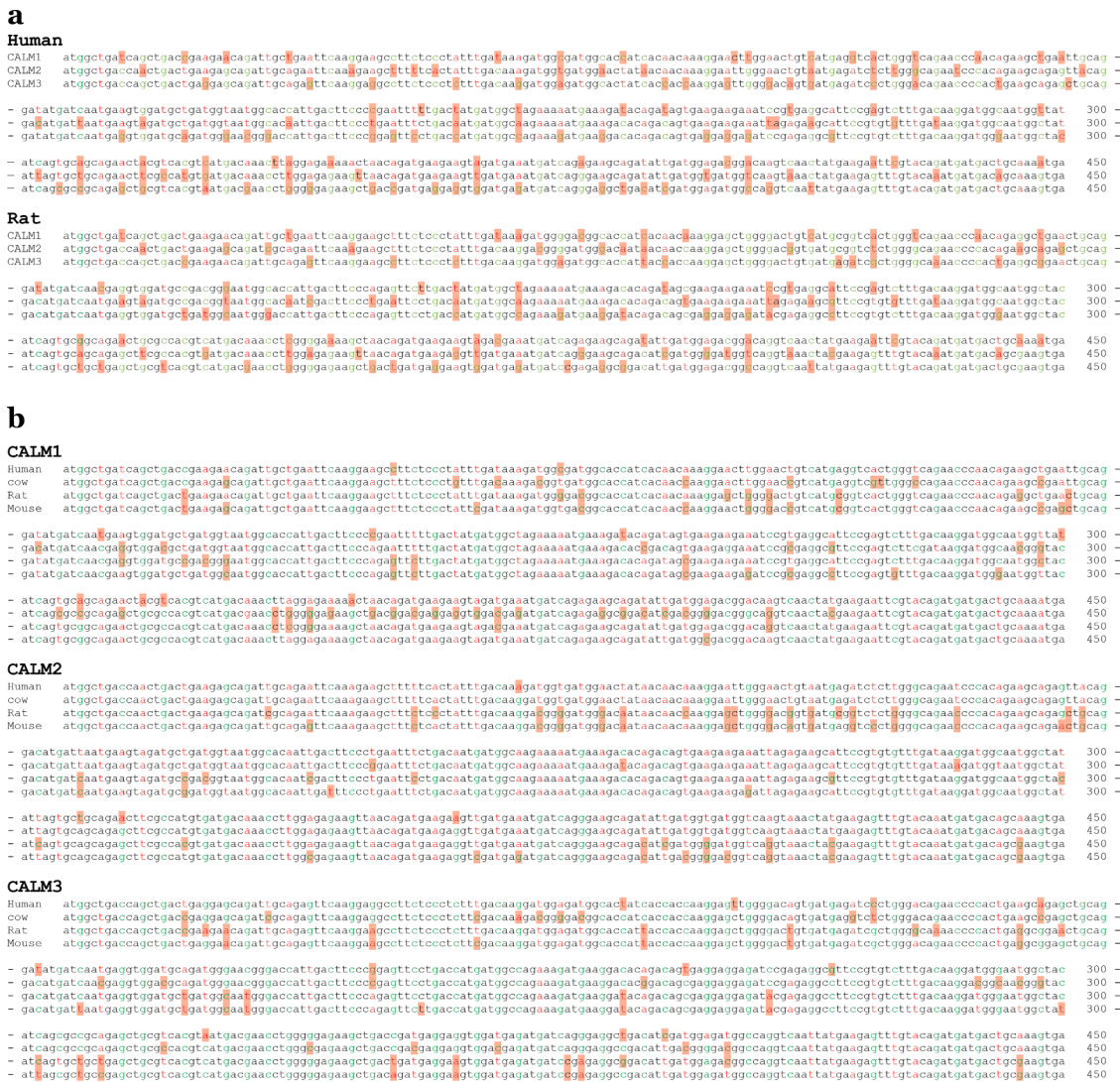
human and murine *CALM2* and *CaMII/Calm2* genes, respectively, as well as other unrelated genes, were shown to be present in different species, and the evolutionary divergence time between human and other primates, and among different rodent species, was established [64,65].

Four processed CaM pseudogenes of rats, denoted *ISC8*, *ISC9*, *ISC19*, and *ISC27* are derived from the genuine *CaMI/Calm1* gene (*ISC9*) or the *CaMII/Calm2* gene (*ISC8*, *ISC19* and *ISC27*), respectively. The *ISC9* gene differs from the *ISC8* in that it does not contain frameshift mutations in its CDS [66–68]. Transcripts of pseudogenes that do not contain disabling mutations in the CDS could produce CaM-like proteins if they are correctly transcribed and translated and therefore could potentially exert diverse functional roles. Indeed, a functional intronless CaM-like gene in human has been found [52,60]. This pseudogene, identified in the NCBI database (<https://www.ncbi.nlm.nih.gov/>) as CaM-like protein 3 (*CALML3* or *CLP*) is highly expressed in epithelial cells, and its downregulation during tumor transformation suggests that it may be involved in differentiation processes [60]. The nucleotide coding sequence identity of the human *CALML3* transcript with those of *CALM1*, *CALM2* and *CALM3* is 70.7%, 69.8% and 80%, respectively. In chicken, an intronless CaM pseudogene denoted *cCMI*, also identified as CaM-like protein 3 (*CALML3*), as in human, codes for a 149 residues protein, and is differentially expressed in tissues. Comparing the coding sequence of human and chicken *CALML3* shows that they share 79.1% identity. This chicken protein differs in 19 residues from the one encoded by the CaM gene *cCL1*, 11 of them non-conservative, and it is affected in its Ca<sup>2+</sup>-binding capacity at EF-hands I and IV, but not EF-hands II and III, as critical Ca<sup>2+</sup> coordinating acidic residues are altered in the former EF-hands [55]. The authors of this study suggested that this protein may have a specific role in striated muscle physiology, as this was the only tissue among several others tested, where the authors found expression of this protein. Its subcellular localization and availability may determine its impact on skeletal muscle specific processes.

### 3.2. Regulation of CaM expression

Since CaM is involved in a huge variety of physiological functions, even within a single cell, it is of no surprise that a certain level of CaM expression is not only required in a specific tissue and cell with a given set of CaM targets of different abundance, locations and affinities for CaM, but also in different subcellular compartments. It has been shown in living smooth muscle cells that only 5% of the total CaM concentration is freely diffusible [69] and that, upon increase of cytosolic Ca<sup>2+</sup>, CaM could be translocated to the nucleus. This has been confirmed in cardiomyocytes where soluble free CaM was found in very low concentration (ca. 1% of total) and increasing intracellular Ca<sup>2+</sup> led to translocation of CaM to the nucleus as shown by Bers and collaborators [70]. It is believed that the presence of CaM binding sites exceeds the amount of available CaM, making diffusion of free CaM difficult and indicating that there must be a selection of targets based on affinities and kinetics [71]. Therefore, subcellular CaM pools could play an important role in target regulation [72].

Another factor regulating the available CaM concentration is the presence of CaM targets that bind in the absence of calcium, e.g. GAP-43/neuromodulin, which has been proposed to act as CaM buffer [73] to prevent Ca<sup>2+</sup> dependent signaling through CaM. In addition to this CaM storage hypothesis, it has been proposed that CaM could be temporally “masked” and thereby not anymore able to regulate targets by posttranslational modification, e.g. phosphorylation is shown to reduce its activity in regulating targets (reviewed in [13,14]). Moreover,



**Fig. 3.** a.b. Alignment of nucleotide sequences of the three human and other vertebrate CaM genes. Alignment of the three human *CALM* genes (a), and alignment of each gene compared to the orthologs from four species (b). Nucleotide sequences were retrieved from NCBI database under the following accession numbers: NM\_006888.6, NM\_001743.6, and NM\_005184.4 for human *CALM1*, *CALM2* and *CALM3*, respectively; NM\_031969, NM\_017326.3, and BC063187.1 for rat *Calm1*, *Calm2* and *Calm3*; NM\_001242572.1, NM\_001242587.1, and NM\_001046249.2 for Cow *Calm1*, *Calm2* and *Calm3*, and NM\_001313934.1, NM\_007589.5, and NM\_007590.3 for mouse *Calm1*, *Calm2* and *Calm3*. Only coding sequences were aligned with Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) using default settings. The start and stop codons are included in the alignments as well as the nucleotide position numbers. Non-identical nucleotides are highlighted in red. The third position in each codon is marked in red=AT3 or green=GC3.

as CaM contains two Ca<sup>2+</sup> binding lobes, each with two EF-hands with different Ca<sup>2+</sup> affinities and binding kinetics that could act in a somewhat independent way, this may contribute to its extraordinary flexibility in Ca<sup>2+</sup>/CaM-dependent signaling.

The available information on the presence and especially the significance of *cis*-regulatory elements controlling the differential expression of the three human CaM genes and their alternatively polyadenylated mRNAs is rather scanty. Obviously, a major obstacle is the fact that it is not possible to discriminate on the level of protein accounted for each gene namely, which CaM gene is responsible for the final protein product as CaM from all three *CALM* genes are identical. Whether and how these potential regulatory elements are used by the relevant transcription factors and other regulatory proteins as well as RNAs and whether there is cross-talk among these elements to facilitate or repress the relative expression of a particular CaM gene and its mRNA product over the others, depending on different physiological states of the cell, is only very fragmentarily known so far. As work on CaM gene regulation carried out before and around 2000 has been reviewed by the

group of Strehler [12], we will only discuss some of the major conclusions and focus on a few more recent examples of *cis*-regulatory elements and their functions, as well as cases of epigenetic control of the CaM genes in human and other organisms and list and compare these potential regulatory elements in table form.

**3.2.1. Promoters, insulators**

The promoter regions of the three human CaM genes contain a variable number of regulatory elements, such as TATA, CAAT, GC boxes; AGGGA and CRE motifs, transcription factor-binding sites (e.g. for SP1, AP1, AP2, CREB), as shown in Table 1. This may impinge on the relative efficiency of their transcriptional activity [74]. The function of insulators in vertebrates is mediated upon binding to CTCF, an eleven zinc finger DNA-binding domain protein, which could interfere with the communication between promoters and enhancers or by connecting two CTCF binding sites thereby mediating chromatin looping (reviewed in [75]). Of interest, potential CTCF-binding sites have been identified in all three human CaM genes (see Table 1). Possible functional

significance of the cAMP response element found in human, rat, and the *Xenopus CaM II* promoter has been supported by findings showing that cAMP upregulates *CaM I* and *CaM II* in rat PC12 cells [76].

It was demonstrated that a 1218 bp segment of the promoter and the 5'-leading sequence of the chicken *Calm2* gene containing CpG islands, typical for housekeeping genes and in many cases not containing TATA boxes, was sufficient for the initiation and regulation of transcription [77]. Gene promoters may contain single or multiple dispersed transcriptional start sites (reviewed in [78]), and in mammals CpG island-containing promoters usually are associated with dispersed transcriptional start sites and housekeeping genes (reviewed in [79]). The above mentioned chicken *Calm2* promoter was shown to have lower activity in differentiated cells, as compared to proliferating ones, in agreement with the lower endogenously-produced CaM level observed during differentiation [77]. As the chicken *Calm2* promoter shares similarities with mammalian *CALM2* promoters, it could be predicted that it works in a similar way as the mammalian promoter. Similar studies were conducted using promoter segments of the rat *Calm2* and *Calm3* genes in subsequent studies. In these cases, transgenic mice expressing constructs of promoter segments with the  $\beta$ -galactosidase gene as a reporter were used to monitor the expression in different regions and cell types of the central nervous system. In most but not all cases,  $\beta$ -galactosidase expression matched the expression of the endogenous CaM genes in adult mice [80,81]. Additional studies were performed in testis, specifically demonstrating *Calm2* expression in spermatocytes, and not in other cellular stages during spermatogenesis, when *Calm1* and *Calm3* are expressed [82].

### 3.2.2. Epigenetic control

Epigenetic labeling of DNA and histones plays a central role in the pattern of gene transcription during development, differentiation and the physiological status of the cell [83–86]. Of interest, the cytosine methylation level of human *CALM1* in spermatocytes progressively decreases from 25 to 60 years of age, as observed with other genes (e.g. *CDH13* and *STMN2* encoding T-cadherin and stathmin-2, respectively) [87], suggesting a differential expression level of CaM during aging. Also, a prospective study in a rather small cohort of type 2 diabetes patients demonstrated hypomethylation of CpG islands in the promoter of *CALM2* in peripheral leukocytes, contradicting the hypothesis that methylation of this gene contributes to the onset of type 2 diabetes as previously suspected [88].

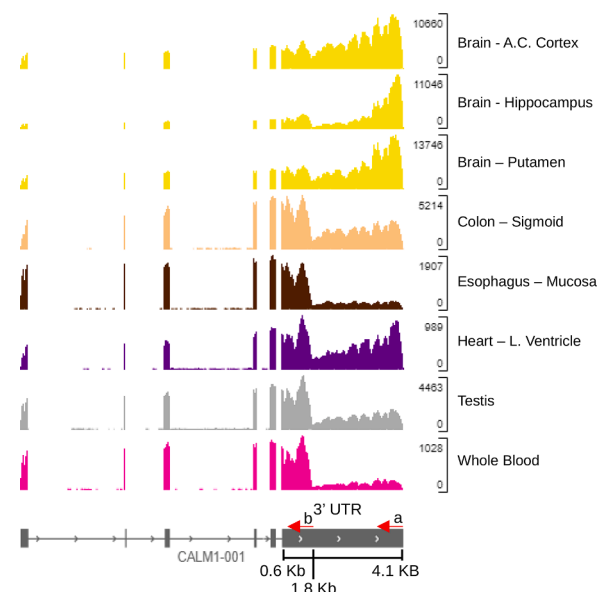
### 3.2.3. Tissue and cell type specific expression of the three mammalian CaM genes

CaM is expressed in all cells of eukaryotes in variable amounts. To determine which CaM gene is expressed in each case in a specific cell type and physiological condition may shed light on how the three CaM genes are contributing to the function of the cell or tissue. Expression of the individual CaM genes can only be measured on the transcript level, because the proteins coded by the three genes are identical. Therefore, transcription data cannot be directly linked to the levels of CaM protein. Human *CALM1* expression was detected in all organs tested including heart, placenta, lung, and kidney, though it was found to be particularly abundant in brain and skeletal muscles, and very scarce in liver and pancreas [52]. *CALM2* also presents very high expression in human brain, followed by similar expression levels in heart, placenta, lung, liver, skeletal muscle and kidney, and in much lesser extent in pancreas [51]. In the heart, the expression of different CaM genes is not homogeneous through its different tissular structures. Thus, in the Purkinje fibers, specialized to conduct the electrical stimuli to the left and right heart ventricles, the expression of *CALM1* is higher than in the working atrial and ventricular non-specialized cardiomyocytes, while the expression of *CALM3* is higher in both the left and right epicardium than in the endocardium [89]. This suggests that each cardiac structure may require a different quantitative and/or temporal output of CaM for their correct functioning. The specific downregulation of only one of the three

CaM genes could give some hints on the functionality exerted on the cell. Several reports have been published in this context [90,91]. However, these studies do not specifically address whether the knocked-down CaM gene was the only one regulating the cell function under study, but rather that the supply of CaM derived from the untouched genes is unable to fully compensate for the observed deficiencies. This could be due to lack of sufficient temporo-spatial supply of CaM to the targeted CaM-binding proteins.

### 3.2.4. Differential expression of CaM induced by signaling processes, during cell differentiation and development

The expression of the CaM genes has been shown to be dependent on the physiological conditions of the cell. For instance, the half-life of *CALM1* and *CALM3* mRNAs is longer than the one of *CALM2* mRNA in mouse myoblasts during both proliferation and differentiation, while the half-life of all three *CALM* mRNAs increases during differentiation, but the actual level of these transcripts decreases, as well as the CaM level, although with distinct kinetics [92]. CaM expression is also subjected to endocrine control. Thus, it has been shown that resection of the adrenal glands decreased the level of *CALM3* mRNA in the rat cerebral cortex and the hippocampus, without affecting *CALM1* and *CALM2* mRNAs, while corticosterone administration prevented *CALM3* downregulation as expected [93]. Evidence that differential transcription of the *CALM* genes in rats provide a means to control the spatial and temporal availability of CaM pools was reviewed in [12]. The *CALM* mRNAs exhibit cell type and cell state-specific mRNA abundance. In rats *CALM1* and *CALM2* appear in neurites and dendrites, while *CALM3* predominates in the cell body. Humans express a long *CALM1* transcript (4.1 kb) as well as a shorter 1.8 kb version. The longer *CALM1* (4.1 kb) human mRNA has been demonstrated to have increased stability. It was suggested that this protects the transcripts from degradation during trafficking to distal locations (reviewed in [12]). We searched for evidence of differential expression of the two *CALM1* RNA isoforms, in the RNA sequence database GTExPortal (gtexportal.org). It appears that the 4.1 kb version is more specific to the brain (Fig. 4). The shorter version appears to be the most abundant in all tissues sequenced, except in brain, where the long form dominates. This may reflect the need for



**Fig. 4.** RNA isoforms from alternative polyadenylation in *CALM1*. RNA-seq reads from representative tissues. The figure shows RNA sequencing results of *CALM1*. The 3' UTR starts at position 0.6 Kb of the transcript. Evidence is seen of reverse transcriptase activity starting from d-TTT primers at position 1.8 Kb (red arrow a) and at position 4.1 Kb (red arrow b). <https://www.gtexportal.org/>.



stable *CALM* mRNA when trafficking to distal dendrites in brain.

### 3.2.5. 5'-UTR and 3'-UTR

The 5'-UTR of mRNAs may be involved in transcriptional or translational regulation whereas the 3'-UTR of mRNAs may act, among other functions, in transport and location of the mRNA to specific intracellular locations where translation is required, as well as controlling mRNA stability by promoting or inhibiting its degradation. In mammals, both UTRs are unique in the three CaM mRNAs, presenting highly conserved sequence homology (76–95%) among human, rat and mouse, and 85–96% between human and several cetaceans.

The importance of the 5'-UTRs was demonstrated in the three human CaM genes in experiments testing transcriptional activity with the use of the luciferase reporter gene in teratoma cells, as its absence strongly decreased transcription [51] and discussed in Tootenhoofd and Strehler [12]. The least conserved 3'-UTR is found in the *CALM3* transcript (79.7% and 78.2% comparing human with rat and mouse, respectively), while the most conserved corresponded to the *CALM2* transcript (90.2% and 92.7% comparing human with rat and mouse, respectively), closely followed by the *CALM1* transcript (89.3% and 86.2% comparing human with rat and mouse, respectively) [94,95]. This is close to the somewhat higher conservation (91.1–92.7%) of the CDSs of the three CaM genes comparing human with rat and mouse transcripts indicating an important function of the 3'non-coding region [94].

There are putative CPEB binding sites in the 3'-UTR of human *CALM1* and *CALM2* mRNAs (see Table 1). The CPEB protein, first identified in *Xenopus* oocyte, binds to the 3'-UTR at specific sites and regulates mRNA polyadenylation and translation (reviewed in [96]). CPEB activity is regulated by phosphorylation with diverse protein kinases, including CaMK-II [97,98]. Nevertheless, CPEB phosphorylation does not affect RNA binding, a process requiring  $Zn^{2+}$ , and which is mediated by two RNA recognition motifs and the Cys/His-rich C-terminus of CPEB needed for interaction with the consensus sequence 5'-UUUUUUAU-3' [98] (see Table 1).

The 3'-UTR of the human *CALM3* transcript contains three polyadenylation sites, accounting for the identified 0.9, 1.9 and 2.3 kb transcripts [94]. The close homology of these UTRs denotes close phylogenetic relationship as well. The 3'-UTR in the long mRNA of mouse *Calm3*, an isoform of the canonical *Calm3* mRNA, contains an intron forming a hairpin duplex where the double-stranded RNA-binding protein, denoted Stauf2, binds. This binding promotes the location of the long mRNA molecule in dendritic terminals, only in activated synapses in mouse hippocampal neurons without affecting its stability [99].

Among the transcripts of *Calm1* in mouse, isoforms containing short and long 3'-UTRs were found. Knocking down the mRNA isoform with the long 3'-UTR, but not the one with the short 3'-UTR, blocks the development of axon of dorsal root ganglion neurons, restricts their rostral migration to the hindbrain in embryos, and impairs activation of hippocampal neurons in the adult [100]. The presence of 3 CaM genes in mammals each with several variants due to differential polyadenylation with different regulatory properties may allow for a high level of orchestration and adaptation to different physiological conditions, pinpointing any specific function or targeted CaM-binding protein(s) involved.

### 3.2.6. Regulation by miRNAs

MicroRNAs usually downregulate the expression of proteins by inducing the degradation of the mRNAs via the dicer/RNA-induced silencing complex. In transgenic mice, it was shown that overexpression of the miRNA *miR-1* downregulates CaM after binding to the 3'-UTR of *CALM1* and *CALM2* [101,102]. Subsequent downregulation of CaM translation leads to the negative control of transcription factors involved in cardiac hypertrophy. This anti-hypertrophic effect is produced by the inhibition of the CaM-dependent calcineurin/NFAT and CaMK/Mef2a pathways [101]. Studies performed in a mouse model have shown that *miR-1* is downregulated during heart failure [101]. On

the other hand, in conjunction with MLCK repression, low CaM levels contribute to sarcomere damage and impaired cardiac contraction [102]. The miRNA *Let7a* has also been shown to hybridize with the 3'-UTR of human *CALM1* at the 5'-UACCUC-3' sequence, a segment that is conserved in rat and mouse. Thus, this interaction also produces the downregulation of CaM translation leading to a cardiac anti-hypertrophic effect [103]. Another miRNA downregulating CaM expression was shown to be *miR-335*, although the authors of this work did not specify on which *CALM* gene it was acting [104]. The microRNA *hsa-miR-4709-3p* was identified to act on *CALM3* in kidney, and this microRNA strongly downregulated CaM mRNA expression in patients suffering chronic kidney disease [105]. Curiously, the decreased plasma levels of *miR-335* in patients suffering an acute ischemic stroke inversely correlated with the plasma levels of CaM [104], indicating the importance of other factors regulating the level of plasma CaM levels. Moreover, the miRNA *miR-4661-3p* potentially interacts with the 3'-UTR of CaM mRNA, although this miRNA stabilizes the CaM mRNA, instead of inducing its degradation. This was due to increasing the stability of the AU-rich elements, where *miR-4661-3p* is expected to bind. Suppression of *miR-4661-3p* by peptide nucleic acid inhibitors lowered CaM levels and suppressed neuroinflammation mediated by microglial cells [106].

### 3.2.7. Differential subcellular localization of CaM mRNAs

Establishment of local mRNA pools, which could be far away from the cell body in neurons, is believed to contribute to the post-transcriptional regulation leading to temporal and spacial fine-tuning of CaM gene expression. The initial discovery of specific subcellular localization for several CaM RNAs in differentiated PC12 cells in 1993 by Zhang et al. [107] showed that *CALM1* and *CALM2* mRNAs mostly localized to the neurites and that *CALM3* mRNA was positioned in the cell body or close to it.

The group of Gulya investigated the differential subcellular localization of the different mRNAs of the 3 CaM genes in neuronal rat tissue, which contains the highest CaM concentration in the mammalian body (reviewed in [108]). The presence of different patterns of mRNAs derived from the same CaM gene, detected on Northern blots from different tissues or cell populations as well as degradation sequence elements in certain transcripts of the same CaM gene but not in others, pointed to posttranscriptional regulation on the level of mRNA stability due to differential polyadenylation (see also chapter 3.2.8 on codon usage).

Targeting specific mRNA to dendritic sublocalization, where CaM mRNA is colocalized together with other transcripts, including alpha CaMKII, and assumed to be translated there, has been shown in a number of articles indicating an increased need of CaM signaling for postsynaptic activities (reviewed in [108]) and [109]. Differential transport of mRNAs from the 3 CaM genes to functional different compartments in neurons and glia cells of the rat hippocampus [110] as well as in the midbrain-brain stem region has been shown by the Gulya group [111].

The intracellular distribution of the mRNAs derived from the 3 CaM genes also varies in isolated mouse ventricular cardiomyocytes. It was shown that the mRNA from the *CALM2* gene is more concentrated in the central part of the cell, close to the nucleus, and the *CALM3* mRNA is concentrated in the cell periphery, while the *Calm1* mRNA is mostly located between the other two [112]. It is possible that the differential intracellular distribution of the three mRNAs is correlated with their translation in specific cellular regions where CaM is transiently required. Moreover, treatment of animals with isoproterenol, an agonist of adrenergic- $\beta$  receptors, enhanced the level of *Calm3* mRNA, but not *Calm1* and *Calm2* mRNAs, coinciding with enhanced overlap of this mRNA with *Ryr2* mRNA at the intercalated discs [112]. This is in agreement with the expected increase of  $Ca^{2+}$  mobilization required for heart contraction during sympathetic stimulation, and therefore suggestive of a specialized role of a given CaM gene for particular cell functions.

Saucerman and Bers [113] added further evidence of local regulation of  $\text{Ca}^{2+}$  signaling depending on the presence and properties of  $\text{Ca}^{2+}$ -binding proteins. They hypothesized that based on the known properties of three important cardiac myocyte channels, the L-type calcium channel, the ryanodine, and the  $\text{IP}_3$  receptors, there exists two functional different pools of CaM, one of them containing CaM permanently attached to the channels, called “dedicated” CaM, and another one consisting of “promiscuous” CaM sensing local  $\text{Ca}^{2+}$ . It appears that local  $\text{Ca}^{2+}$  concentrations may be much higher than the overall cytosolic levels, as the affinity of CaM for  $\text{Ca}^{2+}$  binding is quite low ( $K_d$  ca. 5 mM) further indicating that relatively high CaM concentrations possibly in the apo form sequestered by “storage” CaM-binding proteins, such as GAP43/neuromodulin, must be readily available at distinct places in the cell and may be differentially regulated by the 3 CaM genes.

A general concept, as a result of these investigations, has been put forward by Palfi et al. [109] that in the rat brain *CaMI* and *CaMII* are the genes that are mostly regulated in response to a variety of signals and stimuli in contrast to *CaMIII* that is rather considered a housekeeping gene, even though there is no clear distinction and most of the data about differential CaM gene expression derives from neuronal studies. This partial specificity of the expression of the 3 CaM genes enabling the control of local CaM concentration in microdomains, which is important for its availability to interact with targets with high accuracy, may explain at least in part the evolutionary maintenance of the 3 genes.

### 3.2.8. Codon usage affecting RNA stability and translational efficiency

mRNA stability has long been considered as a factor influencing the final protein production from a gene. Cis-regulatory elements in the 3'-UTR, which can bind proteins and miRNAs, and destabilize the mRNA or repress translation, are covered (Section 3.2.6), as features, which separates the different transcripts from human *CALM1,2,3*. Here, we will focus on the codon usage bias (CUB) of the three *CALM* coding regions, and relate these to (predicted) differences in stability. We postulate that the CUBs may have evolved as a means for the cell and organism to regulate the CaM production required depending on time, location, and physiological conditions.

It has been known that the availability of tRNAs has an effect on mRNA stability, but no well-defined way to predict stability from RNA sequences has been described before 2015. This changed when Presnyak et al. [114], studying yeast codon usage, introduced the codon stabilization coefficient (CSC) as a way to give each codon a score, which links the codon usage of mRNA to its half-life. If a codon is overrepresented in stable mRNAs the CSC score is positive, and vice versa the CSC score is negative for codons overrepresented in unstable mRNAs. Several studies have since then expanded on this, so that we now also have CSC scores for zebrafish, *Xenopus* [115], and several human cells types [116,117]. The CSC scores are based on experimental data, in which RNA-seq is done at different time points after transcription was blocked by actinomycin D. Experimental approaches vary, and can relate to endogenous mRNA or to mRNA expressed from lentiviral infection [117].

In animals, CSC appear to be somewhat shared between different species, but not with yeast [118]. In addition, the CSC score is not an entirely stable factor in an organism, one reason being that tRNA levels can vary in different cell types of the same organism [117]. Interestingly, genes with similar function appear to coordinate their mRNA stabilities in part by similar codon usage [119].

It has been proposed that synonymous codons are translated at different speed, and that slowdown of translation can stimulate mRNA decay [116]. This appears to be linked to longer dwell times at the amino acid entry (binding) site of the ribosome during translation. This slowdown is sensed by the decay machinery leading to deadenylation by the CCR4-NOT complex and eventual decapping of the 5'-end. Hia et al. [120] proposed that interleukin enhancer binding factor 2 (ILF2) mediates mRNA stability by binding to transcripts with low GC3 content (that is, a G or C at the wobbler position 3 in a codon), and noted that GC3 codons tend to have higher CSC scores. This observation is our

starting point for comparing the three *CALM* transcripts. Fig. 3ab shows that human *CALM3* has a remarkably higher GC3 content (70%), compared to the 37% in *CALM1* and 30% in *CALM2*. When only considering synonymous codons the values are 33%, 25%, 68% for *CALM1,2,3*, respectively. As discussed above, *CALM1* and *CALM2* appear to be closer related to each other than to *CALM3*, based on similar promoter, enhancer and UTR regions and, as we show here, they are also both predicted to have relatively unstable mRNAs based on their low GC3 content. Next, we applied published CSC scores from eight independent datasets for four human cell lines, from two publications using three different experimental approaches [116,117] (Fig. 5a,b). As shown in Fig. 5a, when averaging the CSC scores from four experiments done in 293T HEK cells, the GC3/AT3 discrimination does indeed appear to be very highly linked to CSC scores, so that there is only a single AT3 codon among the positive CSC codons. Indeed, when calculating the CSC scores for *CALM1,2,3* for all cell types, the same pattern as in Fig. 3ab for GC3 content is seen, so that *CALM3* has a much higher average CSC score, with *CALM2* having the lowest.

In Fig. 5c, we have applied the 293T HEK cell CSC scores along the three transcripts, using the color gradients from Fig. 5a. It seems likely that what is important is not only the average CSC score of a transcript, but also the presence of one or more rows of low scoring codons as seen in *CALM1* and *CALM2*. Hypothetically, a longer row of CSC low scoring codons could slow down translation and trigger the decay machinery, compared to if the low scoring codons are more dispersed. Not only does *CALM3* have a higher average CSC score, but the slow codons are also more dispersed and not clustered as compared to *CALM1* and *CALM2*.

Since all three CaMs are identical in their amino acid sequence, the codon usage is limited. For instance, CaM contains nine methionines, which are coded for by only one codon, thereby fixing the CSC scores at those positions. On the other hand, some amino acids are coded for by six different codons, leaving a wide window of CSC scores to be selected for by evolution. Considering the commonly preferred codon usage in human, the coding region of the 148 amino acids CaM protein would be expected to have a GC3 content of 60% (Table 3). This again points to the most salient observation being the low GC3 content of *CALM1* and *CALM2*.

It can be calculated that any *CALM* mRNA selected for the highest possible CSC score at all sites would have a CSC score of 0.077; on the other hand, any *CALM* mRNA with a minimal stability would have a CSC score of -0.085. The average CSC score for a theoretical *CALM* coding region is 0.007. Considering this, it appears that human *CALM3* is selected for stability, but more evident, that *CALM1* and especially *CALM2* are selected for instability.

In Fig. 5d, we show graphically, by accumulative CSC scores, how the codon stability is distributed along the mRNAs of the three *CALMs*, and for a theoretical *CALM* mRNA, which has CSC scores calculated by the average score of the codons at each AA position, when considering the common human codon usage (see Table 3). *CALM3* has a higher stability than the average, while *CALM1,2* are below, as illustrated by arrows. The *CALM1* mRNA is only slightly unstable in the first half of the protein, after which it is highly unstable. The *CALM2* mRNA is highly unstable throughout the coding region, except a neutral region at the 5' end. Conversely, *CALM3* is stable throughout the coding region, except for the very last part of the 3' end.

Fig. 6 shows the GC3 content of the coding sequence of the three *CALM* genes for eight mammals. They all share the pattern of *CALM3* having the highest GC3 content, and *CALM2* the lowest, except for rat, which has a slightly higher score for *CALM2* than *CALM1*. Mice/rats have the least *CALM* differences in GC3 content of the seven species, given their high *CALM1,2* scores. Cow has the highest scoring *CALM3* gene, and also a relatively high *CALM1* GC3 content, which is even higher than the GC3 content of mice/rat *CALM3*, while human/gorilla has the lowest scoring *CALM2* gene.

Interestingly, it has been reported that the GC3-rich and GC3-poor gene coding regions predict distinct sub-cellular spatial distributions



**Fig. 5.** Comparative stability prediction of the three human *CALM* transcripts. **a.** Codons are separated from high to low CSC score, by using the average of four independent measurements of the 293T HEK cells transcriptome [116,117]. Color gradient from dark green to dark red is used to discriminate the CSC scores, and GCS/AT3 wobbler position is indicated by green or red for each codon. **b.** Data from four different human cell lines and three experimental approaches (Endog, ORF and SLAM). In the top section CSC score data from four cell lines are shown. The middle section shows how the data from transcriptomes correlate between different cell lines, and experimental setups. Data from [116] in dark gray, and [117] in light gray. The bottom section shows statistics when applying the CSC scores on the three *CALM* transcripts. **c.** The color gradients and data from panel a is used to show how predicted high and low stability codons are distributed along the transcripts. **d.** Accumulative CSC scores of *CALM1,2,3* transcripts. The X-axis shows the codon number of the *CALM* coding region and the Y-axis the codon stabilization coefficients for each codon (CSC scores) are accumulatively added. The gray *CALM* is theoretical to how the CSC would be distributed based on the average CSC score for each codon position, based on average human codon usage.

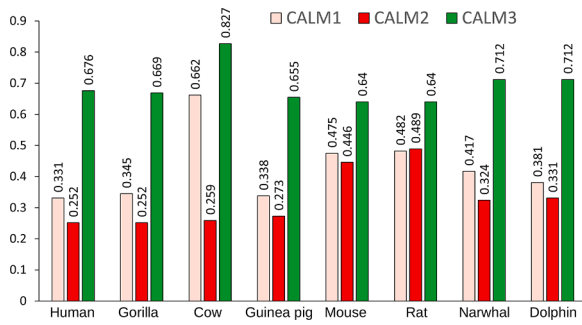
**Table 3**  
Codon usage in *CALM1,2,3*.

Number of amino acids in CaM seq.	Amino acid	Codon	<i>CALM 1</i> codon frequency	<i>CALM 2</i> codon frequency	<i>CALM 3</i> codon frequency	Codon CSC score (293T cells) <sup>a</sup>	Average codon usage <sup>b</sup>	Average amino acid CSC score <sup>c</sup>	Average amino acid GC3 codon content <sup>d</sup>
#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
11	A	gca	5	7	5	-0.0260	0.23	0.0456	0.51
		gcc	1		3	0.1144	0.40		
		gcg			1	0.0381	0.11		
		gct	5	4	2	0.0063	0.26		
0	C	tgc				0.0329	0.55	-0.0308	0.55
		tgt				-0.1085	0.45		
		gac	3	6	7	0.0530	0.54		
17	D	gat	14	11	10	-0.0736	0.46	-0.0053	0.54
		gaa	20	17	2	-0.1250	0.42		
21	E	gag	1	4	19	0.0279	0.58	-0.0363	0.58
		ttc	5	3	5	0.1328	0.55		
8	F	ttt	3	5	3	-0.0622	0.45	0.0451	0.55
		gga	4	3	3	-0.0328	0.25		
11	G	ggc	4	3	3	0.1082	0.34	0.0482	0.6
		ggg	1	1	5	0.0779	0.26		
		ggt	3	4		-0.0038	0.16		
		cac	1		1	-0.0214	0.59		
1	H	cat		1		-0.1517	0.41	-0.0748	0.59
		ata		1		-0.1038	0.16		
8	I	atc	5	1	6	0.1330	0.48	0.0232	0.48
		att	3	6	2	-0.0666	0.36		
		aaa	5	4		-0.1514	0.42		
8	K	aag	3	4	8	-0.0021	0.58	-0.0648	0.58
		cta	3	1		-0.0653	0.07		
9	L	ctc			1	0.1319	0.20	0.0591	0.74
		ctg	2	2	7	0.1609	0.41		
		ctt	1	3		-0.0789	0.13		
		tta	1	2		-0.1365	0.07		
		ttg	2	1	1	-0.0681	0.13		
		atg	9	9	9	0.0687	1		
9	M							0.0687	1
		aac	3	2	3	0.0374	0.54	-0.0321	0.54
6	N	aat	3	4	3	-0.1135	0.46		
		cca	1			-0.0576	0.27	-0.0119	0.44
2	P	ccc	1	1	1	0.0515	0.33		
		ccg		1	1	0.0102	0.11		
		cct				-0.0516	0.28		
		caa	1	3		-0.1257	0.25		
6	Q	cag	5	3	6	-0.0016	0.75	-0.0326	0.75
		aga	2	3	2	-0.1287	0.20		
6	R	agg	1	1	1	-0.0587	0.20	-0.0188	0.6
		cga	1		1	-0.0414	0.11		
		cgc		1		0.0509	0.19		
		cgg				0.0668	0.21		
		cgt	2	1	2	-0.0058	0.08		
		agc			1	0.0119	0.24		
		agt	2	2	1	-0.1506	0.15		
		tca	1	1		-0.1325	0.15		
4	S	tcc	1		2	0.0437	0.22	-0.0512	0.52
		tcg				0.0023	0.06		
		tct		1		-0.1187	0.18		
		aca	6	9	2	-0.0803	0.28		
		acc	3		5	0.0982	0.36		
		acg			1	0.0574	0.12		
7	V	act	3	3	4	-0.1046	0.24	0.0812	0.71
		gta	2	4	2	-0.0724	0.11		
		gtc	4		2	0.1354	0.24		
		gtg	1	2	3	0.1456	0.47		
0	W	gtt		1		-0.0656	0.18	0.0824	1
		tgg				0.0824	1		
2	Y	tac			1	0.0960	0.57	0.0244	0.57
		tat	2	2	1	-0.0707	0.43		

The first three columns show the number of amino acids and codons present in *CALM1,2,3*. Columns #4–6 show codon frequencies in respective genes of codons in column #3. Column #7 (<sup>a</sup>) shows the average CSC score of 4 independent experiments in human 293T cells. Column #8 (<sup>b</sup>) gives average human codon usage for all genes <https://www.genscript.com/tools/codon-frequency-table>. The average amino acid CSC score (<sup>c</sup>) in column #9 is obtained by multiplying the numbers in (<sup>a,b</sup>). Column #10 gives the average GC3 content for each amino acid, by adding GC3 fractions from column #8.

of proteins [121], so that GC3 rich gene products are located more in the cell periphery/membrane/projections, while GC3 poor products are closer to the nucleus. As covered in Section 3.2.7, the *CALM2* mRNA was reported to be located close to the nucleus in human cells, and *CALM3*

mRNA in the periphery [112] thus reflecting what would be predicted based on the GC3 content here, provided the mRNA locations correspond to the local translation locations. However, as also mentioned in 3.2.7, in rats *CALM1,2* mRNAs are reported to be localised more in the



**Fig. 6.** GC3 content in the three *CALM* genes in eight mammals. The GC3 content when only counting synonymous codons is shown by columns and numbers.

neurites (peripheral) compared to *CALM3* mRNA. This apparent difference compared to humans, may reflect the relatively small differences in GC3 content between the 3 rat genes (Fig. 6), giving less locational predictive power.

If these theoretical estimations reflect real properties of the three *CALM* mRNAs, what could be the functional significance for an organism in having two low and one high stable *CALM*? Variation of mRNA stability could give the cell an additional mechanism in fine-tuning gene expression, besides the role of promoters, enhancers, and other UTR features. It could be that having low stability *CALM* coding regions would allow for rapid regulation of CaM production at the transcriptional level, since the overall mRNA level would be mainly regulated by newly transcribed mRNA. Thus, *CALM3* having a more stable transcript, could be necessary when transcription is halted, as occurs during spermatogenesis, when a more stable *CALM3* may be necessary to prevent the sperm cell from running low of *CALM* mRNA. Another factor could be the need for stable *CALM* mRNA in certain cell types, even though transcription is not halted. For instance, translational activity in brain dendrites may need a more stable mRNA, given the relatively long physical distance from the nucleus, making the mRNA more prone to degradation. Epithelial cells might also need more stable mRNAs as a guard against UV light exposure or other mutagens as it is known that chemical damage to mRNA is likely to result in accumulation of aberrant protein products [122].

### 3.2.9. Transcriptome data

Given the apparent dominant negative effects from single allele *CALM* mutations causing heart arrhythmia, it would be desirable to know the relative protein levels produced from each *CALM* gene. This is, however, not a simple task, as the amino acid sequence of CaM produced by 3 independent genes is identical. The relative *CALM1,2,3* expression in adult human heart (left ventricle) was reported, by qPCR analysis, to be approximately 14%, 25% and 61%, respectively [123]. If these numbers reflect the actual protein levels, and every allele is expressed at an equal level, it would mean that 7% mutated *CALM1* (one of six alleles) is sufficient to cause heart arrhythmia (see also part 4.1). In this section we will review human public RNA-seq data from high-throughput studies where the expression levels of large numbers of genes were measured simultaneously. We will first cover data from bulk tissue gene expression (Gtxportal.org) and then data from single cells (scrRNA, proteinatlas.org).

Average *CALM1,2,3* expression from 53 tissues is 302, 661 and 329 TPM (transcripts per million), respectively (Fig. 7a). All three genes show the highest expression across thirteen brain tissues with averages of 686, 1012 and 692 TPM (29%, 42%, 29%). The overall expression in testis is, on average, 490 TPM, but here the relative contribution from *CALM3* is higher than normal with 34% (506 TPM), and only 8% (124 TPM) from *CALM1*. Average data from two heart tissues (atrium and left ventricle) are 113, 220 and 136 TPM (24%, 47%, 29%), and in all three genes heart tissue is among the tissues with the lowest expression having



**Fig. 7.** a. Expression data for *CALM1,2,3* from 53 human tissues. Y-axes show transcripts per million (TPM) for a given tissue. (Source, <https://www.gtxportal.org/>). b. Expression data from 76 single cell types. Axes shows normalized transcripts per million (nTPM) for a given tissue. (source, <https://www.proteinatlas.org/>).

an average of 156 TPM.

Bulk RNAseq measures the average expression of genes from multiple cells, but is limiting when one wishes to understand gene expression patterns within the cell, and among subpopulations of cells, like for instance cardiomyocytes in heart tissue and the spermatocytes in testis. Looking across 76 single cell types (Fig. 7b), the overall rank between the three *CALMs* is not changed compared to bulk tissue with averages of 596, 943 and 218 (34%, 54%, 12%), respectively, but the contribution from *CALM3* is lower compared to tissue samples. In neuronal cells the *CALM1* contribution is unusually high with 58% (567 nTPM, N normalized transcripts per million), while *CALM3* is relatively low with 9% (86 nTPM). Interestingly, when looking at cardiomyocytes, the mRNA load from *CALM1* is also the highest (868 nTPM, 51%), and here the *CALM3* contribution is only 11% (178 nTPM).

However, the most intriguing observation from these scRNA data is the *CALM3* contribution of germ cells. As mentioned above, *CALM3* is upregulated in testis tissue (34%), but when looking at certain single germ cell types, *CALM3* is dramatically overexpressed. Across 4 different germ cells, the *CALM3* contribution is 58% (1393 nTPM), and in early and especially late spermatids the contribution is even higher with 80% (1864 nTPM) and 87% (2697 nTPM), respectively. Since transcription can be expected to be low, or absent at late stages of spermatogenesis, it can be speculated that *CALM3* given its theoretically higher stability, becomes more important for the organism than *CALM1,2* which would degrade faster due to a lower mRNA half-life.

### 3.3. *CaM gene transcription in pathological conditions*

The three human CaM genes have been shown to be expressed at different levels in cancer cells. For example, *CALM3* presents higher expression than *CALM1* and *CALM2* in lymphoblastic cells [124]. Also, the expression of the three CaM genes tested in human teratoma cells follows the sequence *CALM3* >> *CALM2* = *CALM1* [51].

*CALM1* and the EGFR gene are overexpressed in esophageal squamous cell carcinoma. Knocking-down *CALM1* by CRIPR-Cas9 and inhibiting EGFR activation with Afatinib in these tumor cells synergistically inhibited cell proliferation, migration and invasion *in vitro* and *in vivo* and also induced apoptosis [90]. The observed synergy may be due, at least in part, to the stimulatory effect that CaM exerts on EGFR activation [125]. Likewise, in hepatocellular carcinoma the downregulation of *CALM2* using siRNA induced the inhibition of cell migration and proliferation, inhibiting as well the formation of colonies and the invasive capacity of these tumor cells via the induction of apoptosis [91]. Increased expression of the *CALM1* gene has been shown to occur as well in gastric adenocarcinomas, as compared to normal peritumoral tissue, and this gene was considered a potential inflammation-related target to treat this type of tumors [126]. Also, increased expression of the *CALM2* gene has been detected in a series of cell lines derived from anaplastic large cell lymphomas, but to a highly variable extent [127].

*CALM1* is upregulated about 5-fold in ovary tissues of patients with polycystic ovary syndrome [128]. In a transgenic mouse model of ALS, it was found that, among other genes, the expression of *CALM1* in skeletal muscle during the early symptomatic stage and the progression of the disease negatively correlates with longevity, despite the somewhat irregular expression pattern of this gene its expression increased during the terminal stage of the disease [129].

Finally, intracerebral administration of antisense oligodeoxynucleotides to downregulate receptor-5 for neuropeptide Y within the lateral ventricles of the brain with the goal to treat obesity by decreasing food intake in obese rats, induced the downregulation of *CALM2* in adipose tissues, among other genes regulating Ca<sup>2+</sup> signaling [130].

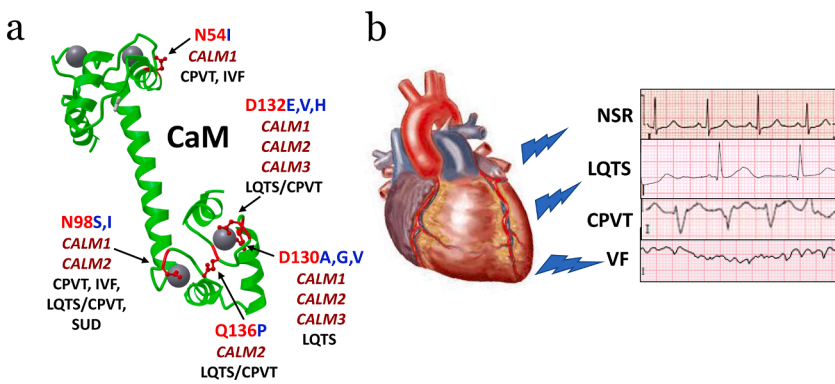
## 4. Human CaM mutations and gene polymorphisms

### 4.1. *CaM mutations leading to heart arrhythmias and polymorphisms associated with cardiovascular disease*

A surprising discovery was made and published in 2012, as for the first time mutations in human CaM were found to be linked to life-threatening arrhythmias [131]. As CaM is an essential and extremely conserved protein taking a central role as a calcium sensor in many basic cell-signaling processes [5], it was earlier assumed that mutations in *CALM* genes would not be tolerated and propagated. Two mutations in *CALM1*, one at position N54I (this numbering includes the start Met), inherited in a large Swedish family and characterized by repeated events of syncope and heart stop and another *de novo* mutation at position N98S in a person from Iraq were identified. Both mutations were found to cause catecholaminergic polymorphic ventricular tachycardia (CPVT), which is a severe arrhythmia known to cause sudden cardiac arrest and death during exercise or emotional stress without previously known heart defect. This was followed by the finding of further CaM mutations affecting the *CALM1* and *CALM2* genes causing long-QT syndrome (LQTS), which affected the patients in a similar way, including strongly enhanced QT episodes of T wave alternans and heart stop in the first year of life [123]. A number of additional CaM mutations were then reported to cause the mentioned syndromes and also idiopathic ventricular fibrillation (IVF) (reviewed in [35,132–135]). Based on these findings, pathologies caused by *CALM* mutations were termed calmodulinopathy in the International Calmodulinopathy Registry by Schwartz and collaborators [132]. As reported by the registry in 2019, 74 patients from 51 families with *CALM* mutations had been identified with a distribution among the three *CALM* genes of 36/23/15 (rel. percentages of 49%/31%/20%) in *CALM1,2,3*, respectively [132]. Most of these rare mutations were life-threatening and found to be *de novo* mutations causing early onset of adrenergically induced arrhythmia. The majority of these mutations caused LQTS, and interestingly the same mutations had different outcomes in different patients.

CaM mutations causing heart arrhythmia were to a major degree localized to the C-terminal half of CaM at residues involved in calcium coordination as part of the calcium binding domains EF-hand III and IV (see Fig. 8). Some of these mutations were found in more than one patient with the most prevalent ones being (N98, D130, D132, and F142). These residues were altered to different amino acids in different patients. It was noted that even considering the relatively small number of patients with arrhythmogenic CaM mutations, most *CALM*-LQTS mutations were found in EF-hand IV replacing evolutionary highly conserved residues important for Ca<sup>2+</sup> binding. On the other hand, *CALM*-CPVT mutations were located in EF-hand III or in the EF-hand I-II linker. Interestingly, some mutations including N98S, D132E and Q136P caused both CVT and LQTS (reviewed in [35,132]).

CaM has many binding partners, interacting with them in a huge number of different ways, working as linker and/or adaptor protein due to its structural flexibility [8]. In the heart, several CaM targets have been shown to be affected by CaM mutations. Most prominent are the ryanodine receptor 2 (RyR2) governing Ca<sup>2+</sup> release from the sarcoplasmic reticulum (reviewed in [136]), and the L-type Ca<sup>2+</sup> channel, responsible for voltage-gated Ca<sup>2+</sup> entry into the sarcoplasm of the cardiomyocyte (reviewed in [137]). As calcium signaling is the main small signaling molecule controlling and triggering cardiomyocyte contraction and relaxation, unbalanced calcium handling is an important reason for heart failure. Since CaM is an essential player in calcium signaling of all eukaryotic cells and has many important targets outside the heart, one can only speculate why CaM mutations primarily affect the heart and no other organ functions. It can be speculated, given the low overall expression of CaM and the high abundance of important CaM binding proteins in the heart, the pool of free CaM might be very low, and therefore the function of mutated CaM becomes more pronounced, since it cannot be substituted by free WT CaM. However, there



**Fig. 8.** Arrhythmogenic effects of CaM mutants. a. Human CaM (PDB ID 1CLL) [170] showing some residues numbered (highlighted in red) which have been shown to be mutated (red  $\rightarrow$  blue) leading to arrhythmias [35,135]. Amino acid residues are numbered considering the N-terminus methionine, absent in the mature protein. The gray spheres are  $\text{Ca}^{2+}$  ions. The affected CaM genes and the mutation-induced arrhythmias are indicated: CPVT, catecholaminergic polymorphic ventricular tachycardia; IVF, idiopathic ventricular fibrillation; LQTS, long-QT syndrome; LQTS/CPVT, overlapping LQTS and CPVT features; SUD, sudden unexplained death. For clarity and to prevent overlapping only one mutation located in the N-lobe and four mutations in the C-lobe of CaM of those mentioned in the text are shown. Additional arrhythmogenic mutations have been reviewed [35,135]. b. Deregulation of intracellular  $\text{Ca}^{2+}$  due to RyR2 and Cav1.2 dysfunction causes heart arrhythmias [171–174]. Electrocardiogram traces taken in derivation I showing a normal sinus rhythm (NSR), and traces from patients

suffering LQTS, CPVT, and ventricular fibrillation (VF) leading to sudden death, are presented (adapted from reference [175]).

are some reports suggesting that *CALM* gene mutations may be causing other pathologies (see Sections 4.2–4.4).

Interestingly, it was found that two newly discovered *CALM* mutations, causing heart failure, were present only in some tissues (genetic mosaicism) of parents from affected individuals [138]. This was documented in one case with a mutation in *CALM3*-E41K with 25% of sequencing reads and in another case with *CALM3*-D130G with 6% mutated alleles, in the peripheral blood of the mother in the first case and the father in the second case. This indicates that germline parental mosaicism may explain unexpected fetal arrhythmia in some children of the same parents and multiple intrauterine fetal demise observed in one case. This finding will be of potential relevance for consulting families with CaM mutation causing heart disorders.

Limited information is available about genetic polymorphisms in the three *CALM* genes that are correlated with cardiovascular disease [139]. The polymorphisms *rs3814843* and *rs3179089* in the *CALM1* 3'-UTR of the gene (CC and GG genotypes, respectively) have been associated with enhanced risk of ischemic stroke, particularly the *rs3179089* polymorphism in females [140,141], and have also been found to be a potential marker for sudden cardiac death [142].

In addition, a polymorphism of unassigned number, in the *CALM3* promoter region at position  $-34\text{T}>\text{A}$ , directly adjacent to the 5'-end of a putative SP1 transcription factor binding site, was found to affect the level of transcription of this gene, with a lower transcriptional activity of the T-*CALM3* version in a promoter analysis in HEK293 cells and cardiomyocytes compared to the wild-type and to exhibit a clearly higher frequency in patients with familiar hypertrophic cardiomyopathy (FHC) compared to controls [143]. According to this, a small decrease in wild-type CaM would increase the probability of causing FHC.

How can a mutation in one out of 6 alleles of the *CALM* genes cause severe and life-threatening heart diseases? It is known that mutant CaM can pre-associate at low calcium with the two most prominent CaM targets in cardiomyocytes, the  $\text{Ca}_v1.2$  and RyR2 channels which regulate  $\text{Ca}^{2+}$  influx into the sarcoplasm and therefore may compete with wild-type CaM in closing the channels after calcium is increased, as shown for the  $\text{Ca}_v1.2$  channel which has an IQ-motif CaM binding site (reviewed in [133–135]). Therefore, it was proposed that mutations affecting calcium binding (the majority of all known arrhythmia causing mutations) would mostly affect the L-type calcium channels leading to LQTS and, on the other hand, mutations with minor effect on calcium binding (N98S and N54I) would rather interfere with the more complex situation in the regulation of the RyR2 channel, which could lead to CPVT. In the case of the RyR2, one explanation of the reported dominant negative effect of *CALM* mutations (e.g. of CaM F90L causing IVF and sudden heart arrest [144]) could be based on the fact that RyR2 consists of 4 subunits each of which have a CaM binding domain. Binding of only one mutated CaM to the complex could affect the other subunits and

interferes with the known clustering of RyR2 into 2D oligomers. A similar situation of dominant negative effect of *CALM* mutations has been demonstrated for another complex CaM target, the CaMK-II, which plays as well an important role in heart physiology, and consists of 12 or 14 subunits, which interact with each other depending on the  $\text{Ca}^{2+}/\text{CaM}$  binding [38,145]. Interestingly, CaMK-II regulates both the L-type calcium channel  $\text{Ca}_v1.2$  and the RyR2 by phosphorylation, thus facilitating channel opening (reviewed in [125,126]). The effect of individual CaM mutations occurring in one out of 6 alleles is of course dependent on the relative expression of the three genes. Measuring mRNA levels of the three genes in cardiac tissue has been done by several groups and is so far inconclusive as largely divergent results were obtained (reviewed in [35]). It cannot be ruled out that these divergent data are the result of the theoretically large differences in mRNA stability of three *CALMs*, resulting in higher lab to lab variations than normal. As covered in Section 3.2.9, *CALM2* appears to be the generally highest expressed *CALM*, also in heart tissue, with 47% contribution. The codon stability data covered in Section 3.2.8 suggest that *CALM1,2* produce unusually unstable mRNAs, while *CALM3* is predicted to have high stability. It can be imagined how this may result in higher degradation of *CALM1,2* compared to *CALM3* during sample preparation, and so give a general trend towards lab results overestimating *CALM3* contribution. Unpublished data from our lab point into this direction. When we analyzed relative expression from the three *CALM* genes, we have observed striking differences comparing fresh and old tissue samples. qPCR results from cells and fresh tissue showed *CALM3* being the least expressed, but in old rat brain tissue, *CALM1,2* were virtually absent, whereas *CALM3* seemed in the normal range, resulting in a relative *CALM3* expression of 98% and 100%, respectively from two independent measurements. Also, other sources report that the *CALM1,2* transcripts are especially unstable at room temperature [146]. *CALM1,2* were found to correlate with RNA integrity, so that when the RNA quality is low, the relative expression of *CALM1,2* decreased, indicating that *CALM1,2* are more prone to RNA damage compared to genes producing more stable RNA.

The long 4.1 kb mRNA isoform of *CALM1* does appear to be relatively overexpressed compared to the short form in heart left ventricular tissue (no scRNA data for the two forms in cardiomyocytes are available, see Fig. 4). But another, perhaps, more important point when considering mRNA expression in relation to calmodulinopathy, is the nature of the information given when sampling a bulk heart tissue, since cardiomyocytes are estimated to account for only 30% of the 4 cell types in the human heart [147]. As covered in Section 3.2.9, cardiomyocytes appear to express only 11% of the mRNA pool from *CALM3* and 51% from *CALM1*. How these numbers are translated into actual relative protein amounts is not known. As stated above, most patients with mutations in *CALM1* and fewest in *CALM3* are registered, possibly reflecting the cardiomyocyte expression levels of the three genes,

perhaps because a disease variant with higher expression is more likely to lead to severe symptoms. However, due to the small number of patients with CaM mutations this remains speculative.

Even if a clearer picture of mRNA levels will arrive, it is still an open question whether this would relate to the level of proteins, as mutant CaM may differ from the wild-type one in stability, subcellular location, target specificity, and might be post-translationally modified (e.g. phosphorylation, methylation, acetylation, oxidation) in different ways.

Even though a potential connection between the type of *CALM* mutations and the caused pathology has emerged, as outlined above, a clear correlation between phenotype and genetic alteration or properties of the resulting proteins has not been established yet. As CaM has many important targets in cardiomyocytes, it is likely that the pathophysiological effect of mutant CaM is the result of improper regulation of several critical players.

#### 4.2. CaM mutations potentially leading to neural diseases

Ion channels are essential for neural function, and CaM has an important role in the regulation of many of these ion channels, including Ca<sup>2+</sup>-dependent L-type Ca<sup>2+</sup> channels [148] (reviewed in [134,149]), voltage-gated Na<sup>+</sup> channels [150] (reviewed in [134,151]), voltage-gated K<sup>+</sup> channels (reviewed in [134]), Ca<sup>2+</sup>-activated K<sup>+</sup> channels (reviewed in [152]), among others (reviewed in [35,134,149,152]) Additionally, other CaM-dependent systems in neurons and/or glial cells, including the important signaling molecules CaMK-II and calcineurin, could also be affected likewise when CaM mutations occur and contribute to neural pathogenesis.

Patients with *CALM1* and *CALM2* (*CALM3* have 29% average relative expression in brain tissue, but only 8% across 11 glial and neuronal single cell RNA measurements) mutations affecting residues at EF-hands III and IV, showing decreased Ca<sup>2+</sup> affinity, exhibited in addition to cardiac arrhythmias as described above, neurodevelopmental delay, epilepsy, cognitive, and motor deficiencies. This was first ascribed to be secondary to brain injury after transient anoxic episodes, resulting from arrhythmias leading to syncope and cardiac arrest during early life, that was successfully resuscitated [123]. However, the authors of this report could not exclude primary neuronal affection associated to circulatory insufficiency. The same group hypothesized thereafter that perhaps only in cases where multiple cardiac arrests occurred, post-anoxic neural sequelae remain [153]. Nevertheless, this hypothesis was based considering only few cases with divergent neural clinical features, ranging from mental retardation, autism, developmental delay and seizures, and without history of cardiac arrest (as recorded in the International Calmodulinopathy Registry) [132,135,154]. This suggests, not unexpectedly, that CaM mutations may directly affect brain functions.

#### 4.3. CaM mutations potentially leading to endocrine-associated illness

A female patient affected with sinus bradycardia presenting heterozygous mutation of *CALM3* (E141K) suffered as well a serious hypoglycemic condition, that was the ultimate cause of her death, and this was attributed to hyperinsulinemia [138]. The authors of this study suggested that this CaM mutation could be responsible for the dysregulated insulin secretion by affecting L-type Ca<sup>2+</sup> channels in pancreatic  $\beta$ -cells [138]. This report opens the possibility that CaM mutations could cause or contribute to some endocrine disorders.

#### 4.4. CaM polymorphisms potentially associated with osteoarthropathies

Some single nucleotide polymorphisms (SNP) in *CALM1* have been associated with osteopathies. Mototani et al. [155] reported that a SNP in intron 3 (IVS3 -293C>T), as well as another SNP in the core promoter region of *CALM1* (*rs12885713* -16C>T) strongly correlated with hip osteoarthritis. Particularly, these authors claimed that the latter SNP decreased the rate of transcription of this gene, which lowers the CaM

level and reduces the expression of other genes essential for cartilage matrix formation. These include *COL2A1*, encoding the pro- $\alpha$ 1 chain of type II collagen [156], and *AGC1*, encoding chondroitin sulfate proteoglycan 1, also denoted aggrecan [157]. The study of Mototani et al. [155] was performed with Japanese patients. However, analogue studies with British and Greek Caucasian patients did not find a correlation between the occurrence of polymorphisms in the core promoter region of *CALM1* and susceptibility to hip or knee osteoarthritis [158,159]. Hoaglund [160], commented that this apparent inconsistency could be explained by the different underlying pathology in the two patient groups, being mostly primary or idiopathic osteoarthritis in the Caucasian group, while being congenital hip dislocations or developmental acetabular dysplasia in the Japanese group, therefore giving validity to both studies. Nevertheless, two SNPs (*i-CALM2-5 rs10153674* A>G; and *i-CALM2-6* novel G>A) both located in intron 1 of *CALM2* were significantly associated with hip osteoarthritis without acetabular dysplasia in the Japanese patients [161]. The authors of this study pointed out that this was observed only after stratification of the patients according to their acetabular dysplasia status, otherwise no significant correlation was found. In addition, SNP *rs12885713* -16C>T occurrence in the core promoter region of *CALM1*, and SNP *rs5871* located in the 3'-UTR, were associated with idiopathic scoliosis in adolescents Chinese patients [162]. Nevertheless, the genetic basis of these pathological conditions are multifactorial and many SNPs associated with idiopathic scoliosis have been found in other genes (reviewed in [163]). Furthermore, the association of SNP *rs12885713* (-16C>T) in the core promoter region of *CALM1* with osteoarthritis predisposition has been further questioned in a study carried out in Han Chinese patients and in a meta-analysis pooling Asian and Caucasian patients [164,165]. These findings suggest that ethnicity may play a role in the pathogenesis of these osteoarthropathies, although further studies are required to clarify this issue.

### 5. Concluding remarks

Based on a multiplicity of discoveries described earlier in [108], and extended in this article, it can be concluded that the three CaM genes in mammals cannot solely function as redundant copies replacing each other in case of need. The following issues would argue against this possibility:

- (1) Non-coding regions including promoter/enhancers 5'- and 3'- non-coding mRNA sequences, may lead to different stabilities of the RNA, e.g., due to target sequences of proteins and miRNA. The importance of these sequences is highlighted by the fact that they are almost as highly conserved among the corresponding genes in different vertebrates as the coding sequences.
- (2) Every individual CaM gene has highly conserved non-coding sequences, when different vertebrate species are compared. The conservation of these sequences among genes in the same species is lower, allowing the synthesis of mRNAs in different cell types with different stabilities and subcellular localizations. This can be responsible, at least in part, of multitude modes of fine-tuning regulation.
- (3) Another factor not previously focused on, but highlighted in this article, is the potential significance of differential codon usage by the 3 CaM genes, which may lead to different stabilities and translational efficiency. It is captivating to observe how the three *CALMs* have differentiated in their GC3 usage, considering that they originated from one single *CALM* gene by duplication. The pattern of two low GC3 *CALM1,2* genes, and one high GC3 *CALM3* gene in humans is also found in seven other mammals investigated. Especially, *CALM2* appears to be highly selected for instability throughout its coding region. CaM is a very important protein, as highlighted by its many interacting partners, and uniquely 100% sequence conservation in all investigated



vertebrates. Higher organisms may have evolved to take advantage of three *CALM* genes with different mRNA stabilities. A low stability mRNA can be in some cell types where fast regulation of the available mRNA pool is especially advantageous, but can be a trade-off in other cell types. With three *CALM* genes evolution may have more freely tailored the three mRNAs. Following this hypothesis, it is interesting to observe how the high stability *CALM3* is the main mRNA in late spermatids, where transcription is very low, and high stability is therefore needed. The organism ability to use an available stable *CALM3* mRNA in spermatids may then have allowed the *CALM1,2* mRNAs to evolve in a less stable direction, without evolutionary pressure selecting in opposite direction.

- (4) Every individual CaM gene seems to be irreplaceable, as deduced from pathological conditions in human. First, mutations in a single CaM gene can lead to illness, despite co-expression of non-mutated paralogs. Second, the specific expression patterns of CaM genes change in cancer cells.

It would be interesting to know whether different subcellular pools of CaM created by the presence of specific mRNA pools would be prone to different post-translational modifications, such as phosphorylation, methylation, carboxymethylation or acetylation [13,166–169]. It can be imagined that the cellular location where the mRNA is translated, and in addition the speed of the translation machinery could result in different modifications of *CALM1,2,3* products. Codon usage predict faster translation of the *CALM3* compared to the other CaM genes. This could be relevant to determine the contribution of a specific CaM gene mutation or an isoform for the development of a pathology.

The advanced gene editing methods, using CRISPR technology, would allow to slightly modify CaM genes, for example adding different tags. This can be the basis for gene discrimination in mammals in order to study the differential expression of CaM from the 3 genes. However, this may of course be difficult, as the modification of CaM, even if minor, could lead to changes in function.

## Data Availability

Data will be made available on request.

## Acknowledgments

This work was funded by grants from the Danish Research Council (DFR-4004-00560) the Wedell Wedellsborg Foundation and the Candys Foundation, Denmark, to MWB; grants from the program *Consolidación de Grupos de la Junta de Andalucía* (research group BIO-210) and from the Plan Estatal 2017–2020 (PID2019-103850GB-I00), Spain, to EV; and grant from the *Consejería de Educación, Juventud y Deportes – Comunidad de Madrid* (B2017/ BMD-36), Spain, involving contributions from the European Funds for Regional Development and the Social European Fund, to AV.

## References

- [1] E. Carafoli, J. Krebs, Why calcium? How calcium became the best communicator, *J. Biol. Chem.* 291 (2016) 20849–20857.
- [2] R.H. Kretsinger, E. Carafoli, F. Clementi, W. Drabikowski, A. Margreth, Hypothesis: calcium modulated proteins contain EF hands. *Calcium Transport, Contraction, Secretion*, Elsevier, Amsterdam, 1975.
- [3] D. Chin, A.R. Means, Calmodulin: a prototypical calcium sensor, *Trends Cell Biol.* 10 (2000) 322–328.
- [4] H. Tidow, P. Nissen, Structural diversity of calmodulin binding to its target sites, *FEBS J.* 280 (2013) 5551–5565.
- [5] M.W. Berchtold, A. Villalobo, The many faces of calmodulin in cell proliferation, programmed cell death, autophagy, and cancer, *Biochim. Biophys. Acta* 1843 (2014) 398–435.
- [6] A. Villarroel, M. Tagliatalata, G. Bernardo-Seisdedos, A. Alaimo, J. Agirre, A. Alberdi, C. Gomis-Perez, M.V. Soldovieri, P. Ambrosino, C. Malo, P. Areso, The ever changing moods of calmodulin: how structural plasticity entails transductional adaptability, *J. Mol. Biol.* 426 (2014) 2717–2735.
- [7] A. Villalobo, M.W. Berchtold, The role of calmodulin in tumor cell migration, invasiveness, and metastasis, *Int. J. Mol. Sci.* 21 (2020) 765.
- [8] A. Villalobo, H. Ishida, H.J. Vogel, M.W. Berchtold, Calmodulin as a protein linker and a regulator of adaptor/scaffold proteins, *Biochim. Biophys. Acta Mol. Cell Res.* 1865 (2018) 507–521.
- [9] A.P. Yamniuk, M. Rainaldi, H.J. Vogel, Calmodulin has the potential to function as a Ca<sup>2+</sup>-dependent adaptor protein, *Plant Signal. Behav.* 2 (2007) 354–357.
- [10] L.A. Jurado, P.S. Chockalingam, H.W. Jarrett, Apocalmodulin, *Physiol. Rev.* 79 (1999) 661–682.
- [11] M. Bahler, A. Rhoads, Calmodulin signaling via the IQ motif, *FEBS Lett.* 513 (2002) 107–113.
- [12] S.L. Toutenhoofd, E.E. Strehler, The calmodulin multigene family as a unique case of genetic redundancy: multiple levels of regulation to provide spatial and temporal control of calmodulin pools? *Cell Calcium* 28 (2000) 83–96.
- [13] G. Benaim, A. Villalobo, Phosphorylation of calmodulin. Functional implications, *Eur. J. Biochem.* 269 (2002) 3619–3631.
- [14] A. Villalobo, The multifunctional role of phospho-calmodulin in pathophysiological processes, *Biochem. J.* 475 (2018) 4011–4023.
- [15] T.N. Davis, M.S. Urdea, F.R. Masiarz, J. Thorner, Isolation of the yeast calmodulin gene: calmodulin is an essential protein, *Cell* 47 (1986) 423–431.
- [16] T. Takeda, M. Yamamoto, Analysis and *in vivo* disruption of the gene coding for calmodulin in *Schizosaccharomyces pombe*, *Proc. Natl. Acad. Sci. USA* 84 (1987) 3580–3584.
- [17] C.D. Rasmussen, R.L. Means, K.P. Lu, G.S. May, A.R. Means, Characterization and expression of the unique calmodulin gene of *Aspergillus nidulans*, *J. Biol. Chem.* 265 (1990) 13767–13775.
- [18] R.C. Simao, S.L. Gomes, Structure, expression, and functional analysis of the gene coding for calmodulin in the chytridiomycete *Blastocladiella emersonii*, *J. Bacteriol.* 183 (2001) 2280–2288.
- [19] H.B. LeJohn, Structure and expression of fungal calmodulin gene, *J. Biol. Chem.* 264 (1989) 19366–19372.
- [20] D.H. O'Day, R.J. Taylor, M.A. Myre, Calmodulin and calmodulin binding proteins in dictyostelium: a primer, *Int. J. Mol. Sci.* 21 (2020) 1210. Art. No.
- [21] K.J. Robson, M.W. Jennings, The structure of the calmodulin gene of *Plasmodium falciparum*, *Mol. Biochem. Parasitol.* 46 (1991) 19–34.
- [22] M. Katoh, M. Hirono, T. Takemasa, M. Kimura, Y. Watanabe, A micronucleus-specific sequence exists in the 5'-upstream region of calmodulin gene in *tetrahymena thermophila*, *Nucleic Acids Res.* 21 (1993) 2409–2414.
- [23] W.E. Zimmer, J.A. Schloss, C.D. Silflow, J. Youngblom, D.M. Watterson, Structural organization, DNA sequence, and expression of the calmodulin gene, *J. Biol. Chem.* 263 (1988) 19370–19383.
- [24] S.H. Chung, J. Swindle, Linkage of the calmodulin and ubiquitin loci in *Trypanosoma cruzi*, *Nucleic Acids Res.* 18 (1990) 4561–4569.
- [25] C. Tschudi, A.S. Young, L. Ruben, C.L. Patton, F.F. Richards, Calmodulin genes in trypanosomes are tandemly repeated and produce multiple mRNAs with a common 5' leader sequence, *Proc. Natl. Acad. Sci. USA* 82 (1985) 3998–4002.
- [26] T.N. Davis, J. Thorner, Vertebrate and yeast calmodulin, despite significant sequence divergence, are functionally interchangeable, *Proc. Natl. Acad. Sci. USA* 86 (1989) 7909–7913.
- [27] A. Perochon, D. Aldon, J.P. Galaud, B. Ranty, Calmodulin and calmodulin-like proteins in plant calcium signaling, *Biochimie* 93 (2011) 2048–2053.
- [28] X. Zhu, C. Dunand, W. Snedden, J.P. Galaud, CaM and CML emergence in the green lineage, *Trends Plant Sci.* 20 (2015) 483–489.
- [29] V. La Verde, P. Dominici, A. Astegno, Towards understanding plant calcium signaling through calmodulin-like proteins: a biochemical and structural perspective, *Int. J. Mol. Sci.* 19 (2018) 1331.
- [30] R. Fischer, M. Koller, M. Flura, S. Mathews, M.A. Strehler-Page, J. Krebs, J. T. Penniston, E. Carafoli, E.E. Strehler, Multiple divergent mRNAs code for a single human calmodulin, *J. Biol. Chem.* 263 (1988) 17055–17062.
- [31] M.W. Berchtold, R. Egli, J.A. Rhyner, H. Hameister, E.E. Strehler, Localization of the human bona fide calmodulin genes *CALM1*, *CALM2*, and *CALM3* to chromosomes 14q24-q31, 2p21.1-p21.3, and 19q13.2-q13.3, *Genomics* 16 (1993) 461–465.
- [32] Q. Ye, M.W. Berchtold, Structure and expression of the chicken calmodulin I gene, *Gene* 194 (1997) 63–68 [Corrigendum *Gene*, 213 (1998) 1225].
- [33] D.J. Mangelsdorf, B.S. Komm, D.P. McDonnell, J.W. Pike, M.R. Haussler, Immunoselection of cDNAs to avian intestinal calcium binding protein 28 K and a novel calmodulin-like protein: assessment of mRNA regulation by the vitamin D hormone, *Biochemistry* 26 (1987) 8332–8338.
- [34] B. Kovacs, K. Gulya, Differential expression of multiple calmodulin genes in cells of the white matter of the rat spinal cord, *Brain Res. Mol. Brain Res.* 102 (2002) 28–34.
- [35] H.H. Jensen, M. Brohus, M. Nyegaard, M.T. Overgaard, Human calmodulin mutations, *Front. Mol. Neurosci.* 11 (2018) 396.
- [36] H.A. Ledford, S. Park, D. Muir, R.L. Woltz, L. Ren, P.T. Nguyen, P. Sirish, W. Wang, C.R. Sihn, A.L. George, B.C. Knollmann, E.N. Yamoah, V. Yarov-Yarovoy, X.D. Zhang, N. Chiamvimonvat, Different arrhythmia-associated calmodulin mutations have distinct effects on cardiac SK channel regulation, *J. Gen. Physiol.* 52 (2020), e202012667. Art. No.
- [37] W.B. Limpitkul, I.E. Dick, R. Joshi-Mukherjee, M.T. Overgaard, A.L. George Jr., D.T. Yue, Calmodulin mutations associated with long QT syndrome prevent inactivation of cardiac L-type Ca(2+) currents and promote proarrhythmic behavior in ventricular myocytes, *J. Mol. Cell Cardiol.* 74 (2014) 115–124.

- [38] M.W. Berchtold, T. Zacharias, K. Kulej, K. Wang, R. Torggler, T. Jespersen, J. N. Chen, M.R. Larsen, J.M. la Cour, The arrhythmogenic calmodulin mutation D129G dysregulates cell growth, calmodulin-dependent kinase II activity, and cardiac function in zebrafish, *J. Biol. Chem.* 291 (2016) 26636–26646.
- [39] R.H. Kretsinger, C.E. Nockolds, Carp muscle calcium-binding protein. II. Structure determination and general description, *J. Biol. Chem.* 248 (1973) 3313–3326.
- [40] N.C. Strynadka, M.N. James, Crystal structures of the helix-loop-helix calcium-binding proteins, *Annu. Rev. Biochem.* 58 (1989) 951–998.
- [41] D.C. Dominguez, M. Guragain, M. Patrauchan, Calcium binding proteins and calcium signaling in prokaryotes, *Cell Calcium* 57 (2015) 151–165.
- [42] N.D. Moncrief, R.H. Kretsinger, M. Goodman, Evolution of EF-hand calcium-modulated proteins. I. Relationships based on amino acid sequences, *J. Mol. Evol.* 30 (1990) 522–562.
- [43] S. Nakayama, N.D. Moncrief, R.H. Kretsinger, Evolution of EF-hand calcium-modulated proteins. II. Domains of several subfamilies have diverse evolutionary histories, *J. Mol. Evol.* 34 (1992) 416–448.
- [44] S. Nakayama, R.H. Kretsinger, Evolution of EF-hand calcium-modulated proteins. III. Exon sequences confirm most dendrograms based on protein sequences: calmodulin dendrograms show significant lack of parallelism, *J. Mol. Evol.* 36 (1993) 458–476.
- [45] R.H. Kretsinger, S. Nakayama, Evolution of EF-hand calcium-modulated proteins. IV. Exon shuffling did not determine the domain compositions of EF-hand proteins, *J. Mol. Evol.* 36 (1993) 477–488.
- [46] F. Friedberg, A.R. Rhoads, Evolutionary aspects of calmodulin, *IUBMB Life* 51 (2001) 215–221.
- [47] NCBI, *CALM1* calmodulin 1 [Homo sapiens (human)] <https://www.ncbi.nlm.nih.gov/gene/801>.
- [48] NCBI, *CALM3* calmodulin 3 [Homo sapiens (human)] <https://www.ncbi.nlm.nih.gov/gene/808>.
- [49] NCBI, *CALM2* calmodulin 2 [Homo sapiens (human)] <https://www.ncbi.nlm.nih.gov/gene/805>.
- [50] Ensemble Database, <https://www.ensembl.org>.
- [51] S.L. Toutenhoofd, D. Foletti, R. Wicki, J.A. Rhyner, F. Garcia, R. Tolon, E. E. Strehler, Characterization of the human *CALM2* calmodulin gene and comparison of the transcriptional activity of *CALM1*, *CALM2* and *CALM3*, *Cell Calcium* 23 (1998) 323–338.
- [52] J.A. Rhyner, M. Ottiger, R. Wicki, T.M. Greenwood, E.E. Strehler, Structure of the human *CALM1* calmodulin gene and identification of two *CALM1*-related pseudogenes *CALMIP1* and *CALMIP2*, *Eur. J. Biochem.* 225 (1994) 71–82.
- [53] B. SenGupta, F. Friedberg, S.D. Detera-Wadleigh, Molecular analysis of human and rat calmodulin complementary DNA clones. Evidence for additional active genes in these species, *J. Biol. Chem.* 262 (1987) 16663–16670.
- [54] B. SenGupta, S.D. Detera-Wadleigh, O.W. McBride, F. Friedberg, A calmodulin pseudogene on human chromosome 17, *Nucleic Acids Res.* 17 (1989) 2868.
- [55] J.P. Stein, R.P. Munjaal, L. Lagace, E.C. Lai, B.W. O'Malley, A.R. Means, Tissue-specific expression of a chicken calmodulin pseudogene lacking intervening sequences, *Proc. Natl. Acad. Sci. USA*, 80 (1983) 6485–6489.
- [56] D. Zheng, M.B. Gerstein, The ambiguous boundary between genes and pseudogenes: the dead rise up, or do they? *Trends Genet.* 23 (2007) 219–224.
- [57] S.W. Cheetham, G.J. Faulkner, M.E. Dinger, Overcoming challenges and dogmas to understand the functions of pseudogenes, *Nat. Rev. Genet.* 21 (2020) 191–201.
- [58] R.K. Singh, D. Singh, A. Yadava, A.K. Srivastava, Molecular fossils "pseudogenes" as functional signature in biological system, *Genes Genom.* 42 (2020) 619–630.
- [59] M. Koller, E.E. Strehler, Characterization of an intronless human calmodulin-like pseudogene, *FEBS J.* 239 (1988) 121–128.
- [60] M.W. Berchtold, M. Koller, R. Egli, J.A. Rhyner, H. Hameister, E.E. Strehler, Localization of the intronless gene coding for calmodulin-like protein CLP to human chromosome 10p13-ter, *Hum. Genet.* 90 (1993) 496–500.
- [61] M. Koller, E.E. Strehler, Functional analysis of the promoters of the human CaMIII calmodulin gene and of the intronless gene coding for a calmodulin-like protein, *Biochim. Biophys. Acta* 1163 (1993) 1–9.
- [62] M. Koller, A. Baumer, E.E. Strehler, Characterization of two novel human retrospliced genes related to the calmodulin-encoding gene, *CaMII*, *Gene* 97 (1991) 245–251.
- [63] M.N. Szmulewicz, G.E. Novick, R.J. Herrera, Effects of Alu insertions on gene function, *Electrophoresis* 19 (1998) 1260–1264.
- [64] F. Friedberg, A.R. Rhoads, Calculation and verification of the ages of retroprocessed pseudogenes, *Mol. Phylogenet. Evol.* 16 (2000) 127–130.
- [65] F. Friedberg, A.R. Rhoads, Age and detection of retroprocessed pseudogenes in murine rodents, *J. Mol. Evol.* 54 (2002) 102–106.
- [66] H. Nojima, H. Sokabe, Structure of rat calmodulin processed genes with implications for a mRNA-mediated process of insertion, *J. Mol. Biol.* 190 (1986) 391–400.
- [67] H. Nojima, K. Kishi, H. Sokabe, Multiple calmodulin mRNA species are derived from two distinct genes, *Mol. Cell. Biol.* 7 (1987) 1873–1880.
- [68] H. Nojima, H. Sokabe, Genes and pseudogenes for calmodulin in the spontaneously hypertensive rat, *J. Hypertens. Suppl.* 6 (1988) S231–S233.
- [69] K. Luby-Phelps, M. Hori, J.M. Phelps, D. Won, Ca(2+)-regulated dynamic compartmentalization of calmodulin in living smooth muscle cells, *J. Biol. Chem.* 270 (1995) 21532–21538.
- [70] X. Wu, D.M. Bers, Free and bound intracellular calmodulin measurements in cardiac myocytes, *Cell Calcium* 41 (2007) 353–364.
- [71] Q.K. Tran, D.J. Black, A. Persechini, Intracellular coupling via limiting calmodulin, *J. Biol. Chem.* 278 (2003) 24247–24250.
- [72] A. Persechini, B. Cronk, The relationship between the free concentrations of Ca<sup>2+</sup> and Ca<sup>2+</sup>-calmodulin in intact cells, *J. Biol. Chem.* 274 (1999) 6827–6830.
- [73] R.P. Estep, K.A. Alexander, D.R. Storm, Regulation of free calmodulin levels in neurons by neuromodulin: relationship to neuronal growth and regeneration, *Curr. Top. Cell. Regul.* 31 (1990) 161–180.
- [74] T.H. Kim, Z.K. Abdullaev, A.D. Smith, K.A. Ching, D.I. Loukinov, R.D. Green, M. Q. Zhang, V.V. Lobanenko, B. Ren, Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome, *Cell* 128 (2007) 1231–1245.
- [75] S.J. Holwerda, W. de Laat, CTCF: the protein, the binding partners, the binding sites and their chromatin loops, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368 (2013), 20120369.
- [76] G. Bai, R.A. Nichols, B. Weiss, Cyclic AMP selectively up-regulates calmodulin genes I and II in PC12 cells, *Biochim. Biophys. Acta* 1130 (1992) 189–196.
- [77] P.N. Epstein, M.A. Christenson, A.R. Means, Chicken calmodulin promoter activity in proliferating and differentiated cells, *Mol. Endocrinol.* 3 (1989) 193–202.
- [78] T. Juven-Gershon, J.T. Kadonaga, Regulation of gene expression via the core promoter and the basal transcriptional machinery, *Dev. Biol.* 339 (2010) 225–229.
- [79] C. Li, B. Lenhard, N.M. Luscombe, Integrated analysis sheds light on evolutionary trajectories of young transcription start sites in the human genome, *Genome Res.* 28 (2018) 676–688.
- [80] K. Matsuo, H. Ikeshima, K. Shimoda, A. Umezawa, J. Hata, K. Maejima, H. Nojima, T. Takano, Expression of the rat calmodulin gene II in the central nervous system: a 294-base promoter and 68-base leader segment mediates neuron-specific gene expression in transgenic mice, *Brain Res. Mol. Brain Res.* 20 (1993) 9–20.
- [81] K. Shimoda, H. Ikeshima, K. Matsuo, J. Hata, K. Maejima, T. Takano, Spatial and temporal regulation of the rat calmodulin gene III directed by a 877-base promoter and 103-base leader segment in the mature and embryonal central nervous system of transgenic mice, *Brain Res. Mol. Brain Res.* 31 (1995) 61–70.
- [82] H. Ikeshima, K. Shimoda, K. Matsuo, J. Hata, K. Maejima, T. Takano, Spermatocyte-specific transcription by calmodulin gene II promoter in transgenic mice, *Mol. Cell. Endocrinol.* 99 (1994) 49–53.
- [83] C.T. Ong, V.G. Corces, Enhancer function: new insights into the regulation of tissue-specific gene expression, *Nat. Rev. Genet.* 12 (2011) 283–293.
- [84] A.L. Roy, D.S. Singer, Core promoters in transcription: old problem, new insights, *Trends Biochem. Sci.* 40 (2015) 165–171.
- [85] F. Muller, L. Torra, Chromatin and DNA sequences in defining promoters for transcription initiation, *Biochim. Biophys. Acta* 1839 (2014) 118–128.
- [86] S. Ogbourne, T.M. Antalis, Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes, *Biochem. J.* 331 (Pt 1) (1998) 1–14.
- [87] J.M. Flanagan, V. Popenkitye, N. Pozdniakovaite, M. Sobolev, A. Assadzadeh, A. Schumacher, M. Zangeneh, L. Lau, C. Virtanen, S.C. Wang, A. Petronis, Intra- and interindividual epigenetic variation in human germ cells, *Am. J. Hum. Genet.* 79 (2006) 67–84.
- [88] J. Cheng, L. Tang, Q. Hong, H. Ye, X. Xu, L. Xu, S. Bu, Q. Wang, D. Dai, D. Jiang, S. Duan, Investigation into the promoter DNA methylation of three genes (*CAMK1D*, *CRY2* and *CALM2*) in the peripheral blood of patients with type 2 diabetes, *Exp. Immunol. Ther. Med.* 8 (2014) 579–584.
- [89] N. Gaborit, S.L. Bouter, V. Szuts, A. Varro, D. Escande, S. Nattel, S. Demolombe, Regional and tissue specific transcript signatures of ion channel genes in the non-diseased human heart, *J. Physiol.* 582 (2007) 675–693.
- [90] T. Liu, X. Han, S. Zheng, Q. Liu, A. Tuerxun, Q. Zhang, L. Yang, X. Lu, *CALM1* promotes progression and dampens chemosensitivity to EGFR inhibitor in esophageal squamous cell carcinoma, *Cancer Cell Int.* 21 (2021) 121.
- [91] S.Y. Park, Y.R. Seo, M.J. Ko, J.H. Lee, K.S. Chun, M.J. Kim, Y.K. Choo, T.J. Lee, Y. H. Lee, Targeting *CALM2* inhibits hepatocellular carcinoma growth and metastasis by suppressing E2F5-mediated cell cycle progression, *Anticancer Res.* 41 (2021) 1315–1325.
- [92] M.A. Christenson, A.R. Means, Coordinate regulation of mRNAs from multiple calmodulin genes during myoblast differentiation *in vitro*, *J. Cell. Physiol.* 154 (1993) 343–349.
- [93] M.N. Gannon, B.S. McEwen, Distribution and regulation of calmodulin mRNAs in rat brain, *Brain Res. Mol. Brain Res.* 22 (1994) 186–192.
- [94] F. Friedberg, A.R. Rhoads, Sequence homology of the 3'-untranslated region of calmodulin III in mammals, *Mol. Biol. Rep.* 28 (2001) 27–30.
- [95] F. Friedberg, A.A. Saunders, A.R. Rhoads, Conservation of the 3'-untranslated regions of calmodulin mRNAs in cetaceans, *Mol. Biol. Rep.* 30 (2003) 193–198.
- [96] J.D. Richter, CPEB: a life in translation, *Trends Biochem. Sci.* 32 (2007) 279–285.
- [97] C.M. Atkins, N. Nozaki, Y. Shigeri, T.R. Soderling, Cytoplasmic polyadenylation element binding protein-dependent protein synthesis is regulated by calcium/calmodulin-dependent protein kinase II, *J. Neurosci.* 24 (2004) 5193–5201.
- [98] L.E. Hake, R. Mendez, J.D. Richter, Specificity of RNA binding by CPEB: requirement for RNA recognition motifs and a novel zinc finger, *Mol. Cell. Biol.* 18 (1998) 685–693.
- [99] T. Sharangdhar, Y. Sugimoto, J. Heraud-Farlow, S.M. Fernandez-Moya, J. Ehses, I. Ruiz de Los Mozos, J. Ule, M.A. Kiebler, A retained intron in the 3'-UTR of *Calm3* mRNA mediates its Staufen2- and activity-dependent localization to neuronal dendrites, *EMBO Rep.* 18 (2017) 1762–1774.
- [100] B. Bae, H.N. Gruner, M. Lynch, T. Feng, K. So, D. Oliver, G.S. Mastick, W. Yan, S. Pierant, P. Miura, Elimination of *Calm1* long 3'-UTR mRNA isoform by CRISPR-Cas9 gene editing impairs dorsal root ganglion development and hippocampal neuron activation in mice, *RNA* 26 (2020) 1414–1430.
- [101] S. Ikeda, A. He, S.W. Kong, J. Lu, R. Bejar, N. Bodyak, K.H. Lee, Q. Ma, P.M. Kang, T.R. Golub, W.T. Pu, MicroRNA-1 negatively regulates expression of the hypertrophy-associated calmodulin and *Mef2a* genes, *Mol. Cell. Biol.* 29 (2009) 2193–2204.

- [102] J. Ai, R. Zhang, X. Gao, H.F. Niu, N. Wang, Y. Xu, Y. Li, N. Ma, L.H. Sun, Z.W. Pan, W.M. Li, B.F. Yang, Overexpression of microRNA-1 impairs cardiac contractile function by damaging sarcomere assembly, *Cardiovasc. Res.* 95 (2012) 385–393.
- [103] X. Zhou, F. Sun, S. Luo, W. Zhao, T. Yang, G. Zhang, M. Gao, R. Lu, Y. Shu, W. Mu, Y. Zhuang, F. Ding, C. Xu, Y. Lu, Let-7a is an antihypertrophic regulator in the heart via targeting calmodulin, *Int. J. Biol. Sci.* 13 (2017) 22–31.
- [104] B. Zhao, Z. Zhu, J. Hao, Z. Wan, X. Guo, Decreased plasma miR-335 expression in patients with acute ischemic stroke and its association with calmodulin expression, *J. Int. Med. Res.* 44 (2016) 1331–1338.
- [105] J. Yu, C. Yu, B. Feng, X. Zhan, N. Luo, X. Yu, Q. Zhou, Intrarenal microRNA signature related to the fibrosis process in chronic kidney disease: identification and functional validation of key miRNAs, *BMC Nephrol.* 20 (2019) 336.
- [106] P.N. Brown, H. Yin, PNA-based microRNA inhibitors elicit anti-inflammatory effects in microglia cells, *Chem. Commun.* 49 (2013) 4415–4417 (Camb).
- [107] S.P. Zhang, N. Natsukari, G. Bai, R.A. Nichols, B. Weiss, Localization of the multiple calmodulin messenger RNAs in differentiated PC12 cells, *Neuroscience* 55 (1993) 571–582.
- [108] E. Kortvely, K. Gulya, Calmodulin, and various ways to regulate its activity, *Life Sci.* 74 (2004) 1065–1070.
- [109] A. Palfi, E. Kortvely, E. Fekete, B. Kovacs, S. Varszegi, K. Gulya, Differential calmodulin gene expression in the rodent brain, *Life Sci.* 70 (2002) 2829–2855.
- [110] A. Palfi, E. Kortvely, E. Fekete, K. Gulya, Multiple calmodulin mRNAs are selectively transported to functionally different neuronal and glial compartments in the rat hippocampus. An electron microscopic *in situ* hybridization study, *Life Sci.* 77 (2005) 1405–1415.
- [111] I. Orojan, L. Bakota, K. Gulya, Differential calmodulin gene expression in the nuclei of the rat midbrain-brain stem region, *Acta Histochem.* 108 (2006) 455–462.
- [112] V. Bogdanov, A.M. Soltisz, B. Hernández-Aragón, G. Sakuta, R. Veerarahgavan, J. P. Davis, S. Gyorko, Spatially and functionally distinct pools of calmodulin mRNA in cardiac myocytes, *Biophys. J.* 121 (1) (2022) 257A. Suppl.
- [113] J.J. Saucerman, D.M. Bers, Calmodulin binding proteins provide domains of local  $Ca^{2+}$  signaling in cardiac myocytes, *J. Mol. Cell Cardiol.* 52 (2012) 312–316.
- [114] V. Presnyak, N. Alhusaini, Y.H. Chen, S. Martin, N. Morris, N. Kline, S. Olson, D. Weinberg, K.E. Baker, B.R. Graveley, J. Collier, Codon optimality is a major determinant of mRNA stability, *Cell* 160 (2015) 1111–1124.
- [115] S.G. Medina-Munoz, G. Kushawah, L.A. Castellano, M. Diez, M.L. DeVore, M.J. B. Salazar, A.A. Bazzini, Crosstalk between codon optimality and cis-regulatory elements dictates mRNA stability, *Genome Biol.* 22 (2021) 14.
- [116] A. Narula, J. Ellis, J.M. Taliaferro, O.S. Rissland, Coding regions affect mRNA stability in human cells, *RNA* 25 (2019) 1751–1764.
- [117] Q. Wu, S.G. Medina, G. Kushawah, M.L. DeVore, L.A. Castellano, J.M. Hand, M. Wright, A.A. Bazzini, Translation affects mRNA stability in a codon-dependent manner in human cells, *eLife* 8 (2019), e45396, <https://doi.org/10.7554/eLife.45396>.
- [118] A.A. Bazzini, F. Del Viso, M.A. Moreno-Mateos, T.G. Johnstone, C.E. Vejnar, Y. Qin, J. Yao, M.K. Khokha, A.J. Giraldez, Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition, *EMBO J.* 35 (2016) 2087–2103.
- [119] M.E. Forrest, O. Pinkard, S. Martin, T.J. Sweet, G. Hanson, J. Collier, Codon and amino acid content are associated with mRNA stability in mammalian cells, *PLoS One* 15 (2020), e0228730.
- [120] F. Hia, S.F. Yang, Y. Shichino, M. Yoshinaga, Y. Murakawa, A. Vandenberg, A. Fukao, T. Fujiwara, M. Landthaler, T. Natsume, S. Adachi, S. Iwasaki, O. Takeuchi, Codon bias confers stability to human mRNAs, *EMBO Rep.* 20 (2019) e48220.
- [121] W. Shen, D. Wang, B. Ye, M. Shi, L. Ma, Y. Zhang, Z. Zhao, GC3-biased gene domains in mammalian genomes, *Bioinformatics* 31 (2015) 3081–3084.
- [122] L.L. Yan, H.S. Zaher, How do cells cope with RNA damage and its consequences? *J. Biol. Chem.* 294 (2019) 15158–15171.
- [123] L. Crotti, C.N. Johnson, E. Graf, G.M. De Ferrari, B.F. Cuneo, M. Ovidia, J. Papagiannis, M.D. Feldkamp, S.G. Rathi, J.D. Kunic, M. Pedrazzini, T. Wieland, P. Lichtner, B.M. Beckmann, T. Clark, C. Shaffer, D.W. Benson, S. Kaab, T. Meitinger, T.M. Strom, W.J. Chazin, P.J. Schwartz, A.L. George Jr., Calmodulin mutations associated with recurrent cardiac arrest in infants, *Circulation* 127 (2013) 1009–1017.
- [124] J. Colomer, N. Agell, P. Engel, O. Bachs, Expression of calmodulin and calmodulin binding proteins in lymphoblastoid cells, *J. Cell. Physiol.* 159 (1994) 542–550.
- [125] H. Li, S. Panina, A. Kaur, M.J. Ruano, P. Sanchez-Gonzalez, J.M. la Cour, A. Stephan, U.H. Olesen, M.W. Berchtold, A. Villalobo, Regulation of the ligand-dependent activation of the epidermal growth factor receptor by calmodulin, *J. Biol. Chem.* 287 (2012) 3273–3281.
- [126] Y. Liu, Y. Xu, F. Xiao, J. Zhang, Y. Wang, Y. Yao, J. Yang, Comprehensive analysis of a circRNA-miRNA-mRNA network to reveal potential inflammation-related targets for gastric adenocarcinoma, *Mediat. Inflamm.* 2020 (2020), 9435608.
- [127] R. Rust, L. Visser, J. van der Leij, G. Harms, T. Blokzijl, J.C. Deloulme, P. van der Vlies, W. Kamps, K. Kok, M. Lim, S. Poppema, A. van den Berg, High expression of calcium-binding proteins, S100A10, S100A11 and CALM2 in anaplastic large cell lymphoma, *Br. J. Haematol.* 131 (2005) 596–608.
- [128] L. Li, H. Mo, J. Zhang, Y. Zhou, X. Peng, X. Luo, The role of heat shock protein 90B1 in patients with polycystic ovary syndrome, *PLoS One* 11 (2016), e0152837.
- [129] A.C. Calvo, R. Manzano, G. Atencia-Cibreiro, S. Oliván, M.J. Muñoz, P. Zaragoza, P. Cordero-Vazquez, J. Esteban-Perez, A. Garcia-Redondo, R. Osta, Genetic biomarkers for ALS disease in transgenic SOD1(G93A) mice, *PLoS One* 7 (2012) e32632.
- [130] J. Qiu, Y.H. Ni, R.H. Chen, C.B. Ji, F. Liu, C.M. Zhang, C.L. Gao, X.H. Chen, M. L. Tong, X. Chi, X.Y. Zhou, X.R. Guo, Gene expression profiles of adipose tissue of obese rats after central administration of neuropeptide Y-Y5 receptor antisense oligodeoxynucleotides by cDNA microarrays, *Peptides* 29 (2008) 2052–2060.
- [131] M. Nyegaard, M.T. Overgaard, M.T. Sondergaard, M. Vranas, E.R. Behr, L. L. Hildebrandt, J. Lund, P.L. Hedley, A.J. Camm, G. Wettrell, I. Fosdal, M. Christiansen, A.D. Borglum, Mutations in calmodulin cause ventricular tachycardia and sudden cardiac death, *Am. J. Hum. Genet.* 91 (2012) 703–712.
- [132] L. Crotti, C. Spazzolini, D.J. Tester, A. Ghidoni, A.E. Baruteau, B.M. Beckmann, E. R. Behr, J.S. Bennett, C.R. Bezzina, Z.A. Bhuiyan, A. Celiker, M. Cerrone, F. Dagradi, G.M. De Ferrari, S.P. Etheridge, M. Fatah, P. Garcia-Pavia, S. Al-Ghamdi, R.M. Hamilton, Z.N. Al-Hassnan, M. Horie, J. Jimenez-Jaimez, R. J. Kanter, J.P. Kaski, M.C. Kotta, N. Lahrouchi, N. Makita, G. Norrish, H. H. Odland, S. Ohno, J. Papagiannis, G. Parati, N. Sekarski, K. Tveten, M. Vatta, G. Webster, A.A.M. Wilde, J. Wojcicki, A.L. George, M.J. Ackerman, P.J. Schwartz, Calmodulin mutations and life-threatening cardiac arrhythmias: insights from the International Calmodulinopathy Registry, *Eur. Heart J.* 40 (2019) 2964–2975.
- [133] W.J. Chazin, C.N. Johnson, Calmodulin mutations associated with heart arrhythmia: a status report, *Int. J. Mol. Sci.* 21 (2020) 1418.
- [134] B. Badone, C. Ronchi, M.C. Kotta, L. Sala, A. Ghidoni, L. Crotti, A. Zaza, Calmodulinopathy: functional effects of CALM mutations and their relationship with clinical phenotypes, *Front. Cardiovasc. Med.* 5 (2018) 176.
- [135] M. Nyegaard, M.T. Overgaard, The International Calmodulinopathy Registry: recording the diverse phenotypic spectrum of un-CALM hearts, *Eur. Heart J.* 40 (2019) 2976–2978.
- [136] K.A. Woll, F. Van Petegem, Calcium release channels: structure and function of IP3 receptors and ryanodine receptors, *Physiol. Rev.* 102 (2022) 209–268.
- [137] C. Christel, A. Lee,  $Ca^{2+}$ -dependent modulation of voltage-gated  $Ca^{2+}$  channels, *Biochim. Biophys. Acta* 1820 (2012) 1243–1252.
- [138] L.M. Wren, J. Jimenez-Jaimez, S. Al-Ghamdi, J.Y. Al-Aama, A. Bdeir, Z.N. Al-Hassnan, J.L. Kuan, R.Y. Foo, F. Potet, C.N. Johnson, M.C. Aziz, G.L. Carvill, J. P. Kaski, L. Crotti, F. Perin, L. Monserrat, P.W. Burridge, P.J. Schwartz, W. J. Chazin, Z.A. Bhuiyan, A.L. George Jr., Genetic mosaicism in calmodulinopathy, *Circ. Genom. Precis. Med.* 12 (2019) 375–385.
- [139] S. Beghi, F. Cavaliere, A. Buschini, Gene polymorphisms in calcium-calmodulin pathway: focus on cardiovascular disease, *Mutat. Res. Rev. Mutat. Res.* 786 (2020), 108325.
- [140] M.M. Luke, E.S. O'Meara, C.M. Rowland, D. Shiffman, L.A. Bare, A.R. Arellano, W.T. Longstreth Jr., T. Lumley, K. Rice, R.P. Tracy, J.J. Devlin, B.M. Psaty, Gene variants associated with ischemic stroke: the cardiovascular health study, *Stroke* 40 (2009) 363–368.
- [141] L. Gu, J. Huang, J. Li, S. Huang, M. Li, L. Gong, T. Li, L. Su, Association of CALM1 rs3179089 polymorphism with ischemic stroke in Chinese Han population, *Neuromol. Med.* 20 (2018) 271–279.
- [142] Z. Liu, X. Liu, H. Yu, J. Pei, Y. Zhang, J. Gong, J. Pu, Common variants in TRDN and CALM1 are associated with risk of sudden cardiac death in chronic heart failure patients in Chinese Han population, *PLoS One* 10 (2015), e0132459.
- [143] F.W. Friedrich, P. Bausero, Y. Sun, A. Treszl, E. Kramer, D. Juhr, P. Richard, K. Wegscheider, K. Schwartz, D. Brito, E. Arbustini, A. Waldenström, R. Isnard, M. Komajda, T. Eschenhagen, L. Carrier, E.H.F. Project, A new polymorphism in human calmodulin III gene promoter is a potential modifier gene for familial hypertrophic cardiomyopathy, *Eur. Heart J.* 30 (2009) 1648–1655.
- [144] M. Thomikos, A. Thanassoulas, K. Beck, V. Vassilakopoulou, H. Hu, B.L. Calver, M. Nedicoridou, J. Kashir, L. Blayney, E. Livianou, P. Rizkallah, G. Nounesis, F. A. Lai, Altered RyR2 regulation by the calmodulin F90L mutation associated with idiopathic ventricular fibrillation and early sudden cardiac death, *FEBS Lett.* 588 (2014) 2898–2902.
- [145] M.W. Berchtold, M. Munk, K. Kulej, I. Porth, L. Lorentzen, S. Panina, T. Zacharias, M.R. Larsen, J.M. la Cour, The heart arrhythmia-linked D130G calmodulin mutation causes premature inhibitory autophosphorylation of CaMKII, *Biochim. Biophys. Acta Mol. Cell Res.* 1868 (2021), 119119.
- [146] E. Jia, Y. Zhou, H. Shi, M. Pan, X. Zhao, Q. Ge, Effects of brain tissue section processing and storage time on gene expression, *Anal. Chim. Acta* 1142 (2021) 38–47.
- [147] S. Doll, M. Drensen, P.E. Geyer, D.N. Itzhak, C. Braun, S.A. Doppler, F. Meier, M. A. Deutsch, H. Lahm, R. Lange, M. Krane, M. Mann, Region and cell-type resolved quantitative proteomic map of the human heart, *Nat. Commun.* 8 (2017) 1469.
- [148] B.Z. Peterson, C.D. DeMaria, J.P. Adelman, D.T. Yue, Calmodulin is the  $Ca^{2+}$  sensor for  $Ca^{2+}$ -dependent inactivation of L-type calcium channels, *Neuron* 22 (1999) 549–558.
- [149] E. Nanou, W.A. Catterall, Calcium channels, synaptic plasticity, and neuropsychiatric disease, *Neuron* 98 (2018) 466–481.
- [150] S.B. Gabelli, A. Boto, V.H. Kuhns, M.A. Bianchet, F. Farinelli, S. Aripirala, J. Yoder, J. Jakoncic, G.F. Tomaselli, L.M. Amzel, Regulation of the Nav1.5 cytoplasmic domain by calmodulin, *Nat. Commun.* 5 (2014) 5126.
- [151] S. Nathan, S.B. Gabelli, J.B. Yoder, L. Srinivasan, R.W. Aldrich, G.F. Tomaselli, M. Ben-Johny, L.M. Amzel, Structural basis of cytoplasmic Nav1.5 and Nav1.4 regulation, *J. Gen. Physiol.* 153 (2021), e202012722, <https://doi.org/10.1085/jgp.202012722>.
- [152] Y. Saimi, C. Kung, Calmodulin as an ion channel subunit, *Annu. Rev. Physiol.* 64 (2002) 289–311.
- [153] D.C. Pipilas, C.N. Johnson, G. Webster, J. Schlaepfer, F. Fellmann, N. Sekarski, L. M. Wren, K.V. Ogorodnik, D.M. Chazin, W.J. Chazin, L. Crotti, Z.A. Bhuiyan, A. L. George Jr., Novel calmodulin mutations associated with congenital long QT syndrome affect calcium current in human cardiomyocytes, *Heart Rhythm* 13 (2016) 2012–2019.

- [154] S.P. Etheridge, M.C. Niu, Calmodulinopathies: throwing back the veil on the newest life-threatening genetic arrhythmia syndrome, *Curr. Opin. Cardiol.* 36 (2021) 61–66.
- [155] H. Mototani, A. Mabuchi, S. Saito, M. Fujioka, A. Iida, Y. Takatori, A. Kotani, T. Kubo, K. Nakamura, A. Sekine, Y. Murakami, T. Tsunoda, K. Notoya, Y. Nakamura, S. Ikegawa, A functional single nucleotide polymorphism in the core promoter region of *CALM1* is associated with hip osteoarthritis in Japanese, *Hum. Mol. Genet.* 14 (2005) 1009–1017.
- [156] K.S. Cheah, N.G. Stoker, J.R. Griffin, F.G. Grosveld, E. Solomon, Identification and characterization of the human type II collagen gene (*COL2A1*), *Proc. Natl. Acad. Sci. USA*, 82 (1985) 2555–2559.
- [157] K.J. Doege, M. Sasaki, T. Kimura, Y. Yamada, Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, aggrecan. Human-specific repeats, and additional alternatively spliced forms, *J. Biol. Chem.* 266 (1991) 894–902.
- [158] J. Loughlin, J.S. Sinsheimer, A. Carr, K. Chapman, The *CALM1* core promoter polymorphism is not associated with hip osteoarthritis in a United Kingdom Caucasian population, *Osteoarthr. Cartil.* 14 (2006) 295–298.
- [159] M. Poulou, M. Kaliakatsos, A. Tsezou, E. Kanavakis, K.N. Malizos, M. Tzetzis, Association of the *CALM1* core promoter polymorphism with knee osteoarthritis in patients of Greek origin, *Genet. Test.* 12 (2008) 263–265.
- [160] F.T. Hoaglund, *CALM1* promoter polymorphism gene and Japanese congenital hip disease, *Osteoarthr. Cartil.* 15 (2007) 593.
- [161] H. Mototani, A. Iida, Y. Nakamura, S. Ikegawa, Identification of sequence polymorphisms in *CALM2* and analysis of association with hip osteoarthritis in a Japanese population, *J. Bone Miner. Metab.* 28 (2010) 547–553.
- [162] D. Zhao, G.X. Qiu, Y.P. Wang, J.G. Zhang, J.X. Shen, Z.H. Wu, H. Wang, Association of calmodulin1 gene polymorphisms with susceptibility to adolescent idiopathic scoliosis, *Orthop. Surg.* 1 (2009) 58–65.
- [163] K.F. Gorman, C. Julien, A. Moreau, The genetic epidemiology of idiopathic scoliosis, *Eur. Spine J.* 21 (2012) 1905–1919.
- [164] D. Shi, H. Ni, J. Dai, J. Qin, Y. Xu, L. Zhu, C. Yao, Z. Shao, D. Chen, Z. Xu, L. Yi, S. Ikegawa, Q. Jiang, Lack of association between the *CALM1* core promoter polymorphism (-16C/T) and susceptibility to knee osteoarthritis in a Chinese Han population, *BMC Med. Genet.* 9 (2008) 91.
- [165] J. Shi, S.T. Gao, Z.T. Lv, W.B. Sheng, H. Kang, The association between rs12885713 polymorphism in *CALM1* and risk of osteoarthritis, *Medicine* 97 (2018) e12235 (Baltimore).
- [166] J. Banerjee, R. Magnani, M. Nair, L.M. Dirk, S. DeBolt, I.B. Maiti, R.L. Houtz, Calmodulin-mediated signal transduction pathways in Arabidopsis are fine-tuned by methylation, *Plant Cell* 25 (2013) 4493–4511.
- [167] H.L. Zhang, B. Zhao, W. Han, Y.B. Sun, P. Yang, Y. Chen, D. Ni, J. Zhang, D. M. Yin, Acetylation of calmodulin regulates synaptic plasticity and fear learning, *J. Biol. Chem.* 297 (2021), 101034.
- [168] H.L. Zhang, W. Han, Y.Q. Du, B. Zhao, P. Yang, D.M. Yin, SRC3 acetylates calmodulin in the mouse brain to regulate synaptic plasticity and fear learning, *J. Biol. Chem.* 297 (2021), 101044.
- [169] P. Thiry, A. Vandermeers, M.C. Vandermeers-Piret, J. Rathe, J. Christophe, The activation of brain adenylate cyclase and brain cyclic-nucleotide phosphodiesterase by seven calmodulin derivatives, *Eur. J. Biochem.* 103 (1980) 409–414.
- [170] R. Chattopadhyaya, W.E. Meador, A.R. Means, F.A. Quijcho, Calmodulin structure refined at 1.7 Å resolution, *J. Mol. Biol.* 228 (1992) 1177–1192.
- [171] M. Lankaputhra, A. Voskoboinik, Congenital long QT syndrome: a clinician's guide, *Intern. Med. J.* 51 (2021) 1999–2011.
- [172] S.G. Priori, A. Mazzanti, D.J. Santiago, D. Kukavica, A. Trancuccio, J.C. Kovacic, Precision medicine in catecholaminergic polymorphic ventricular tachycardia: JACC focus seminar 5/5, *J. Am. Coll. Cardiol.* 77 (2021) 2592–2612.
- [173] M. Visser, J.F. van der Heijden, P.A. Doevendans, P. Loh, A.A. Wilde, R. J. Hassink, Idiopathic ventricular fibrillation: the struggle for definition, diagnosis, and follow-up, *Circ. Arrhythmia Electrophysiol.* 9 (2016), e003817.
- [174] D.J. Tester, M.J. Ackerman, Cardiomyopathic and channelopathic causes of sudden unexplained death in infants and children, *Annu. Rev. Med.* 60 (2009) 69–84.
- [175] R.J. Koene, W.O. Adkisson, D.G. Benditt, Syncope and the risk of sudden cardiac death: Evaluation, management, and prevention, *J. Arrhythmia* 33 (2017) 533–544.
- [176] J.D. Karkera, F. Friedberg, Sequence of the human *CALM II* calmodulin gene promoter region, *Mol. Biol. Rep.* 25 (1998) 193–195.
- [177] A. Charlesworth, H.A. Meijer, C.H. de Moor, Specificity factors in cytoplasmic polyadenylation, *Wiley Interdiscip. Rev. RNA* 4 (2013) 437–461.
- [178] J.D. Ziebarth, A. Bhattacharya, Y. Cui, CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization, *Nucleic Acids Res.* 41 (2013) D188–D194.