Article

A Multi-Layered Study on Harmonic Oscillations in Mammalian Genomics and Proteomics

Nikolai Genov [1,2,†]**, Stefano Castellana** [3,†]**, Felix Scholkmann** [4,5,†]**, Daniele Capocefalo** [3]**,
Mauro Truglio** [3]**, Jessica Rosati** [6]**, Elisa Maria Turco** [6]**, Tommaso Biagini** [3]**, Annalucia Carbone** [7]**,
Tommaso Mazza** [3]**, Angela Relógio** [1,2,†,*]** and Gianluigi Mazzoccoli** [7,†,*]

[1] Institute for Theoretical Biology (ITB), Charité–Universitätsmedizin Berlin, corporate member of Freie
   Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, 1011 Berlin, Germany;
   nikolai.genov@fu-berlin.de

[2] Medical Department of Hematology, Oncology, and Tumor Immunology, and Molekulares
   Krebsforschungszentrum (MKFZ), Charité–Universitätsmedizin Berlin, corporate member of Freie
   Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, 13353 Berlin, Germany

[3] Bioinformatics Unit, IRCCS "Casa Sollievo della Sofferenza", 71013 San Giovanni Rotondo (FG), Italy;
   s.castellana@css-mendel.it (S.C.); d.capocefalo@css-mendel.it (D.C.); m.truglio@css-mendel.it (M.T.);
   t.biagini@css-mendel.it (T.B.); t.mazza@css-mendel.it (T.M.)

[4] Research Office for Complex Physical and Biological Systems (ROCoS), 8006 Zurich, Switzerland;
   felix.scholkmann@usz.ch

[5] Department of Neonatology, University Hospital Zurich, University of Zurich, 8091 Zurich, Switzerland

[6] Cell Reprogramming Unit, Fondazione IRCCS "Casa Sollievo della Sofferenza", 71013 San Giovanni
   Rotondo (FG), Italy; j.rosati@css-mendel.it (J.R.); e.turco@css-mendel.it (E.M.T.)

[7] Division of Internal Medicine and Chronobiology Unit, Fondazione IRCCS "Casa Sollievo della Sofferenza",
   71013 San Giovanni Rotondo (FG), Italy; annalucia.carbone@gmail.com

* Correspondence: angela.relogio@charite.de (A.R.); g.mazzoccoli@operapadrepio.it (G.M.)

† These authors contributed equally to this article.

check for updates

Abstract: Cellular, organ, and whole animal physiology show temporal variation predominantly featuring 24-h (circadian) periodicity. Time-course mRNA gene expression profiling in mouse liver showed two subsets of genes oscillating at the second (12-h) and third (8-h) harmonic of the prime (24-h) frequency. The aim of our study was to identify specific genomic, proteomic, and functional properties of ultradian and circadian subsets. We found hallmarks of the three oscillating gene subsets, including different (i) functional annotation, (ii) proteomic and electrochemical features, and (iii) transcription factor binding motifs in upstream regions of 8-h and 12-h oscillating genes that seemingly allow the link of the ultradian gene sets to a known circadian network. Our multifaceted bioinformatics analysis of circadian and ultradian genes suggests that the different rhythmicity of gene expression impacts physiological outcomes and may be related to transcriptional, translational and post-translational dynamics, as well as to phylogenetic and evolutionary components.

Keywords: biological clock; rhythmic gene expression; rhythmic protein expression; circadian rhythms; ultradian rhythms; electrochemical features

1. Introduction

Tissue and cellular functions underlying physiology of living organisms show time-dependent variations predominantly featured by 24-h (circadian) periodicity and driven by molecular clockworks operated by rhythmically expressed genes and proteins hardwiring transcriptional/translational feedback loops (TTFL) [1–4].

Circadian expressed genes include a handful of core-clock genes that in turn drive thousands of downstream clock-controlled genes [5,6]. Experiments performed in animal models showed that approximately half of the transcriptome shows 24-h oscillations that manage crucial biological processes such as the cell cycle, proliferation, metabolism, DNA damage repair, apoptosis and autophagy [7–10].

The interacting positive and negative limbs of the TTFL regulate gene transcription through sequential cycles of transcriptional activation of the expression of clock genes followed by transcriptional suppression by their protein products [11,12]. The positive limb is operated by CLOCK and BMAL1 that heterodimerize and activate the transcription of cryptochrome genes (*Cry1* and *Cry2*) and period genes (*Per1*, *Per2* and *Per3*), which operate the negative limb encoding repressors hindering gene transcription. Conversely, *Bmal1* rhythmic expression is driven by the nuclear receptors REV-ERBα and RORα through competitive binding at its promoter region [13,14].

Gene expression profiling performed by means of high-throughput measurements with DNA microarrays and quantitative PCR in mouse liver specimens collected at regular time intervals showed that two groups of genes oscillate at the second (12-h) and third (8-h) harmonic of the fundamental (24-h) frequency [15].

The different periodicity of gene expression impacts physiological outcomes [16] and may be related to transcriptional, translational and post-translational dynamics [17–21], as well as to phylogenetic and evolutionary components [22].

The aim of our study was to characterize genomic and proteomic features of the clusters of genes oscillating with harmonics of circadian periodicity. We exploited bioinformatics tools for functional prediction to identify the biological functions and enriched signalling pathways and to perform comparative qualitative proteomic analysis. Finally, we implemented several computational strategies in order to detect the presence of significant de-novo regulatory motifs and known transcription factor binding sites in the promoter region of clock genes, with respect to the whole mouse genome.

We investigated the following working hypotheses: (i) Circadian genes and genes oscillating with harmonic frequencies show dissimilar biological facets and encode different proteome profiles; (ii) canonical and non-canonical DNA structures are found within the upstream regions of the oscillating genes subsets; (iii) ultradian genes connect to an identified circadian network through distinctive upstream short nucleotide sequences and DNA binding sites. Our results show that the three subsets of oscillating genes are hallmarked by very different functional annotation and proteomic features, as well as peculiar transcription factor binding motives, in addition to canonical binding sites. These are found within the upstream regions of rhythmically expressed target genes and seemingly allow for the link of the ultradian gene sets to a known circadian network.

2. Results

To characterize particular features of the gene sets with ultradian and circadian periodicity (8-h, 12-h, 24-h gene sets), we used a variety of computational and bioinformatics methods including a comprehensive analysis at the gene expression level namely: a sequence analysis for known transcription factor binding sites, multiple sequence alignment and phylogenetic analysis, enrichment analysis of the three gene sets, as well as the analysis of epigenetic and non-epigenetic regulation of oscillating gene expression. We further carried out an analysis at the protein level and investigated the electrochemical properties of oscillating proteins and completed our analysis by generating chromosomal co-localization networks created upon homology mapping of oscillating genes.

2.1. Known Transcription Factor Binding Sites Are Enriched in the Promoter Regions of the 8-h, 12-h, and 24-h Rhytmically Expressed Genes

To characterize a putative differential functionality of ultradian genes, we searched for enriched transcription factor (TF) binding sites in the promoter regions of 8-h and 12-h gene sets as compared to 24-h rhythmically expressed gene set, using the MEME SUITE AME tool [23]. We found several significantly enriched binding sites ($p > 0.05$), Tables S1–S6. The top 5 enriched TF binding sites for the

3500 bp upstream promoter region of the 8-h, 12-h and 24-h gene sets, as well as the 300 bp downstream promoter region enriched TF binding sites are depicted in Figure 1. Interestingly, several of the transcription factors binding to the upstream promoter region are shared between the gene sets, which might point towards common mechanisms controlling the time-dependent expression of these genes. The TF found include MAZ (MYC Associated Zinc Finger Protein) that regulates inflammation-induced expression of serum amyloid A proteins, PATZ1 (POZ/BTB And AT Hook Containing Zinc Finger 1), involved in chromatin modelling and transcription regulation and postulated to be a repressor of gene expression, and ZNF770 (Zinc Finger Protein 770), which were identified in all three gene sets and EGR2 (Early Growth Response 2), VEZF1 (Vascular Endothelial Zinc Finger 1), which are shared between the 12-h and the 24-h rhythmically expressed gene sets. Interestingly, VEZF is a paralog of MAZ. The downstream regions of the gene sets show little correlation. This partially confirms that the results for the upstream region are not occurring by random chance. Additionally, we performed an enrichment analysis for the Reactome pathways related to the TFs with enriched known binding sites (Figure S1). Our results show that the TFs associated with the 12-h gene set show an enrichment of terms related to the cell cycle, and the TFs associated with the 8-h gene set show an enrichment of terms associated with myogenesis and senescence.
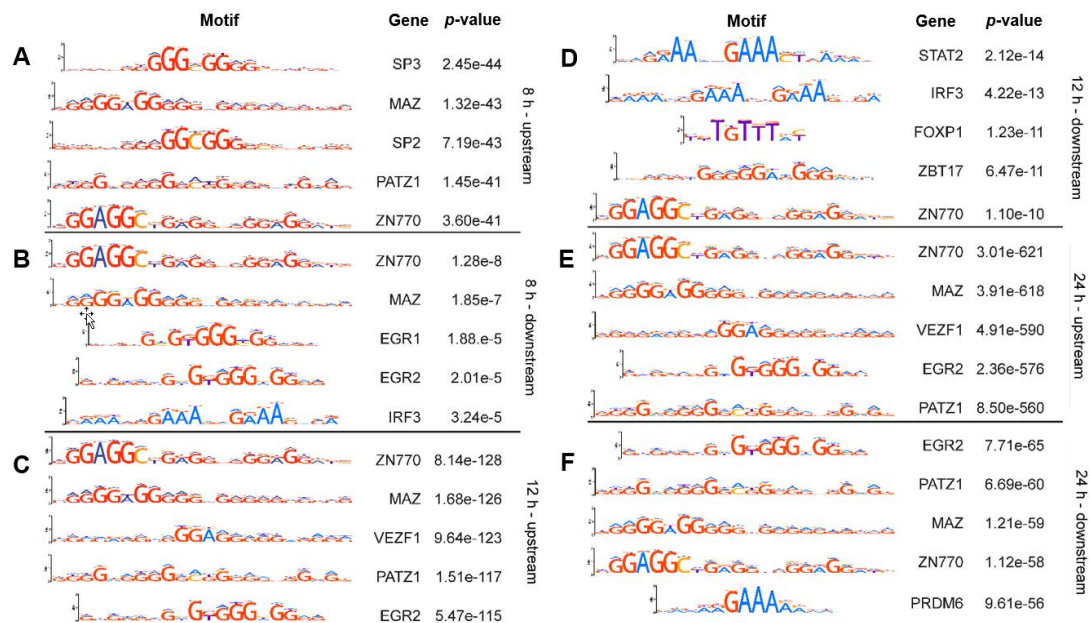

Figure 1. The upstream and downstream regions of the promoters of the different. gene sets show enriched binding sites for known transcription factors. The five transcription factor binding sites with the lowest adjusted *p*-value are displayed for the upstream promoter regions of the (**A**) 8-h (Table S1), (**C**) 12-h (Table S3) (**E**) and 24-h (Table S6) gene sets and for the downstream promoter region of the (**B**) 8-h (Table S2), (**D**) 12-h (Table S4) and (**F**) the 24-h (Table S5) gene sets.

In addition, using the AME tool of the MEME suite we searched for E-boxes and D-boxes, conserved motifs known to be the present in the promoter region of clock-controlled genes, and bound by core-clock elements. Both E-boxes and D-boxes were detected in the upstream promoter region of the 8-h gene set and of the 12-h gene set ($p < 0.05$). However, other motifs are more significantly enriched (Figure 1). We detected E-boxes in the 8-h gene set in 38.3% of the upstream promoter sequences (adj. $p = 4.99e-2$), and in the 12-h gene set in 6.9% of the upstream promoter sequences (adj. $p = 3.75e-8$). In particular, we detected CLOCK (adj. $p = 1.25e-16$) and BMAL1 (adj. $p = 5.83e-06$) binding motifs in the upstream promoter regions of the 12-h gene set (Table S3). For the 24-h gene set we detected E-boxes in 26.7% of the upstream promoter sequences (adj. $p = 3.04e-48$). Likewise, we detected binding motifs for both CLOCK (adj. $p = 4.81e-90$) and BMAL1 (adj. $p = 2.67e-55$) in the upstream promoter regions

of the 24-h gene set (Table S6). Interestingly, CLOCK (adj. $p = 0.02$) and BMAL1 (adj. $p = 3.28e-05$) binding motif are also present in the downstream promoter region of the 24-h gene set (Table S5).

2.2. Phylogenetic Analysis Shows Similarity within the Promoter Regions of the 8 h and 12 h Gene Sets

The significantly enriched motifs found point to a common regulatory system for the 8-h and 12-h gene sets, hence we hypothesized the existence of an evolutionary connection between the promoter regions of both gene sets as the key mechanism of activation is most likely evolutionary ancient and well conserved (as the clock itself). To further investigate this hypothesis we generated phylogenetic trees, as an output visualization of the multiple sequence alignments of the 3500 bp upstream promoter region of the 8-h and 12-h gene sets (Figure 2). First, we produced a multiple sequence alignment of the promoter region sequences together with 10 control sequences of non-oscillating genes. Second, we created phylogenetic trees from the resulting alignment using the Felsenstein (F84) (Figure 2A) and Jukes-Cantor (JC) (Figure 2B) nucleotide substitution models [24]. We tagged the 8-h oscillating genes with red markers in the resulting trees and the 10 control genes with green markers. When utilizing 10 non-oscillating genes as control sequences, the 8-h and 12-h gene sets do not show a clustering pointing at strong evolutionary conservation of the entire sequence regardless of the substitution model used. Hence, while individual binding sites for TFs are highly enriched and conserved, the promoter sequence itself varies greatly and may allow for the fine-tuning of expression for individual genes.
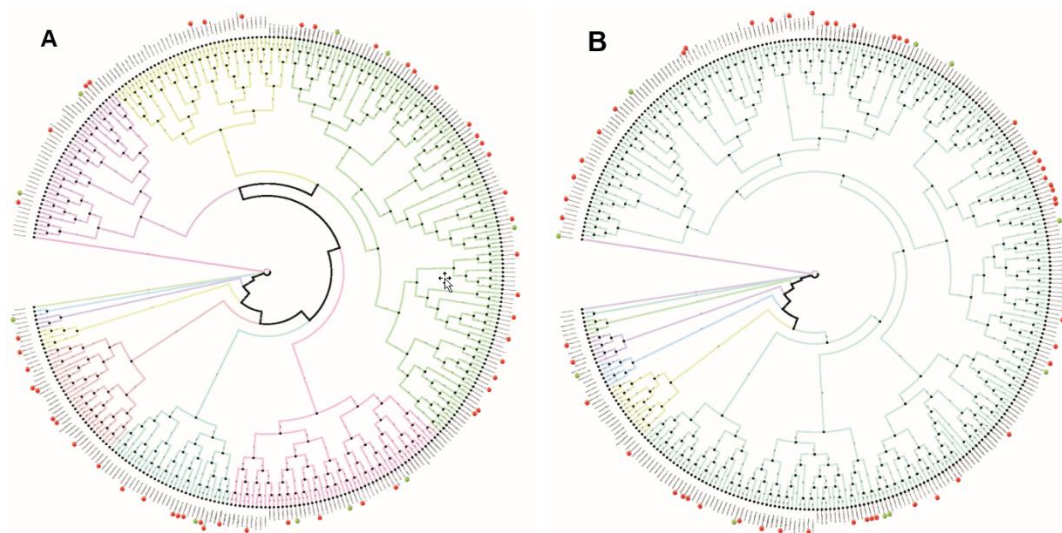

Figure 2. Phylogenetic analysis of the promoter regions shows little variation between the promoters of the corresponding gene sets. A multiple sequence alignment and phylogenetic tree construction was performed on the upstream promoter regions of both the 8-h and the 12-h gene sets with Felsenstein nucleotide substitution model (**A**) and a Jukes-Cantor substitution model (**B**). A control set of 10 non-oscillating genes was added to the phylogenetic analysis (green markers). The red markers represent the position of the 8-h oscillating genes. The 12-h oscillating genes are the unmarked positions. A high-resolution figure is provided as Figure S2.

To search for specific functions of the individual gene sets, we investigated a possible enrichment of Gene Ontology and Reactome Pathways terms for the 8-h, 12-h and 24-h rhythmically expressed genes. In all cases, significant enrichments, generated with ConsensusPathDB, are present ($p < 0.01$). While the 8-h gene set showed an enrichment of terms related to metabolism, the 12-h set showed an enrichment of terms related to endoplasmic reticulum (ER)-related processes, splicing, translation and gene expression regulation. The 24-h rhythmically expressed genes showed an enrichment of terms related to meiosis, and splicing (Figure S3), which is in line with our previous findings [25–28].

We further explored the putative connection of the ultradian gene sets to a known circadian (approximately 24 h rhythmically expressed elements) network (NCRG-network of circadian regulated genes [10]) for that, we performed a series of simulations based on randomized protein-protein interaction networks. The random network generation is based on the IntAct database contained in iRefIndex. We quantified the number of interactions between the elements of the 8-h gene set, and 100 random networks of the same size as the NCRG. In addition, we also quantified the number of interactions between the 8-h gene set and the NCRG. While the average interactions between the 8-h gene set and the random networks was 3.99 ± 2.91 (mean and SD), the number of interactions between the same gene set and the NCRG was 12. We applied the same procedure to the 12-h gene set and obtained 19 ± 9.25 (mean and SD) connections to the random networks, while the number of connections between the NCRG and the 12 h gene set was 87. Both sets therefore exhibit a connectivity to the NCRG that was higher than the connectivity displayed by the random gene sets. Thus, the randomized network analysis points to a connection between the ultradian rhythmically expressed genes, the core-clock and clock-controlled genes.

2.3. Epigenetic and Non-epigenetic Regulation of Oscillating Gene Expression

We performed an enrichment analysis for histone modifications associated with the 8-h, 12-h and 24-h gene sets available in the public Encode 2015 project data [29]. The enriched histone methylation pattern of H3K79me2 associated with the 12-h and 24-h gene sets is tissue-specific for the liver in agreement with the original data, generated from liver cells. This points to a tissue specificity of the methylation pattern and the corresponding expression of rhythmically expressed genes. The methylation pattern of the 8-h gene set is associated with a different cell line (*M. Musculus* MEL, $p = 3.555e-02$).

The H3K79me2 histone modification is associated with the function of the RNA polymerase II (Figure S4). RNA Polymerase II plays a major role in the transcription regulation of the 12-h and 24-h rhythmically expressed genes based on ChIPSeq data from the ENCODE project and as suggested by the previous histone modification data. However, GABPA (GA Binding Protein Transcription Factor Subunit Alpha, $p = 4.315e-08$) TF scores higher in the 12-h oscillating gene set than RNA Polymerase II. GABPA is related to the mitochondrial gene expression pathway, thus pointing again at the potential metabolic role of the genes with ultradian oscillations. For the 8-h gene set we detected an enrichment of TCF12 (Transcription Factor 12) binding. This transcription factor recognizes E-boxes and is involved in the formation of lineage-specific gene expression. This enrichment illustrates the important role of the E-boxes, which even though not being the most enriched motif seem to attract the strongest TF activity (Figure S4).

The enrichment analysis for computationally predicted miRNA targets from the TargetSCAN 2017 database provides significant results for the 24-h and 8-h gene sets. For the 8-h gene set the targets for miRNA 1295 are the dominant signal ($p = 6.563e-03$). Computationally determined targets for the miRNA 4637 are enriched in the 24-h gene set ($p = 1.630e-05$), while in the 12-h gene set four microRNAs are enriched: miRNA 344 ($p = 6.117e-04$), miRNA 344c ($p = 6.117e-04$), miRNA 1244 ($p = 7.081e-04$) and miRNA 499 ($p = 1.096e-03$) Figure S4.

Moreover, we investigated the protein-protein interactions of the transcription factors potentially influencing the gene sets according to the ENRICHR database [30]. This enrichment showed interesting candidates–POLE (DNA polymerase Epsilon, Catalytic Subunit, $p = 1.026e-02$) for the 8-h gene set and ESR1 (Estrogen receptor 1) for the 12-h ($p = 1.137e-07$) and 24-h ($p = 7.235e-13$) gene sets. Figure S4. We further investigated the role of POLE in a potential cancer context. From publicly available data the high expression of POLE is an unfavourable marker in renal cancer and melanoma [31] (Figure S5).

Altogether, the enrichment information points towards very specific processes that govern the regulation and output of the ultradian oscillating genes. Often a single microRNA (such as miRNA-1295 for the 8-h gene set) or a single gene such as POLR2A in the 24-h gene set are predicted to have

the most significant results in terms of interaction with other genes, regulation of transcription or computationally predicted targets in the genomic sequence.

2.4. Electrochemical Properties of Oscillating Proteins

Next, we analyzed the electrochemical features of the proteins encoded by circadian and ultradian genes as compared to a randomly sorted set of proteins encoded by non-oscillating genes (Table 1, Table S7). The overall stability as predicted by the FoldX algorithm [32] was higher in oscillating proteins when compared to non-oscillating proteins (although with a barely detectable statistically significant difference), whereas the terms represented by interresidue Van der Waals' clashes and electrostatic interaction (Table S8) between molecules in the precomplex were significantly different. A correlation analysis showed negative correlation of these terms with free energy values in oscillating proteins and positive correlation in non-oscillating proteins, suggesting a different contribution to the overall protein stability.

Among the oscillating proteins subsets, 8-h oscillating proteins showed statistically significant differences in respect to 12-h and 24-h oscillating proteins, with lower average number of residues and with higher free energy (lower energy of unfolding), suggesting lower overall stability. In this regard, the components that were different in a statistically significant way were represented by solvation of polar and hydrophobic atoms, water binding, Van der Waals energy, steric clashes, hydrogen bonds, electrostatic interactions (Table 1). Gibbs free energy was negatively correlated with these statistically significant variables in the 8-h oscillating proteins, while an inverse correlation was found in the set of 12-h oscillating proteins, hinting at a diverse involvement in the net equilibrium of forces settling on unfolded or folded protein state. On the other hand, similar correlations were found for 8-h and 24-h oscillating proteins, except for the contribution of hydrophobic groups to free energy difference (Figure 3 and Figure S6).
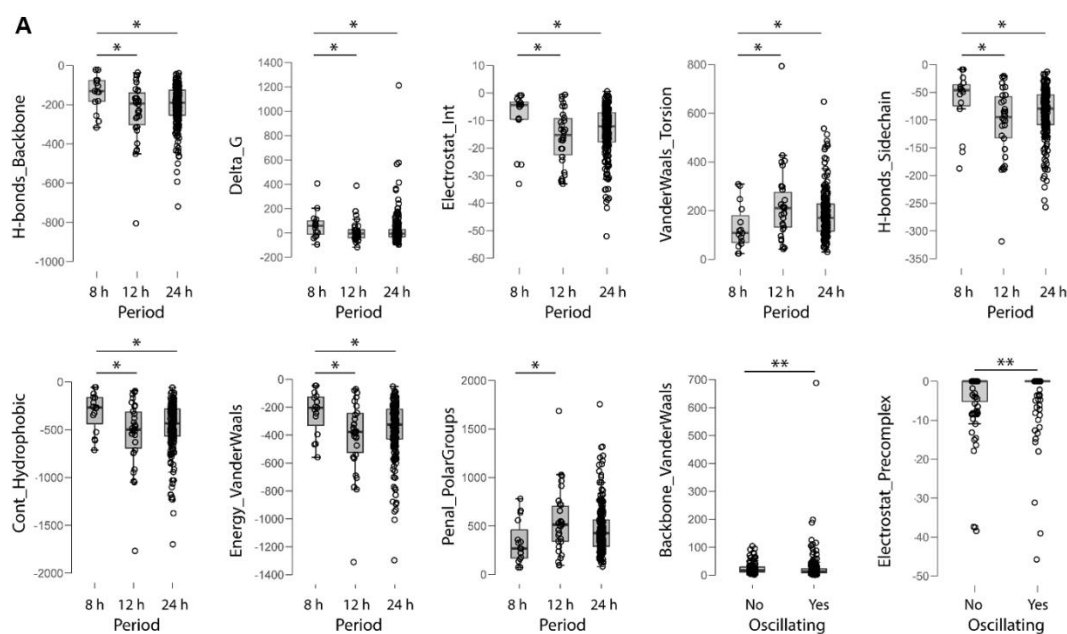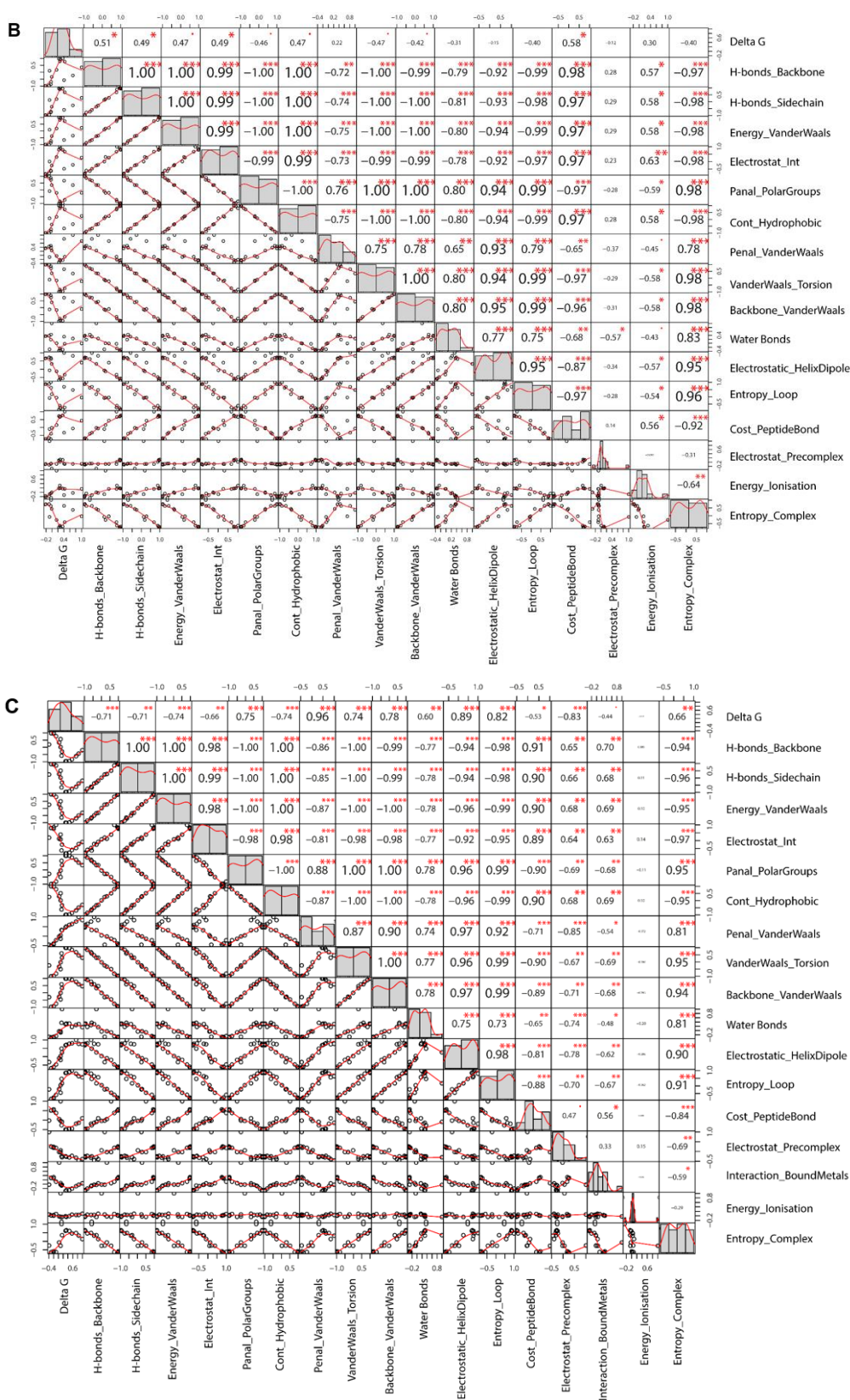

Figure 3. *Cont.*

Figure 3. Electrochemical properties. (**A**) Box plots rendering the interquartile range (IQR) and the horizontal bar the median relative expression. Expression values that do not fall within 1.5 x IQR are outliers and are indicated by circles where appropriate. (**B**) Correlation matrix for parameters of non-oscillating genes. Numbers in the boxes are Spearman's correlation coefficients. (**C**) Correlation matrix for parameters of oscillating genes. Numbers in the boxes are Spearman's correlation coefficients. Asterisks represent statistically significant correlations (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Table 1. Descriptive statistics of energy values as from the FOLD-X energy function (mean ± SD) and results of Kruskal–Wallis one-way analysis of variance with Residue Number as covariate and Dunn's post hoc test with false discovery rate (FDR) correction.

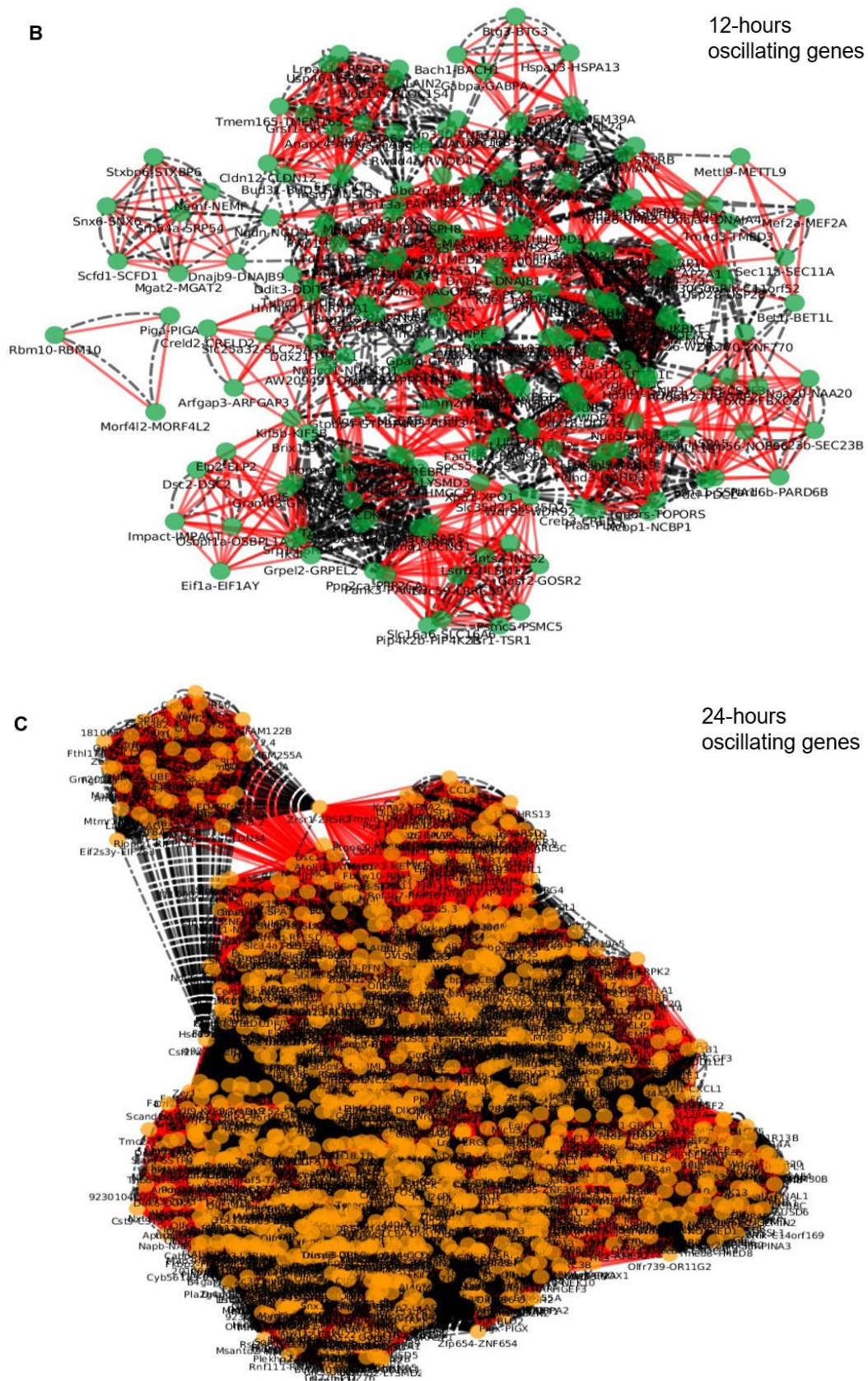| Parameter	Oscillating Proteins	Non-Oscillating Proteins	*p*	Period 8 h	Period 12 h	Period 24 h	*p-value*
Residue Number	323.7 ± 193.6	313.3 ± 165.5	0.642	224.0 ± 136.6	373.3 ± 232.1	321.8 ± 188.5	**0.044*** (8 h vs. 12 h) 0.063 (8 h vs. 24 h) 0.347 (12 vs. 24 h)
Delta_G	23.34 ± 123.04	29.84 ± 91.09	0.065	81.91 ± 133.64	13.53 ± 95.47	21.64 ± 125.74	**0.047*** (8 h vs. 12 h) **0.047*** (8 h vs. 24 h) 0.377 (12 vs. 24 h)
H-bonds_Backbone	−206.6 ± 112.7	−209.4 ± 121.7	0.429	−135.4 ± 95.49	−236.2 ± 157.46	−206 ± 104.26	**0.022*** (8 h vs. 12 h) **0.022*** (8 h vs. 24 h) 0.221 (12 vs. 24 h)
H-bonds_Sidechain	−88.50 ± 49.24	−89.31 ± 52.65	0.496	−61.66 ± 56.19	−103.98 ± 65.35	−87.67 ± 45.53	**0.009**** (8 h vs. 12 h) **0.016*** (8 h vs. 24 h) 0.125 (12 vs. 24 h)
Energy_VanderWaals	−355.8 ± 201.6	−363.0 ± 211.2	0.465	−239.1 ± 168.5	−407.4 ± 260.7	−354.5 ± 191.4	**0.033*** (8 h vs. 12 h) **0.043*** (8 h vs. 24 h) 0.175 (12 vs. 24 h)
Electrostat_Int	−13.75 ± 9.226	−14.03 ± 10.912	0.395	−8.291 ± 9.198	−16.011 ± 9.858	−13.710 ± 9.049	**0.006**** (8 h vs. 12 h) **0.014*** (8 h vs. 24 h) 0.091 (12 vs. 24 h)
Penal_PolarGroups	474.9 ± 267.9	486.5 ± 278.3	0.401	335.0 ± 232.3	541.5 ± 339.8	472.6 ± 255.7	**0.049*** (8 h vs. 12 h) 0.071 (8 h vs. 24 h) 0.162 (12 vs. 24 h)
Cont_Hydrophobic	−472.4 ± 268.4	−480.7 ± 281.7	0.470	−311.7 ± 219.0	−542.3 ± 350.6	−470.8 ± 254.0	**0.029*** (8 h vs. 12 h) **0.037*** (8 h vs. 24 h) 0.180 (12 vs. 24 h)
Penal_VanderWaals	23.16 ± 49.59	24.85 ± 21.93	**0.002***	19.94 ± 27.56	22.23 ± 23.03	23.46 ± 53.31	0.444 (8 h vs. 12 h) 0.444 (8 h vs. 24 h) 0.444 (12 vs. 24 h)
VanderWaals_Torsion	198.7 ± 108.4	191.1 ± 109.7	0.496	126.9 ± 91.05	223.1 ± 151.23	188.2 ± 100.39	**0.023*** (8 h vs. 12 h) **0.038*** (8 h vs. 24 h) 0.140 (12 vs. 24 h)
Backbone_VanderWaals	470.6 ± 272.1	485.1 ± 258.7	0.252	352.1 ± 228.7	532.7 ± 331.3	467.9 ± 263.5	0.081 (8 h vs. 12 h) 0.130 (8 h vs. 24 h) 0.141 (12 vs. 24 h)
Water Bonds	0.699 ± 1.158	0.693 ± 1.024	0.458	0.237 ± 0.451	0.550 ± 0.742	0.745 ± 1.227	0.343 (8 h vs. 12 h) 0.259 (8 h vs. 24 h) 0.343 (12 vs. 24 h)
Electrostatic_HelixDipole	11.66 ± 11.484	12.04 ± 8.639	0.095	8.414 ± 8.079	12.091 ± 9.241	11.769 ± 11.938	0.185 (8 h vs. 12 h) 0.185 (8 h vs. 24 h) 0.284 (12 vs. 24 h)
Cost_PeptideBond	−5.348 ± 4.489	−5.511 ± 6.128	0.183	−3.399 ± 3.684	−5.705 ± 5.276	−5.401 ± 4.399	0.131 (8 h vs. 12 h) 0.126 (8 h vs. 24 h) 0.496 (12 vs. 24 h)
Electrostat_Precomplex	−1.143 ± 4.972	−3.606 ± 7.590	**<0.001***	0.000 ± 0.000	0.000 ± 0.000	−1.373 ± 5.422	0.500 (8 h vs. 12 h) 0.213 (8 h vs. 24 h) 0.073 (12 vs. 24 h)
Interaction_BoundMetals	−0.016 ± 0.155	0.000 ± 0.000	0.200	0.000 ± 0.000	−0.053 ± 0.288	−0.011 ± 0.129	0.594 (8 h vs. 12 h) 0.594 (8 h vs. 24 h) 0.594 (12 vs. 24 h)
Energy_Ionisation	−4.988 ± 12.64	−6.075 ± 15.55	0.473	−1.984 ± 4.856	−8.547 ± 14.491	−4.628 ± 12.589	0.444 (8 h vs. 12 h) 0.444 (8 h vs. 24 h) 0.444 (12 vs. 24 h)
Entropy_Complex	1.220 ± 0.957	1.108 ± 0.820	0.226	0.865 ± 0.876	1.379 ± 1.0124	1.215 ± 0.951	0.134 (8 h vs. 12 h) 0.146 (8 h vs. 24 h) 0.223 (12 vs. 24 h)

* $p < 0.05$; ** $p < 0.01$.

2.5. Chromosome Mapping of Oscillating Genes

All 8-h subset (56) and nearly all 12-h subset of mouse genes (202 out of 205) were mapped to human homologs, while only 1826 out of 2054 mouse circadian genes were suitably mapped. The genes of the three classes were distributed along all chromosomes, with no chromosome left uncovered and no homolog and paralog gene localized on the same chromosome both in human and in mouse (Figure 4). Only a few oscillating genes mapped to chromosome Y, precisely one circadian gene in the mouse set, and one ultradian and one circadian gene in human set (Table S9). The intersection of mouse and human co-localization networks created upon homology mapping of oscillating genes revealed high localization conservation for the 8-h gene sets between both species (65%), a moderate conservation for 12-h gene sets (23%) and poor conservation of chromosomal localization for circadian genes (6%) (Table 2).

Table 2. Topological features of *M. musculus* and *H. sapiens* chromosomal co-localization networks created upon homology mapping of oscillating genes.

M. musculus Network Specifics (after Isolate Nodes Removal)			
Network	**Nodes**	**Edges**	**Components**
8 h	51	89	14
12 h	199	1001	20
24 h	1827	105970	20
H. sapiens Network Specifics (after Isolate Nodes Removal)			
8 h	49	92	14
12 h	198	1085	23
24 h	1826	95657	23
Intersection Networks Characteristics			
8 h	33	29	13
12 h	45	169	312
24 h	103	1801	33439

Figure 4. *Cont.*

Figure 4. Chromosome mapping with genomic clustering of oscillatory genes in *M. musculus* and *H. sapiens* chromosomes for (**A**) 8-h, (**B**) 12-h and (**C**) 24-h oscillating genes (the cytoscape files are provided as Figure S7.

3. Discussion

Frequency multiplication is a common occurrence in rhythmic phenomena observed in multifaceted systems of interest for a variety of scientific disciplines, for instance physics, chemistry, biology, astronomy. In natural and life sciences, harmonics of circadian frequency have been initially reported prior to the foundation of chronobiology as a separate area of scientific research addressing rhythmic phenomena in living beings. Nonetheless, the scientific literature on the multiplication of circadian periodicity in biological processes remains limited at the present time.

The comprehensive bioinformatics analyses performed on transcriptomics and proteomics data in mammalian genes expressed with 24-h periodicity and with harmonics of circadian rhythmicity allowed us to highlight a number of interesting differences among the subsets of oscillating genes: (i) circadian genes and genes oscillating at the second and third harmonic of 24-h periodicity show divergent functional annotation and proteomic characteristics; (ii) within their upstream regions unusual transcription factor binding motives other than canonical binding sites are found; (iii) genes oscillating at the second and third harmonics are connected by specific regulatory motifs and transcription factor binding sites to a recognized circadian network.

In particular, concerning shared enriched transcription factor binding sites in the promoter regions of the circadian and ultradian genes suggest equivalent transcriptional control of time-dependent gene expression. In the upstream promoter region of ultradian genes, in addition to other motifs more significantly enriched, we identified E-boxes and D-boxes, which were not found in their downstream promoter regions. Moreover, the phylogenetic analysis of the promoter regions of the ultradian gene sets showed variability of the entire promoter sequence, which could eventually allow the accurate regulation of expression of the different genes. Furthermore, a randomized network analysis suggested a possible connection between the ultradian genes subset and the circadian clock circuitry. The subsequent enrichment analysis showed that the 8-h oscillating genes were enriched in terms related to metabolism, the 12-h oscillating genes in terms related to ER-related processes, splicing, translation and gene expression regulation, while the 24-h oscillating genes in terms related to meiosis, and splicing. This is in agreement with previous results [25,26,33].

3.1. Oscillating Proteins Are Hallmarked by Higher Overall Stability when Compared to Non-Oscillating Proteins

Bioinformatics analysis of the electrochemical properties of non-oscillating and oscillating proteins showed that the oscillating proteins are hallmarked by higher overall stability when compared to non-oscillating proteins, mainly in relation to significant dissimilarity of two components of free energy calculation in the FoldX protein design algorithm, one related to inter-residue close contacts and the other represented by electrostatic contribution of interactions at interfaces, which differently contributed to the free energy value in the two subsets. Considering the three oscillating proteins subsets, 8-h oscillating proteins showed lower mean residue number and lower overall stability, mainly in relation to different polar and hydrophobic desolvation, water binding, Van der Waals energy, steric clashes, hydrogen bonds, electrostatic interactions, interestingly with opposite correlations when matched up to the other ultradian subset of proteins. Protein folding allows free volume to decrease and considerably impacts protein conformational/binding equilibrium and ultimately physiological function in conditions of macromolecular crowding, such as those hallmarking cellular and sub-cellular volume-restricted compartments [34,35]. The spatio-temporal gathering of oscillating proteins may impact the effects of macromolecular crowding on equilibrium stability of proteins with different folds, cofactors and mechanisms. Protein folding and unfolding kinetics are influenced by crowding, with stabilizing effects whose degree will hinge on intrinsic stability and protein fold [34,35]. In this context, the molecular clockwork could manage the phase relation between subcellular oscillation patterns of folded, intermediate, and unfolded proteins, as well as of molecular chaperones that assist these transitions, especially considering that macromolecular crowding accelerates folding, but over a given limit the folding process will be hindered [34,35].

3.2. Specific Enriched Processes Govern the Regulation and Output of the Ultradian Oscillating Genes

The enrichment analysis for histone modifications showed association with the 12-h and 24-h oscillating genes of H3K79me2, involved in RNA polymerase II function, whereas 8-h oscillating genes showed binding enrichment for TCF12, a transcription factor capable of binding to E-boxes. Altogether, the enrichment information points towards very specific processes that govern the regulation and output of the ultradian oscillating genes. Often a single microRNA (such as miRNA-1295 for the 8 h gene set) or a single gene such as POLR2A in the 24-h gene set are predicted to have the most significant result in terms of targets. The enrichment analysis for computationally predicted miRNA targets pinpointed to the 8-h gene set as targets for miRNA-1295, to the 12-h gene set for miRNA 344, miRNA 344c, miRNA 1244 and miRNA 499, whereas the 24-h oscillating genes appeared as targets for miRNA-4637. Furthermore, the protein-protein interactions of the transcription factors potentially influencing the oscillating gene sets identified as major candidates POLE for the 8-h gene set and ESR1 for the 12-h and 24-h gene sets. Interestingly, elevated POLE expression predicts poorer outcome marker in renal cancer and melanoma patients.

3.3. Homology Mapping of Oscillating Genes Revealed Different Degree of Localization Conservation for the Three Gene Sets

Mapping of the 8-h, 12-h and 24-h oscillating genes in mouse and human chromosomes revealed scattering of the three classes along all chromosomes, with no chromosome left uncovered and homologs and paralogs of core-clock genes and clock-controlled genes never localized on the same chromosome. Nevertheless, in both species only a few oscillating genes mapped to chromosome Y, probably in relation to the peculiar role played by this allosome in male fertility and sex determination in mammals. In addition, we found high localization conservation for the 8-h genes (65%) between both species, a moderate conservation for 12-h genes (23%) and a poor conservation of localization for circadian genes (6%).

4. Materials and Methods

4.1. Primary Dataset

Bioinformatics analyses were performed on publicly available genomic data (GSE11923). Briefly, liver samples were collected every hour for 48 h from *n* = 3-5, 6-week-old male C57BL/6J mice (Jackson) per time point, the specimens were pooled, and high-temporal resolution profiling was performed using Affymetrix arrays to detect cycling genes. Fisher's G-test at a false-discovery rate of < 0.05 and COSOPT were jointly exploited to recognize rhythmic transcripts, which were classified, depending on the length of the oscillation period, as circadian (24 ± 4 h) and ultradian (12 ± 2 h and 8 ± 1 h) [15]. Array probe IDs/nucleotide sequences of 8-h, 12-h and 24-h oscillating genes were registered and by using BioDBnet (https://biodbnet-abcc.ncifcrf.gov/db/db2db.php) 56, 202 and 2396 Ensembl Transcript IDs were recovered from the primary dataset, respectively.

4.2. Sequence Analysis for Known Transcription Factor Binding Sites

For the initial data acquisition, we performed an analysis on pre-selected data sets corresponding to the above-mentioned gene-probes with 8-h, 12-h and 24-h rhythmic oscillations. To perform the sequence analysis, we extracted and analyzed the 3500 bp flanking sequences upstream and the 300 bp flanking sequences downstream of the complete corresponding genes. The mapping and sequence selection were carried out with Ensembl biomaRt (Ensembl revision 84). We searched for enriched known motifs and specific acceptance for gapped motifs with the MEME SUITE software (http://meme-suite.org/) [36]. The length of the motif correlates with its statistical significance. MEME defines the most statistically significant motif based on its E-value (low E-value). The E-value of a motif is based on its log likelihood ratio, width, sites, the background letter frequencies, and the size of the training set. The E-value is an estimate of the expected number of motifs with the given log

likelihood ratio (or higher), and with the same width and site count, as found in a similarly sized set of random sequences. We used the AME tool and the HOCOMOCOv11 [37,38] database as motif sources. The AME tool specifically searches for enrichments of known motifs from the database selected. The 3500 bp upstream promoter region was scanned, as well as the 300 bp downstream motif region.

4.3. Multiple Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignments of the promoter regions of the 8-h and 12-h gene sets were created with MUSCLE (http://www.drive5.com/muscle/). The resulting alignments were used for further phylogenetic analysis of the promoter regions of the 8-h and 12-h gene sets. Phylogenetic tree creation was performed with PHYLIP's neighbor joining method with F84 and Jukes-Cantor substitution models [24] (http://evolution.genetics.washington.edu/phylip.html).

4.4. Enrichment Analysis of the 8-h, 12-h and 24-h Gene Sets

The enrichment analysis for the mouse gene sets was performed with ConsensusPathDB (http://consensuspathdb.org) [39]. An analysis for Enriched Reactome and GO terms was performed. The *p*-value cut-off was set to 0.01 and the GO terms were set to level 4. Each of the 8-h, 12-h and 24-h gene sets was analyzed individually.

4.5. Randomized Network Analysis on Ultradian and Circadian Genes

A network analysis to explore the connection between the ultradian genes, and the core-clock and clock-controlled genes was performed through a series of simulations based on randomized protein-protein interaction networks. For the network creation the probes were mapped to Uniprot and Entrez ID. The network was generated from IntAct data contained in the iRefIndex database (snapshot from 2015) (http://irefindex.org) which summarizes protein-protein interaction data from different sources. The computations were performed using the iRefR R package with 100 random sets of genes of matching size.

4.6. Epigenetic and Non-epigenetic Regulation of Oscillating Gene Expression

All epigenetic and non-epigenetic enrichment analysis was performed with the ENRICHR tool (http://amp.pharm.mssm.edu/Enrichr/enrich), [40] which utilizes publicly available information from projects such as TargetSCAN 2017 (http://www.targetscan.org/vert_72/) and ENCODE 2015 [41]. All enrichment results are *p*-value sorted.

4.7. Impact of Protein Expression on Survival

Survival data associated with protein expression was retrieved from the Protein Atlas database [31].

4.8. Electrochemical Properties of Oscillating Proteins

To predict the electrochemical properties of proteins encoded by ultradian and circadian genes we used the corresponding three-dimensional structural data stored in the protein data bank (PDB; http://www.rcsb.org/pdb/) [42]. All complex analyses were performed with FoldX, which is one of the best stability predictors and is easily implementable in a pipeline [32]. FoldX is an empirical force field that was developed or the rapid evaluation of the stability, folding of proteins and nucleic acids. It is composed of a solvation term, a van der Waals term, H-bond, and electrostatic terms and entropic terms for the backbone and side chains. In the case of protein complexes, an extra term related to the electrostatic contribution is also considered. The software package FoldX includes subroutines, e.g., RepairPDB. The way it operates is the following: first it looks for all Asparagine, Glutamine and Histidine residues and flips them by 180 degrees. This is done to prevent incorrect rotamer assignment in the structure due to the fact that the electron density of Asparagine and Glutamine carboxamide

groups is almost symmetrical and the correct placement can only be discerned by calculating the interactions with the surrounding atoms. The same applies to Histidine. It does a small optimization of the side chains to eliminate small VanderWaals' clashes. This way it prevents moving side chains in the final step. "RepairPDB" identifies the residues that have very bad energies and mutates them and their neighbors' to themselves exploring different rotamer combinations to find new energy minima. Correlations between the parameters were investigated by pairwise correlation analysis (Spearman correlation; R package PerformanceAnalytics). The statistical analysis for electrochemical features of the ultradian and circadian gene sets was conducted using the energy values as from the FOLD-X [32] energy function and performing a Kruskal-Wallis one-way analysis of variance with Residue Number as covariate and Dunn's post hoc test with false discovery rate (FDR) correction.

4.9. Chromosome Mapping of Oscillating Genes

H. sapiens homologs for 8-h, 12-h and 24-h *M. musculus* oscillating genes (GSE11923) were retrieved using biomaRt. Genes not matching between mouse and humans were sought manually using the latest version of Ensembl web portal. In case of multiple homologs (one-to-many or many-to-many relationships), the following scores were considered, in this order of priority: confidence score; gene order conservation (GOC) score; target %ID, which refers to the percentage of the sequence in the target species (human) that matches to the query sequence (mouse); query % ID, which refers to the percentage of the sequence in the query species that matches to the homologue; dN/dS ratio. The number of oscillating genes divided by the total number of genes in each chromosome was represented by bar plots. The extent of gene co-localization overlap was assessed by using networks. Genes of the three subsets, i.e., 8-h, 12-h and 24-h oscillating genes, were represented as networks, where the genes, symbolized as nodes, were linked by edges if they were located on the same chromosomes. For each mouse and human subset of genes networks were built and then intersected by means of Pyntacle (http://pyntacle.css-mendel.it/). An intersection network was built considering only nodes and edges in common between the two original networks. Pairs of intersecting genes were considered to be on the same chromosome in mouse and human, even if the chromosome were not the same between the two species.

5. Conclusions

High-throughput analysis over time-series microarray expression data unveils harmonics in oscillation patterns of omics that, intermingling with spatial hierarchical branching networks lapsing in size-invariant units, could endow a fourth temporal dimension at least complementary to the fourth spatial dimension blueprinted by fractal-like networks broadly pervasive in nature. Our wide-ranging characterization of genomic, proteomic and functional properties of oscillating genes and proteins suggests that ultradian and circadian rhythmicity in omics could subtend or alternatively be related to specific mechanisms underlying the functioning of various and complex biological phenomena crucial to make life possible.

set. Table S7. Descriptive statistics for the data in Table 1. Table S8. Free energy (ΔG) terms included in the core function of FoldX, the empirical force field algorithm aiming to calculate the change of ΔG in kcalmol^{-1}. Table S9. Chromosome mapping of oscillating genes.

Author Contributions: Conceptualization, G.M., A.R.; methodology, T.M., N.G., J.R., M.T., A.C., E.M.T., D.C., T.B., S.C.; formal analysis, T.M., N.G., J.R., M.T., A.C., E.M.T., D.C., T.B., S.C.; investigation, G.M.,T.M., A.R., N.G., M.T., A.C., E.M.T., J.R., D.C., T.B., S.C., F.S.; resources, G.M., A.R.; writing—original draft preparation, G.M., S.C., A.R., N.G.; writing—review and editing, G.M.,T.M., A.R., N.G., M.T., A.C., E.M.T., J.R., D.C., T.B., S.C., F.S.; supervision, G.M., A.R.; funding acquisition, G.M., A.R.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Takahashi, J.S.; Hong, H.K.; Ko, C.H.; McDearmon, E.L. The genetics of mammalian circadian order and disorder: Implications for physiology and disease. *Nat. Rev. Genet.* **2008**, *9*, 764–775. [CrossRef] [PubMed]

2. Bass, J.; Takahashi, J.S. Circadian integration of metabolism and energetics. *Science* **2010**, *330*, 1349–1354. [CrossRef] [PubMed]

3. Dibner, C.; Schibler, U.; Albrecht, U. The mammalian circadian timing system: Organization and coordination of central and peripheral clocks. *Annu. Rev. Physiol.* **2010**, *72*, 517–549. [CrossRef] [PubMed]

4. Dunlap, J.C. Molecular bases for circadian clocks. *Cell* **1999**, *96*, 271–290. [CrossRef]

5. Mermet, J.; Yeung, J.; Naef, F. Systems Chronobiology: Global Analysis of Gene Regulation in a 24-Hour Periodic World. *Cold Spring Harb. Perspect. Biol.* **2017**, *9*, a028720. [CrossRef] [PubMed]

6. Gerber, A.; Saini, C.; Curie, T.; Emmenegger, Y.; Rando, G.; Gosselin, P.; Gotic, I.; Gos, P.; Franken, P.; Schibler, U. The systemic control of circadian gene expression. *Diabetes Obes. Metab.* **2015**, *17* (Suppl. S1), 23–32. [CrossRef]

7. Bozek, K.; Relogio, A.; Kielbasa, S.M.; Heine, M.; Dame, C.; Kramer, A.; Herzel, H. Regulation of clock-controlled genes in mammals. *PLoS ONE* **2009**, *4*, e4882. [CrossRef]

8. Lowrey, P.L.; Takahashi, J.S. Genetics of circadian rhythms in Mammalian model organisms. *Adv. Genet.* **2011**, *74*, 175–230.

9. Zhang, R.; Lahens, N.F.; Ballance, H.I.; Hughes, M.E.; Hogenesch, J.B. A circadian gene expression atlas in mammals: Implications for biology and medicine. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 16219–16224. [CrossRef]

10. Lehmann, R.; Childs, L.; Thomas, P.; Abreu, M.; Fuhr, L.; Herzel, H.; Leser, U.; Relogio, A. Assembly of a comprehensive regulatory network for the mammalian circadian clock: A bioinformatics approach. *PLoS ONE* **2015**, *10*, e0126283. [CrossRef]

11. Hirano, A.; Fu, Y.H.; Ptacek, L.J. The intricate dance of post-translational modifications in the rhythm of life. *Nat. Struct. Mol. Biol.* **2016**, *23*, 1053–1060. [CrossRef] [PubMed]

12. Ye, R.; Selby, C.P.; Chiou, Y.Y.; Ozkan-Dagliyan, I.; Gaddameedhi, S.; Sancar, A. Dual modes of CLOCK: BMAL1 inhibition mediated by Cryptochrome and Period proteins in the mammalian circadian clock. *Genes Dev.* **2014**, *28*, 1989–1998. [CrossRef] [PubMed]

13. Takahashi, J.S. Transcriptional architecture of the mammalian circadian clock. *Nat. Rev. Genet.* **2017**, *18*, 164–179. [CrossRef] [PubMed]

14. Aryal, R.P.; Kwak, P.B.; Tamayo, A.G.; Gebert, M.; Chiu, P.L.; Walz, T.; Weitz, C.J. Macromolecular Assemblies of the Mammalian Circadian Clock. *Mol. Cell* **2017**, *67*, 770–782 e6. [CrossRef] [PubMed]

15. Hughes, M.E.; DiTacchio, L.; Hayes, K.R.; Vollmers, C.; Pulivarthy, S.; Baggs, J.E.; Panda, S.; Hogenesch, J.B. Harmonics of circadian gene transcription in mammals. *PLoS Genet.* **2009**, *5*, e1000442. [CrossRef] [PubMed]

16. Gachon, F.; Nagoshi, E.; Brown, S.A.; Ripperger, J.; Schibler, U. The mammalian circadian timing system: From gene expression to physiology. *Chromosoma* **2004**, *113*, 103–112. [CrossRef] [PubMed]

17. Fuhr, L.; Abreu, M.; Pett, P.; Relogio, A. Circadian systems biology: When time matters. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 417–426. [CrossRef] [PubMed]

18. Westermark, P.O.; Herzel, H. Mechanism for 12 hr rhythm generation by the circadian clock. *Cell Rep.* **2013**, *3*, 1228–1238. [CrossRef]

19. Martins, B.M.; Das, A.K.; Antunes, L.; Locke, J.C. Frequency doubling in the cyanobacterial circadian clock. *Mol. Syst. Biol.* **2016**, *12*, 896. [CrossRef]

20. Shinohara, Y.; Koyama, Y.M.; Ukai-Tadenuma, M.; Hirokawa, T.; Kikuchi, M.; Yamada, R.G.; Ukai, H.; Fujishima, H.; Umehara, T.; Tainaka, K.; et al. Temperature-Sensitive Substrate and Product Binding Underlie Temperature-Compensated Phosphorylation in the Clock. *Mol. Cell* **2017**, *67*, 783–798 e20. [CrossRef]

21. Isojima, Y.; Nakajima, M.; Ukai, H.; Fujishima, H.; Yamada, R.G.; Masumoto, K.H.; Kiuchi, R.; Ishida, M.; Ukai-Tadenuma, M.; Minami, Y.; et al. CKIepsilon/delta-dependent phosphorylation is a temperature-insensitive, period-determining process in the mammalian circadian clock. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15744–15749. [CrossRef] [PubMed]

22. Castellana, S.; Mazza, T.; Capocefalo, D.; Genov, N.; Biagini, T.; Fusilli, C.; Scholkmann, F.; Relogio, A.; Hogenesch, J.B.; Mazzoccoli, G. Systematic Analysis of Mouse Genome Reveals Distinct Evolutionary and Functional Properties Among Circadian and Ultradian Genes. *Front. Physiol.* **2018**, *9*, 1178. [CrossRef] [PubMed]

23. Machanick, P.; Bailey, T.L. MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* **2011**, *27*, 1696–1697. [CrossRef] [PubMed]

24. Arenas, M. Trends in substitution models of molecular evolution. *Front. Genet.* **2015**, *6*, 319. [CrossRef] [PubMed]

25. El-Athman, R.; Fuhr, L.; Relogio, A. A Systems-Level Analysis Reveals Circadian Regulation of Splicing in Colorectal Cancer. *EBioMedicine* **2018**, *33*, 68–81. [CrossRef] [PubMed]

26. Genov, N.; Basti, A.; Abreu, M.; Astaburuaga, R.; Relogio, A. A bioinformatic analysis identifies circadian expression of splicing factors and time-dependent alternative splicing events in the HD-MY-Z cell line. *Sci. Rep.* **2019**, *9*, 11062. [CrossRef] [PubMed]

27. Genov, N.; Basti, A.; Abreu, M.; Relogio, A. Temporal Splicing Switches in Elements of the TNF-Pathway Identified by Computational Analysis of Transcriptome Data for Human Cell Lines. *Int. J. Mol. Sci.* **2019**, *20*, 1182. [CrossRef]

28. El-Athman, R.; Genov, N.N.; Mazuch, J.; Zhang, K.; Yu, Y.; Fuhr, L.; Abreu, M.; Li, Y.; Wallach, T.; Kramer, A.; et al. The Ink4a/Arf locus operates as a regulator of the circadian clock modulating RAS activity. *PLoS Biol.* **2017**, *15*, e2002940. [CrossRef] [PubMed]

29. Roadmap Epigenomics, C.; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; et al. Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317–330.

30. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–97. [CrossRef]

31. Uhlen, M.; Fagerberg, L.; Hallstrom, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, A.; Kampf, C.; Sjostedt, E.; Asplund, A.; et al. Tissue-based map of the human proteome. *Science* **2015**, *347*, 1260419. [CrossRef] [PubMed]

32. Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX web server: An online force field. *Nucleic Acids Res.* **2005**, *33*, W382–388. [CrossRef] [PubMed]

33. Zhu, B.; Zhang, Q.; Pan, Y.; Mace, E.M.; York, B.; Antoulas, A.C.; Dacso, C.C.; O'Malley, B.W. A Cell-Autonomous Mammalian 12 hr Clock Coordinates Metabolic and Stress Rhythms. *Cell Metab.* **2017**, *25*, 1305–1319. [CrossRef] [PubMed]

34. Ellis, R.J. Macromolecular crowding: Obvious but underappreciated. *Trends Biochem. Sci.* **2001**, *26*, 597–604. [CrossRef]

35. Rivas, G.; Minton, A.P. Macromolecular Crowding In Vitro, In Vivo, and In Between. *Trends Biochem. Sci.* **2016**, *41*, 970–981. [CrossRef] [PubMed]

36. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–208. [CrossRef] [PubMed]

37. Kulakovskiy, I.V.; Medvedeva, Y.A.; Schaefer, U.; Kasianov, A.S.; Vorontsov, I.E.; Bajic, V.B.; Makeev, V.J. HOCOMOCO: A comprehensive collection of human transcription factor binding sites models. *Nucleic Acids Res.* **2013**, *41*, D195–202. [CrossRef] [PubMed]

38. Kulakovskiy, I.V.; Vorontsov, I.E.; Yevshin, I.S.; Sharipov, R.N.; Fedorova, A.D.; Rumynskiy, E.I.; Medvedeva, Y.A.; Magana-Mora, A.; Bajic, V.B.; Papatsenko, D.A.; et al. HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **2018**, *46*, D252–D259. [CrossRef]

39. Herwig, R.; Hardt, C.; Lienhard, M.; Kamburov, A. Analyzing and interpreting genome data at the network level with ConsensusPathDB. *Nat. Protoc.* **2016**, *11*, 1889–1907. [CrossRef]

40. Chen, E.Y.; Tan, C.M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G.V.; Clark, N.R.; Ma'ayan, A. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **2013**, *14*, 128. [CrossRef]

41. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74. [CrossRef] [PubMed]

42. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef] [PubMed]