Evidence for DNA resonance signaling via longitudinal hydrogen bonds

Ivan Savelyev [1], Max Myakishev-Rempel [1,2]

1: Localized Therapeutics, San Diego, CA, USA

2: DNA Resonance Lab, San Diego, CA, USA (http://dnaresonance.org)

Email: maxrempel (at) gmail.com

Abbreviations:

HIDERs - Homologous If Decoded Elements, Repetitive.

Protocode: Proton DNA Resonance Code.

H-bonds - Hydrogen bonds

ABSTRACT

The theory of the morphogenic field suggests that chemical signaling is supplemented by electromagnetic signaling governing the structure and shape of tissues, organs and the body. The theory of DNA resonance suggests that the morphogenic field is created by the genomic DNA which sends and receives electromagnetic signals in a sequence-specific manner. Previously, the authors have proposed the existence of HIDERs, genomic elements that serve as antennas in resonance signaling and demonstrated that they occur nonrandomly and are conserved in evolution. Here, it is proposed that longitudinal hydrogen bonds exist in the double helix, that chains of these bonds form delocalized proton clouds, that the shapes of these clouds are sequence-specific and form the basis of sequence-specificity of resonance between HIDERs. Based on longitudinal hydrogen bonds, a proton DNA resonance code was devised and used to identify HIDERs which are enriched 20 fold in the genome and conserved in evolution. It was suggested that these HIDERs are the key elements responsible for DNA resonance signaling and the formation of the morphogenic field.

In addition to chemical signaling, there is electromagnetic signaling that controls the structure and shape of tissues, organs, and the body. Proposed nearly 100 years ago (Gurwitsch, 1922), the morphogenic field was experimentally demonstrated by independent groups (Gurwitsch, 1988; Volodyaev and Beloussov, 2015). In these experiments, perturbing one of the chemically separated biological samples leads to measurable effects in another (Cifra et al., 2011; Scholkmann et al., 2013; Trushin, 2004; Xu et al., 2017). Distorted reflection of the field produced by fish embryos causes developmental abnormalities, thus confirming that the field is electromagnetic and has morphogenic properties. (Burkov et al., 2008; Burlakov et al., 2012). The electromagnetic oscillations in the cells were proposed to be driven by the constant chemical energy flux and are estimated to be in the millimeter-wave region (Frohlich, 1988). Genomic DNA was proposed to be the main source and receiver of the morphogenic field, allowing the genomic program to participate in the morphogenesis directly via DNA resonance signaling (Miller and Webb, 1973).

Coherent oscillations in DNA were detected in the THz range (Sajadi et al., 2011). There are many models for mechanical oscillations in DNA (Scott, 1985; Volkov and Kosevich, 1987). In addition to mechanical oscillations in DNA, we suggested the sequence-dependence of oscillations in delocalized proton clouds in the base stack (Savelyev and Myakishev-Rempel, 2019; Savelyev et al., 2019).

Since the oscillations are sustained better when resonators are present in large numbers, we suggested that it is genomic repeats that are the key resonators in the nucleus and that they are forming a signaling network. This way, conformational changes in the chromatin in one location can lead to conformational changes in the chromatin of similar DNA sequences. We suggested that this process is deliberate and that it was developed by evolution. That would explain the high number of non-coding DNA in complex organisms and the conservation of repetitive elements in evolution (Su et al., 2014). We suggested that the resonance signaling between repeats in untranslated DNA is responsible for the structure and higher functions of complex organisms including the mental functions (Polesskaya et al., 2018; Savelyev et al., 2019).

The idea that transposons comprising over 50% of the genome serve a useful function is not new. The discoverer of transposons, Nobel winner Barbara McClintock ascribed to them the function of universal regulatory units and called them "control elements"(Mcclintock, 1956).

The realization that genomic repeats are key sequences forming a DNA resonance network allowed us to make one step further and hypothesize that there possibly exist sequences which resonate with the repeats but differ from them in primary sequence. We suggested that this could be possible if proton clouds in these sequences are of the same shape and support similar oscillations. We called such sequences HIDERs (Homologous If Decoded Elements, Repetitive) defined as sequences that have different primary structures but support similar oscillations. This came out to be a productive approach. Based on DNA sequence we were able to predict the similarities in delocalized proton clouds structures, formalize the recoding scheme, predict HIDER sequences in the genome and demonstrate that they are enriched in the genome with high statistical significance (Savelyev and Myakishev-Rempel, 2019). One of the advantages of such a computational genomics approach is that it is agnostic to the exact physical mechanism of the resonance, allowing verification of its existence prior to the experimental confirmation. Once the existence of HIDERs is experimentally confirmed, their structure may provide an insight into the modes of their resonance. Here we substantially improve the resolution of the molecular modeling, improve the recoding algorithm and expand computational analysis of HIDERs.

Methods

500 Kb fragments were selected at random from the human genome, version GRCh38. The repeats were masked using RepeatMasker (http://repeatmasker.org/) followed by removal of identical repeats longer than 14 bp. For selected sequences, B-form DNA models in PDB format were produced using 3D-DART (van Dijk and Bonvin, 2009), visualized and annotated in PyMol (DeLano and Others, 2002). Longitudinal H-bonds were selected automatically between the amine nitrogen and the carbonyl oxygen located at most 3.7Å from each other in consecutive base pairs. Recoding was done as described in Fig. [Recoding]. HIDERs were detected in the recoded sequence as identical strings longer than 14 bp and repeated at least 4 times per 500Kb. Randomized controls were made using a custom algorithm R20N: Randomization was done in 20 b.p. bins while the positions of Ns used for masking repeats were retained unchanged. This was done to preserve average nucleotide density in each bin. The significance of enrichment was determined using a paired t-test. Tandem HIDERs were identified in recoded sequences with the requirement of 100% identity of the repeating units.

For conservation measurement, a chromosome was masked and divided into 25Kb bins. In each bin, for each HIDER, a 200bp flanking region was selected centered at the middle of the HIDER, a conservation score was averaged in all HIDER-flanking regions per bin producing average HIDER conservation score per bin. Similarly, an average background conservation score was produced per bin using the remaining sequence of the bin after excluding the HIDER-flanking regions. The masked sequence was excluded from the averaging.

Results

Longitudinal hydrogen bonds

Previously we proposed the existence of clouds of delocalized protons stretching along the base stack (Savelyev et al., 2019) and, based on hydrogen bond patterns, computationally demonstrated significant enrichment of HIDERs (Savelyev and Myakishev-Rempel, 2019). This was done based on the canonical transverse hydrogen bonds, the ones that bind the pairs of bases together. To improve this molecular model, we explored possible delocalized proton patterns in 3D models of DNA in search of longitudinal clouds of delocalized protons. We noticed that in addition to the transverse hydrogen bonds, there likely exist longitudinal hydrogen bonds connecting some of the consecutive bases together. The conditions for the formation of longitudinal H-bonds are quite favorable: in some of the consecutive base pairs, an amino group of one base pair and the carbonyl oxygen of the next base pair come sufficiently close, Fig. [Bonds].
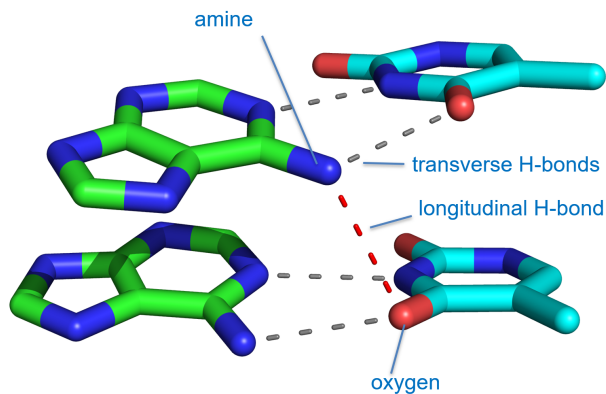

Fig. [Bonds] Transverse (canonical) and longitudinal hydrogen bonds in two sequential AT pairs.

Thus the search for the DNA resonance signaling brought us to the search of delocalized proton clouds and this brought us to the prediction of longitudinal H-bonds in DNA. Our search of the literature didn't reveal any previous mention of such H-bonds likely because they were considered inconsequential, but longitudinal H-bonds for RNA have been described in x-ray structures (Šponer et al., 2003) (the authors called them "neighbor contacts"). Since base stacking structures in RNA and DNA are identical (while the difference is in the sugars of the backbone), we believe that the existence of longitudinal H-bonds is very likely. Here, we propose that they are central to DNA resonance signaling.

According to our hypothesis, chains of longitudinal and transverse H-bonds collectively form clouds of delocalized protons spanning multiple base pairs. At this initial stage, we aimed to record the structure of the longitudinal H-bonds into a simplified code with the assumption that only the general shape of the proton cloud defines its resonance properties. For that purpose, we modeled longitudinal H-bonds for all 16 possible dinucleotide pairs, Supplement Fig. [Dinucleotides] using 3.7Å cutoff for the distance between the amine nitrogen and the carbonyl oxygen of consecutive base pairs. We classified the resulting longitudinal H-bond arrangements into four types: f - no bonds, k - 1 bond, r - 2 bonds in the major groove, m - 1 bond in the major groove and 1 bond in the minor groove, Fig. [Types]. We made a distinction between r and m type because we predict that they would display very different properties: the H-bonds in r-type will be tightly linked by chemical resonance as they are very close to each other while m-type will be very loosely linked as they are far from each other.
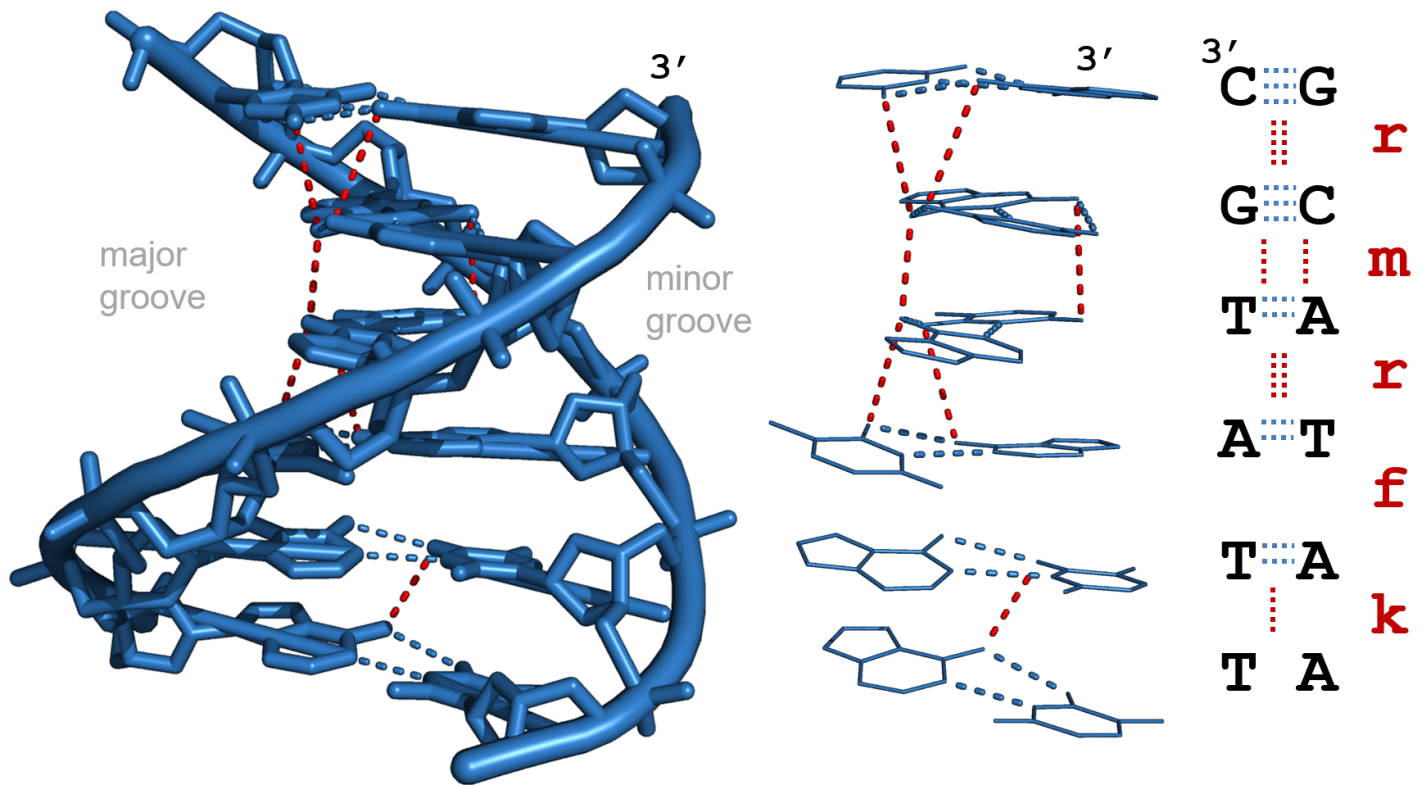
Fig. [Types] *Types of longitudinal H-bond arrangements in dinucleotide pairs: f - no bonds, k - 1 bond, r - 2 bonds in the major groove, m - 1 bond in the major groove and 1 bond in the minor groove.*

The resulting classification of dinucleotides by longitudinal H-bond arrangements is shown in Fig. [Recoding A and B]. This classification made it possible to recode the primary sequence of DNA into longitudinal H-bond types as shown in Fig. [Recoding C]. Note that the types correspond not to individual nucleotides but to connections between two consecutive nucleotides. Since this recoding scheme describes the structure of predicted delocalized proton clouds in the base stack, we will refer to it as a proton code or protocode for brevity.

A

f – no bonds
k – 1 bond
r – 2 bonds in the major groove
m – 1 bond in the major groove and 1 bond in the minor groove

f	r	k	m
at	cg	aa	ac
ta	ca	ag
		ga	cc
		tc	ct
		tg	gc
		tt	gg
			gt

B

C

G C G C T C A T C A
m r m m k k f k k

Fig. [Recoding] *Recoding scheme. (A) Definitions of longitudinal H-bond bond types. (B) Conversion table of dinucleotides to longitudinal H-bond types. (C) Algorithm of recoding a nucleotide sequence into protocode.*

Note that although protocode recodes 4 letters of the primary sequence to 4 letters of the protocode sequence, some of the information is discarded, for example, 7 different dinucleotides are collapsed into m-type. This allows for a certain ambiguity in the primary code for the creation of identical protocode sequences. As outlined above, we hypothesized that sequences having a similar shape of delocalized proton clouds would resonate and be used by nature for resonance signaling even if their primary sequences differ. We also hypothesized that such sequences (HIDERs) are enriched in the genome.

Here, we test this using the protocode. Genomic repeats were masked in a 500Kb genomic fragment, the sequence was converted to protocode (using the scheme from Fig. [Reciding] B), and protocode-specific repeats (HIDERs) were identified. An example of protocode HIDER sequences is shown in Fig. [Alignment]. The same set of five sequences is shown at the top and bottom. At the top, they are colored by nucleotides and it is apparent that the sequences are different. At the bottom, the same sequences are colored by protocode and it is apparent that their protocode sequences are identical.
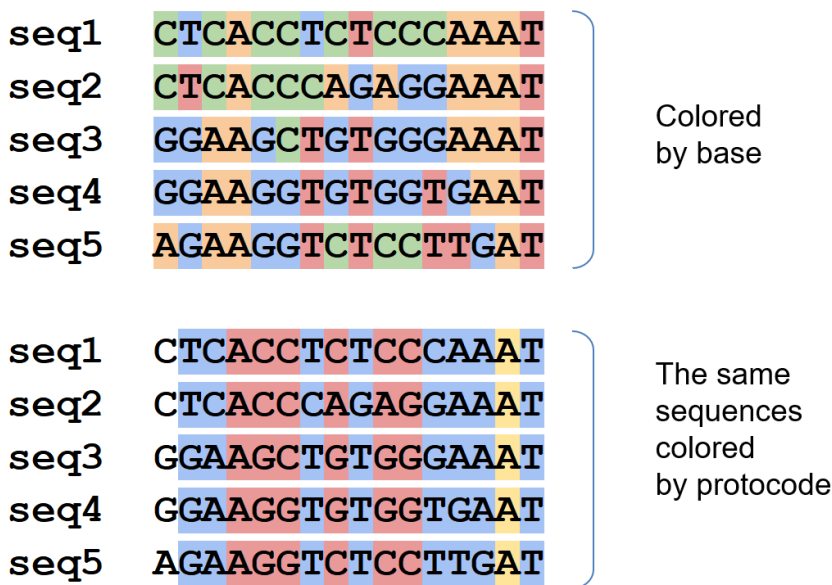

Fig. [Alignment] Example of protocode HIDER sequences.

HIDERs are enriched in genomes.

We selected several mammal species for analysis, an insect (drosophila), and a plant (arabidopsis). In each genome, four 500 kb fragments were selected at random and the repeats were masked. Randomized reference sequences (RAND) were created from the original sequences (ORIG) using our R20N method (see Methods). ORIG and RAND sequences were recoded into protocode, as presented in Fig. [Codes]. In each fragment, HIDERs longer than 14 bases and having a minimum of 4 copies per 500kb fragment were identified. In each fragment, HIDER density was calculated as the length of the sequence covered by HIDERs divided by the total length of the unmasked sequence. Each ORIG and RAND hider densities were normalized by dividing them by the corresponding RAND hider density and plotted in Fig. [Enrichment].
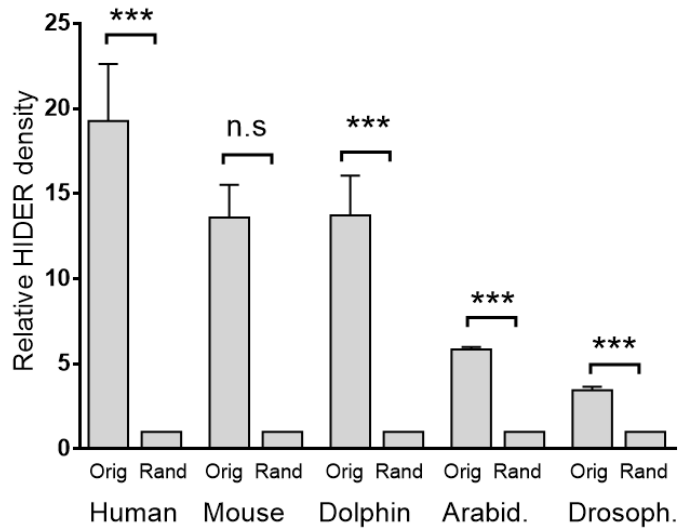
Fig. [Enrichment] HIDER density is enriched in the original over the randomized sequence.

Among the five tested species, the highest HIDER density enrichment was found in mammals and the lowest in drosophila.

We also examined a subset of HIDER sequences which are periodic in protocode. The motivation for that was that periodic molecular structures would support oscillations of proton clouds better than aperiodic ones. We took 4 fragments 500Kb each, masked all repeats (including tandems) in the primary sequence, recoded it to protocode, and identified tandems in the protocode sequence. Included were perfect tandem repeats with an at least 3bp-long repeating unit, the unit was repeated at least 4 times, the whole tandem element was to be found in the 500Kb fragment at least 2 times. An example of protocode tandem HIDERs is shown in Fig. [Period]. Note that the same sequences are aperiodic and different when colored by the base and identical and periodic when colored by protocode.
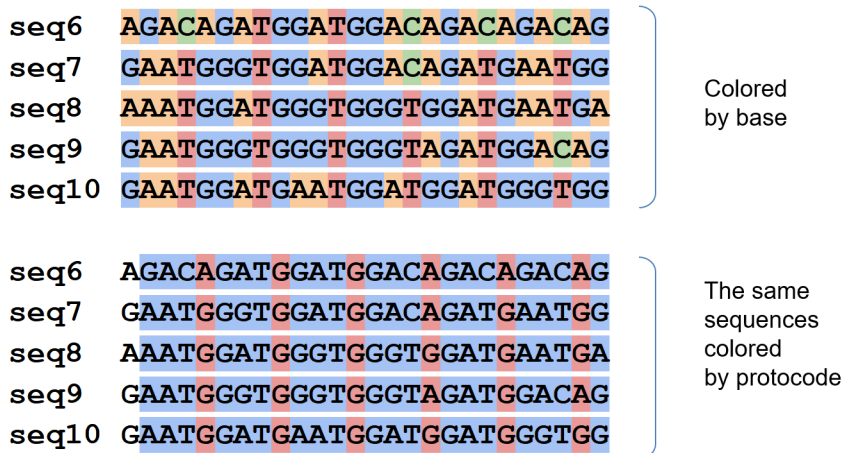

Fig. [Period] Example of protocode tandem HIDER sequences.

The counts of such protocode Tandem HIDERs were compared in the original sequence to the counts in its randomized version, see Fig. [Tandems]. The subset of Protocode Tandem HIDERs was found to be enriched in the original over the randomized sequence, yet the enrichment was not as high as in Fig. [Enrichment] .
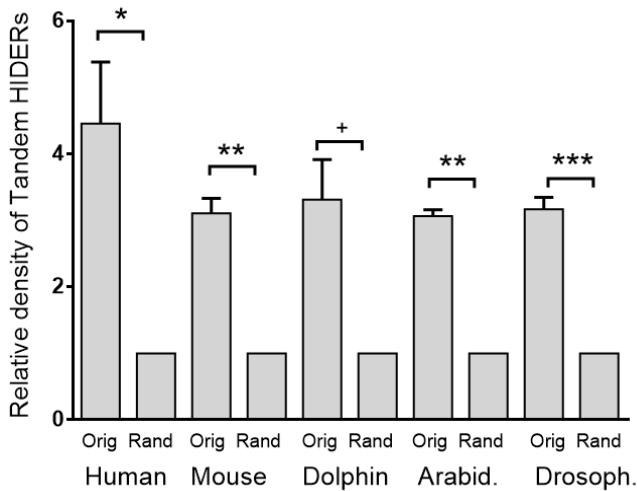
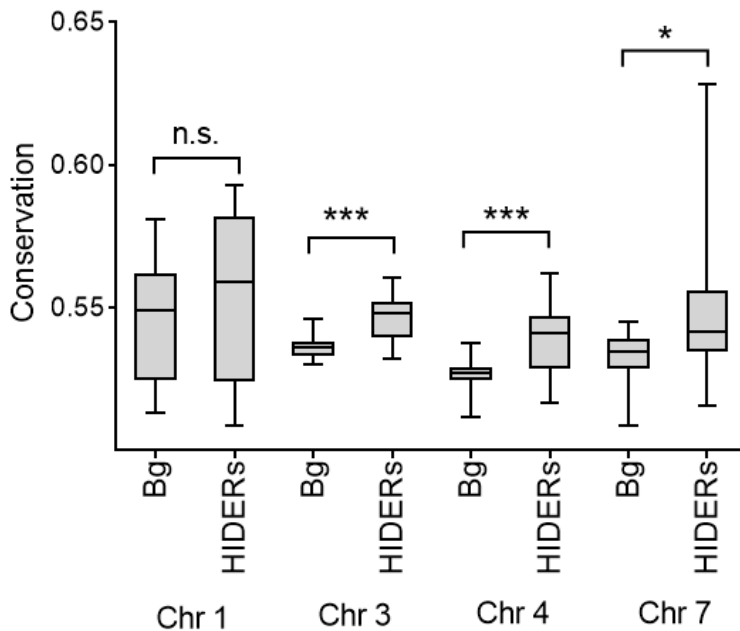Fig. [Tandems]. The density of Tandem HIDERs is enriched in the original over the randomized sequence.

Fig. [Conservation]. Conservation of HIDERs compared to background (Bg).

Since HIDERs were found to be enriched in the genomic sequence over its randomized version (Fig. [Enrichment]), it is likely that they serve a useful function and therefore we tested whether they are conserved in evolution. We took 4 large chromosomes at random, masked repeats, and identified the positions of protocode HIDERs as above. Then we divided each chromosome into 25Kb bins and compared the conservation score in 200bp areas harboring HIDERs to the average conservation score in each bin harboring these HIDERs (background), Fig. [Conservation]. Protocode HIDERs were found conserved in 3 chromosomes out of 4 chromosomes tested.

Discussion

The overall motivation for this work is to get a glimpse into the mechanisms of DNA resonance signaling which to this day remains largely hypothetical. Following the lead of Miller and Webb (Miller and Webb, 1973), we

were attracted by the idea that DNA creates a holographic field that drives and organizes the shape and function of the cell and the organism. Our previous research into the proton clouds in the double helix (Savelyev and Myakishev-Rempel, 2019; Savelyev et al., 2019) motivated us to look closely at the hydrogen bonds. We were particularly interested in finding clouds of delocalized protons that span multiple nucleotides and form patterned resonators that would resonate in a sequence-dependent manner. This allowed us to notice that in addition to canonical transverse H-bonds, there must exist longitudinal H-bonds, connecting consecutive basepairs together. Specifically, we suggested that the longitudinal hydrogen bond is formed every time amino nitrogen of one basepair is located within 3.7Å from carbonyl oxygen of a neighbor base pair, Fig. [Bonds]. As mentioned above, longitudinal H-bonds have been described in RNA structures (Šponer et al., 2003) (the authors called them "neighbor contacts"). Thus we suggest that collectively longitudinal and transverse H-bonds form delocalized proton clouds in a sequence-dependent manner. Such continuous delocalized proton clouds are known in proteins and are called proton wires (Leiderman et al., 2006). Protons are also known to be delocalized in the canonical transverse H-bonds (Sobolewski et al., 2005). Here, we proposed that a combination of transverse and longitudinal hydrogen bonds forms proton wires spanning several nucleotides. Next, we suggest that proton wires or combinations of overlapping proton wires serve as oscillators. Next, we suggested that the patterns of these oscillators are sequence-specific and these oscillators that have similar patterns resonate with each other in a sequence-specific manner.

Although the proton cloud patterns of the identical sequences would be identical, these resonances can be tested only experimentally. Yet, if the sequences are different, they still could have similar proton cloud patterns. Each dinucleotide in the double helix has a specific geometry of longitudinal H-bonds. Here, we ignored minor differences among longitudinal H-bond configurations and summarized 16 possible dinucleotides according to whether longitudinal H-bonds are located in the major or the minor groove, Fig. [Recoding]. That allowed us to simplify the longitudinal H-bond-based code to 4 letters (termed protocode) thus making it possible to find repeats in this code. Next, we searched for such sequences that have similar protocode sequences while differing in the primary (AGCT) sequence - these repeats found only in protocode were called HIDERs (Homologous If Decoded Elements Repetitive). The consequent counting of HIDERs in various species showed significant up to 20 fold enrichment of these hiders in natural sequences over their randomized versions, Fig. [Enrichment], suggesting that HIDERs are indeed functional and selected by evolution. We suggest that HIDERs are harboring patterned delocalized proton clouds and serve for resonance signaling via electromagnetic or possibly electroacoustic waves and currents. Consequent analysis of evolutionary conservation of HIDERs showed that HIDERs are located preferentially in the conserved segments of the genome Fig. [Conservation].

Currently, we don't see any alternative explanation for the observed high enrichment of HIDERs in the genome and their conservation in evolution rather than their proposed function for resonance signaling. It is unlikely that they are bound by the same proteins since it is only proton cloud patterns that they have in common, while their nucleotide sequences are very different. It is unlikely that proteins would recognize proton cloud patterns while ignoring nucleotide sequence differences. Therefore, the obtained enrichment and conservation results offer a good confirmation of the DNA resonance signaling function of protocode HIDERs.

So far, there is no explanation for the uneven distribution of dinucleotides in various branches of the evolutionary tree. This unevenness is pretty well conserved within each branch and various branches differ in the dinucleotide content. Thus, various forms of life would differ in their DNA resonance properties that define their structure and function. Melting temperature is unlikely responsible for the unevenness since CG (30% of expected) and GC (90% of expected in vertebrates (Costantini et al., 2009)) radically differ in frequencies. Amino acid codon bias is unlikely an explanation for the overall unevenness of dinucleotide frequencies because only a small fraction of the genome is made of coding sequences or pseudogenes. Since each dinucleotide has a specific longitudinal H-bond configuration, resonance of delocalized protons in H-bonds may be responsible for the unevenness of dinucleotide frequencies in various branches of the evolutionary

tree. Thus, the least frequent dinucleotides in vertebrates are CG (30% of expected) and TA (70% of expected) which are the only two dinucleotides that have an unusual pattern of longitudinal H-bonds: the two H-bonds are parallel in the major groove. Due to the parallel orientation for the longitudinal H-bonds, this pattern may have a unique oscillation frequency that may be selected against in vertebrates. The third least frequent dinucleotide in vertebrates is AT (82% of expected), which is also unique: it is the only dinucleotide which has no longitudinal H-bonds. Conversely, CA, AA and GA, the only three dinucleotides that have a single longitudinal H-bond have high frequencies (above 105% of expected). Therefore the unevenness of dinucleotide distribution may be mechanistically linked with the organism's structure and function via the vibrational properties of longitudinal H-bonds in vast noncoding parts of the genome.

TATA box (`TATAWAW`) is one of the key transcription factor binding sites located around the 30th base upstream of the transcription start site in vertebrates. (Xu, Gonzalez-Hurtado, and Martinez 2016). We found that the protocode sequence of the TATA box (`TATA=rfr`) is remarkable: it combines two copies of the most stable diamond-shaped resonant H-bond configurations of TA (`TA=r`) with the break in the H-bond cloud of AT, see TAT sequence in Fig. [Types]. It is also remarkable that the protocode sequence of the ATG start codon (`ATG=fk`) also was found to contain a break in the proton cloud (`AT=f`). Also, two of the three stop codons were found to contain the diamond-shaped H-bond configuration of TA (`TA=r`). These observations suggest that diamond-shaped H-bond configurations of TA and the break in the H-bond cloud of AT play a role in transcription initiation, and translation initiation and termination.

Further computational studies of longitudinal H-bonds and protocode HIDERs should explore the structure of proton clouds in better detail, and may utilize computer modeling of oscillation in these clouds since our analysis only used general locations of the longitudinal H-bonds - whether they are located in the major or the minor groove. More precise analysis should produce better signals. Further, it would be prudent to check whether HIDERs are located near genomic annotations related to chromatin function and gene expression.

Experimental studies may explore HIDERs using spectroscopy methods to verify that resonance takes place and various sequences resonate with each other as predicted by protocode.

In practical terms, this work is a step toward the verification of the existence of resonance signaling in the genome. Once resonance DNA signaling is proven, this would allow a better understanding of such fundamental questions as organism development and coordination of its functions. Once the role of DNA resonance in these functions is clarified, it would become possible to start applying DNA-resonance-based technologies in medicine, biotechnology and agriculture.

Acknowledgments.

REFERENCES

Burkov, V.D., Burlakov, A.B., Perminov, S.V., Kapranov, Y.S., Kufal, G.E., 2008. [Correction of Long Range Interaction Between Biological Objects Using Corner-Cube Reflectors]. Biomedical Radioelectronics 41–48.

Burlakov, A.B., Kapranov, Y.S., Kufal, G.E., Perminov, S.V., 2012. [About possible influence on biological object electromagnetic fields]. Weak and ultraweak fields and radiation in biology and medicine-Proceedings of IV International Congress 111–112.

Cifra, M., Fields, J.Z., Farhadi, A., 2011. Electromagnetic cellular interactions. Prog. Biophys. Mol. Biol. 105, 223–246.

Costantini, M., Cammarano, R., Bernardi, G., 2009. The evolution of isochore patterns in vertebrate genomes. BMC Genomics 10, 146.

DeLano, W.L., Others, 2002. Pymol: An open-source molecular graphics tool. CCP4 Newsletter on protein crystallography 40, 82–92.

Frohlich, H., 1988. Theoretical physics and biology. Biological Coherence and Response to External Stimuli. Berlin: Springer-Verlag 1–24.

Gurwitsch, A., 1922. Über den Begriff des Embryonalen feldes. Wilhelm Roux Arch. Entwickl. Mech. Org. 51, 383–415.

Gurwitsch, A.A., 1988. A historical review of the problem of mitogenetic radiation. Experientia 44, 545–550.

Leiderman, P., Huppert, D., Agmon, N., 2006. Transition in the temperature-dependence of GFP fluorescence: from proton wires to proton exit. Biophys. J. 90, 1009–1018.

Mcclintock, B., 1956. Controlling elements and the gene. Cold Spring Harb. Symp. Quant. Biol. 21, 197–216.

Miller, R.A., Webb, B., 1973. Embryonic Holography: An Application of the Holographic Concept of Reality. DNA Decipher Journal 2.

Polesskaya, O., Guschin, V., Kondratev, N., Garanina, I., Nazarenko, O., Zyryanova, N., Tovmash, A., Mara, A., Shapiro, T., Erdyneeva, E., Others, 2018. On possible role of DNA electrodynamics in chromatin regulation. Prog. Biophys. Mol. Biol. 30, 1e5.

Sajadi, M., Furse, K.E., Zhang, X.-X., Dehmel, L., Kovalenko, S.A., Corcelli, S.A., Ernsting, N.P., 2011. Detection of DNA--Ligand Binding Oscillations by Stokes-Shift Measurements. Angew. Chem. Int. Ed. 50, 9501–9505.

Savelyev, I., Myakishev-Rempel, M., 2019. Possible traces of resonance signaling in the genome.

Savelyev, I.V., Zyryanova, N.V., Polesskaya, O.O., Myakishev-Rempel, M., 2019. On The Existence of The DNA Resonance Code and Its Possible Mechanistic Connection to The Neural Code. Neuroquantology 17. https://doi.org/10.14704/nq.2019.17.2.1973

Scholkmann, F., Fels, D., Cifra, M., 2013. Non-chemical and non-contact cell-to-cell communication: a short review. Am. J. Transl. Res. 5, 586–593.

Scott, A.C., 1985. Soliton oscillations in DNA. Phys. Rev. A Gen. Phys. 31, 3518–3519.

Sobolewski, A.L., Domcke, W., Hättig, C., 2005. Tautomeric selectivity of the excited-state lifetime of guanine/cytosine base pairs: the role of electron-driven proton-transfer processes. Proc. Natl. Acad. Sci. U. S. A. 102, 17903–17906.

Šponer, J., Mokdad, A., Šponer, J.E., Špačková, N. 'a, Leszczynski, J., Leontis, N.B., 2003. Unique Tertiary and Neighbor Interactions Determine Conservation Patterns of Cis Watson–Crick A/G Base-pairs. J. Mol. Biol. 330, 967–978.

Su, M., Han, D., Boyd-Kirkup, J., Yu, X., Han, J.-D.J., 2014. Evolution of Alu elements toward enhancers. Cell Rep. 7, 376–385.

Trushin, M.V., 2004. Distant non-chemical communication in various biological systems. Riv. Biol. 97, 409–442.

van Dijk, M., Bonvin, A.M.J.J., 2009. 3D-DART: a DNA structure modelling server. Nucleic Acids Res. 37, W235–9.

Volkov, S.N., Kosevich, A.M., 1987. Conformation oscillations of DNA. Mol. Biol. 21, 797–806.

Volodyaev, I., Beloussov, L.V., 2015. Revisiting the mitogenetic effect of ultra-weak photon emission. Front. Physiol. 6, 241.

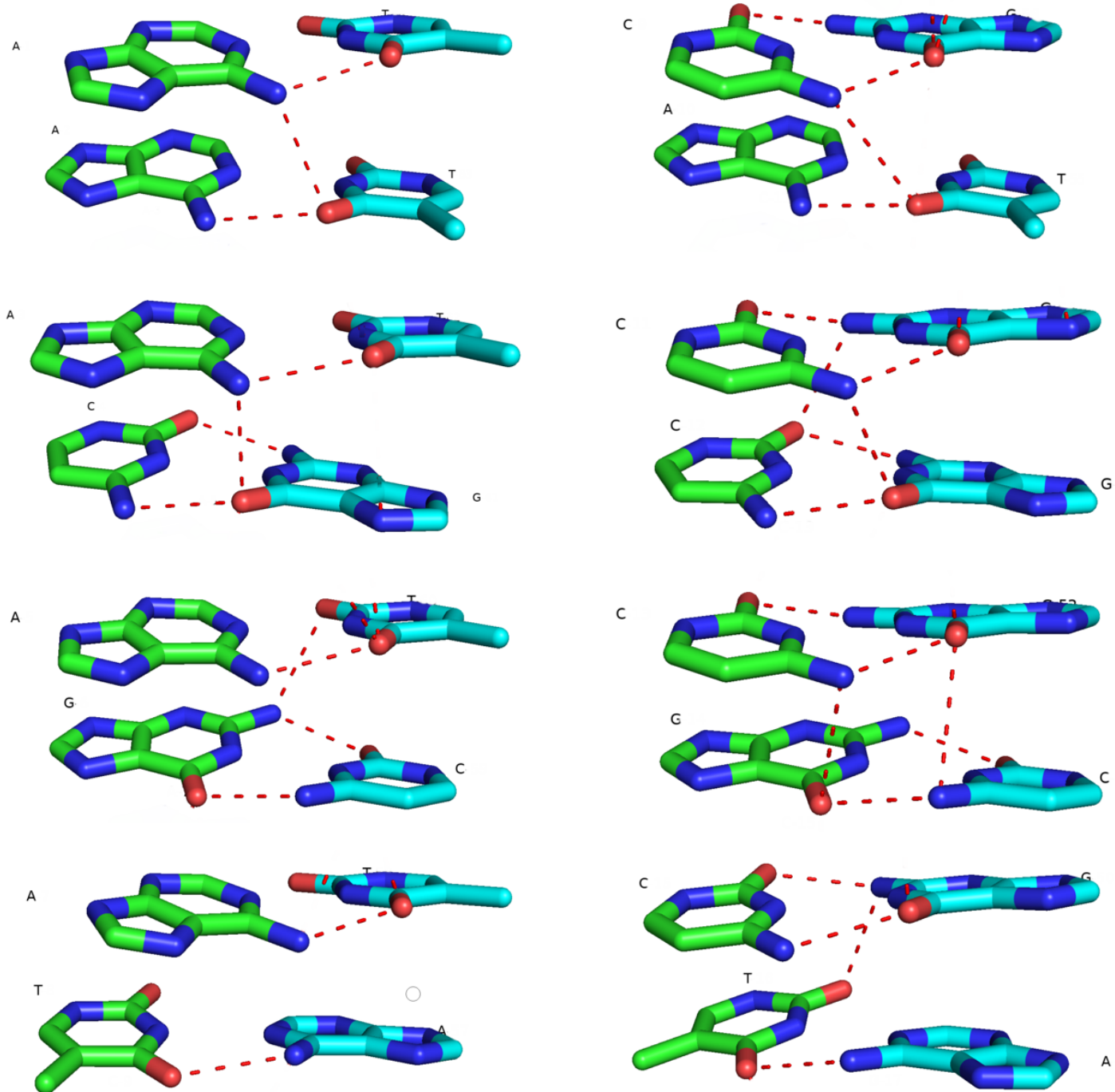Xu, J., Yang, F., Han, D., Xu, S., 2017. Wireless Communication in Biosystems. arXiv [physics.bio-ph].
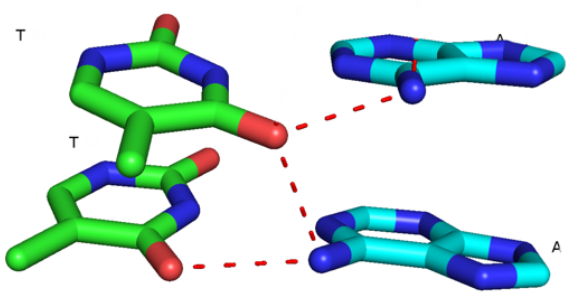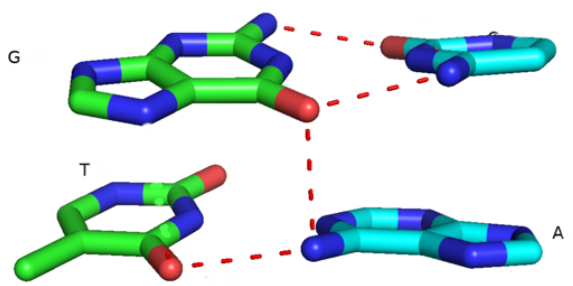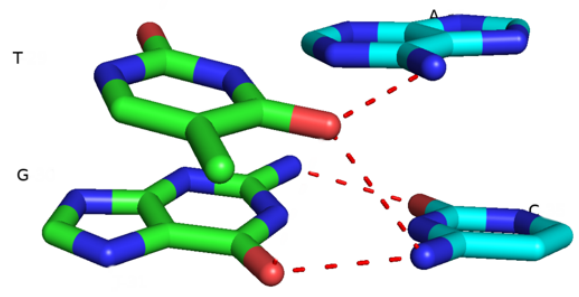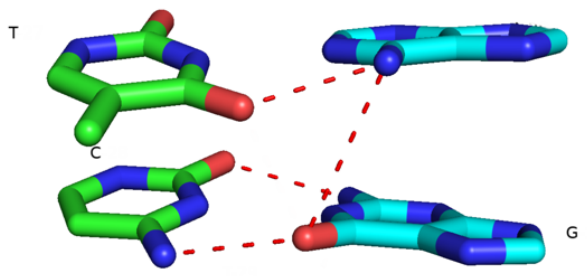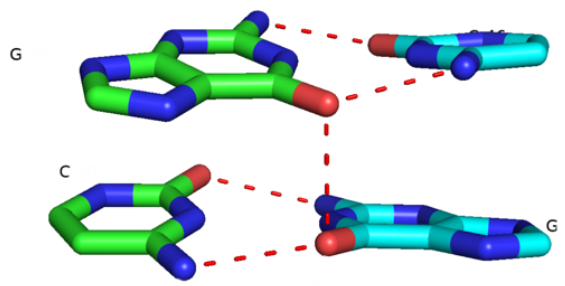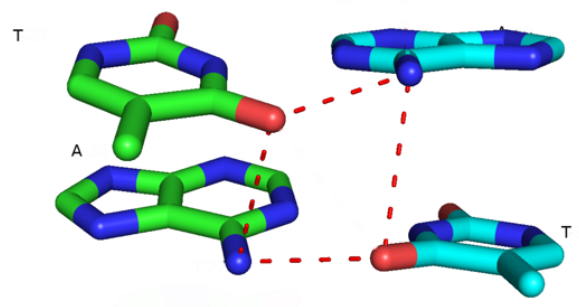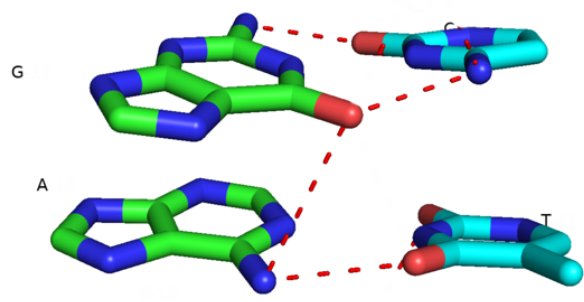
SUPPLEMENT

Fig. [Dinucleotides] Part 1

Fig. [Dinucleotides] Part 2