

DEPPDB v.3 and its use to study electrostatic and other physical properties of genome DNA and its elements, relevant to their biological functions

Alexander A. Osypov

Institute of Higher Nervous Activity and Neurophysiology of RAS, Moscow, Russia, aosypov@gmail.com

DNA is a highly charged molecule and its electrostatic and other physical properties define its shape in the functional space and greatly influence its interactions with different proteins, especially regulating transcription, in particular RNA-polymerases and transcription factors [1-5]. DEPPDB was developed to hold and provide all available information on these properties of genome DNA combined with its sequence and annotation of biological and structural properties of genome elements and whole genomes, organized on a taxonomical basis.

The electrostatic potential around the double-helical DNA molecule was calculated by the original method [6] using a specially developed program package [7,8]. Calculations of other physical properties are based on the di- and trinucleotide content. Different cross-correlation analysis algorithms are applied.

Several calculation algorithms, data processing approaches and database architecture solutions were fundamentally revisited and fine tuned since the previous database releases to withstand the considerably increased data abundance.

DEPPDB contains all completely sequenced archaeal, bacterial, viral, mitochondrial, plastids and eukaryotic genomes according to current release of NCBI RefSeq [9]. Data for promoters, regulation sites, binding proteins *etc.* are incorporated from different databases and compiled from original literature annotation. All data are fully integrated, several tools are provided to support different forms of analysis. Calculation on the fly of the user-provided sequences and selected GeneBank records is available.

DEPPDB can be considered as a portal or collection of databases on the electrostatic and other physical properties of whole genomes and different genome elements in different taxa

and organisms: Promoter DB, Regulatory Sites (Transcription Factors, TF) DB, Gene Starts DB, Terminator DB, *etc.* as well as comprehensive analysis toolbox.

The work for assisting epigenetics research with computations of *e.g.* supercoiling, nucleosome positioning, methylation, *etc.* impact on the DNA physical properties, relevant to its biological functions, is in progress.

The database is available for academic use via the web interface at <http://deppdb.psn.ru> (<http://electrodna.psn.ru> can be used as an alternative address).

The authors are grateful to Institute of Mathematical Problems of Biology of RAS for hosting the DEPPDB database.

The work was supported by RSF grant 14-25 00072.

1. Kamzolova, A.A. Sorokin, T.D. Dzhelyadin P.M., Beskaravainy, A.A. Osypov (2005) Electrostatic potentials of *E. coli* genome DNA, *J Biomol Struct Dyn*, **23(3)**: 341-346.
2. E. A. Krutinina, G. G. Krutinin, S. G. Kamzolova, A. A. Osypov (2011) New Insights into Protein-DNA Electrostatic interactions: Beyond Promoters to Transcription Factors Binding Sites, *J Biomol Struct Dyn*, **28(6)**: 1137-1138.
3. A. A. Osypov, G. Krutinin, E. Krutinina, S. Kamzolova (2014) Electrostatic properties of natural DNA palindromes – transcription factors binding sites and terminators, *FEBS Journal*, **281(s1)**: 681.
4. A. A. Osypov (2014) DNA electrostatics: its role in transcription regulation in prokaryotes, a new natural selection factor, and a base for DNA phenotype and biophysical bioinformatics, *FEBS Journal*, **281(s1)**: 680.
5. E. Krutinina, G. Krutinin, S. Kamzolova, A. A. Osypov (2014) Electrostatics of prokaryotic transcription factors match that of their binding sites, *FEBS Journal*, **281(s1)**: 682.

6. R. V. Polozov et al. (1999) Electrostatic potentials of DNA. Comparative analysis of promoter and nonpromoter nucleotide sequences, *J. Biomol. Struct. Dyn.*, 16(6), 1135-43.
7. A. A. Osypov, G.G. Krutinin, S. G. Kamzolova (2010) DEPPDB – DNA Electrostatic Potential Properties Database. Electrostatic Properties of Genome DNA, *J Bioinform Comput Biol*, **8(3)**: 413-425.
8. A. A. Osypov, G.G. Krutinin, E.A. Krutinina, S. G. Kamzolova (2012) DEPPDB – DNA Electrostatic Potential Properties Database. Electrostatic Properties of Genome DNA elements, *J Bioinform Comput Biol*, **10(2)**: 1241004.
9. The NCBI handbook [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2002 Oct. Chapter 18, The Reference Sequence (RefSeq) Project. Available from <http://www.ncbi.nlm.nih.gov/books/NBK21091>