BIOINFORMATICS OF GENOME REGULATION
AND STRUCTURE/SYSTEMS BIOLOGY
(BGRS/SB-2020)

The Twelfth International Multiconference

Abstracts

06–10 July, 2020
Novosibirsk, Russia
Bioinformatics of Genome Regulation and Structure/Systems Biology (BGRS/SB-2020) :
DOI 10.18699/BGRS/SB-2020-000

Organizers
Institute of Cytology and Genetics, SB RAS (ICG SB RAS)  
Novosibirsk State University (NSU)  
Ministry of Science and Higher Education of the Russian Federation (Minobrnauki of Russia)  
The Vavilov Society of Geneticists and Breeders
State Scientific-Research Institute of Physiology and Basic Medicine (PhBMRI)  
Vavilov Journal of Genetics and Breeding

Contacts
All questions, please contact the official address of the organizing committee of the conference: bgrs2020@bionet.nsc.ru

Organizing Committee Chairman: Zubova Svetlana  
Address: Lavrentieva Avenue, 10, Novosibirsk, 630090  
Phone: +7 (383) 363 4977, e-mail: svetazu@yandex.ru
Evolution of proteins involved in response to ROS

Vassily Lyubetsky  
IITP RAS  
Moscow, Russia  
lyubetsky@iitp.ru

Gregory Shilovsky  
IITP RAS; MSU  
Moscow, Russia  
gregory_sh@list.ru

Oleg Zverkov  
IITP RAS  
Moscow, Russia  
zverkov@iitp.ru

Lev Rubanov  
IITP RAS  
Moscow, Russia  
rubanov@iitp.ru

Alexandr Seliverstov  
IITP RAS  
Moscow, Russia  
slvstv@iitp.ru

Abstract — Original software was used to specify the evolution of transcription factors Nrf2 and Bach1 in Deuterostomia. The transcription factors are antagonistically involved in the response to reactive oxygen species (ROS). The original algorithm shows that Bach emerged by duplication of Nfe2, an ortholog of Nfe2l2, in the chordate ancestor. At the N-terminus, the copy was provided with the BTB domain from a gene orthologous to the European lancelet Branchiostoma lanceolatum gene BL0308_cuf1 followed by a domain typical for the Zinc finger C2H2 superfamily.

Keywords — Nrf2, anti-ageing program, anti-ROS; Bach1, ageing program, pro-ROS; ROS

Introduction

In many vertebrates, Nrf2 (Nfe2l2), Bach proteins (Bach1 and Bach2), as well as Keap1, β-TrCP, c-Myc, and GSK3b are components of the regulatory network (e.g., in Mus musculus) that is, among other functions, associated with regulation of the level of reactive oxygen species (ROS). In particular, the relationship between Nrf2 and Bach1 is critical for the regulation of heme oxygenase 1 expression, which provides for heme degradation. The functioning of this network is related to species-specific lifespan as well as to many human diseases. Nrf2 activates about 200 genes; Nrf2 and Bach1 are antagonists in controlling ROS levels, which is also due to the competition for ARE DNA-binding sites. These proteins usually function by complex formation with the BTB domain (e.g., in Mus musculus) that is, among other functions, associated with regulation of the level of reactive oxygen species (ROS). In particular, the relationship between Nrf2 and Bach1 is critical for the regulation of heme oxygenase 1 expression, which provides for heme degradation. The functioning of this network is related to species-specific lifespan as well as to many human diseases.

Analysis of the search results. Among invertebrate deuterostomes, Bach proteins have been found only in Ciona intestinalis and C. savignyi (tunicates diverged early from other chordates). At least one Bach protein has been found in all invertebrates. It looks like Bach emerged in the common ancestor of vertebrates and tunicates (and was possibly lost in appendicularians), after which it diverged into Bach1 and Bach2 in the common ancestor of cartilaginous and bony fishes. In fishes, the Bach genes were many times duplicated and lost; their genomes commonly have many paralogs that are orthologous to Bach. E.g., the huchen Hucho hucho has four genes orthologous to Bach2. The Bach1 gene is represented by a single copy in all tetrapods. The alignment of Bach proteins indicates considerable conservation of each of them in most mammals. Within Euarchontoglires, Bach1 demonstrates only minor changes in most rodents including the Damaraland mole-rat (DMR) Fukomys damarensis. The only exception is the C-terminal region of all Bach1 isoforms in the naked mole-rat (NMR) Heterocephalus glaber. One Bach1 isoform has an extended deletion, although a short region upstream of the C-terminus is conserved. Another isoform has a long insertion in the same region. The full-length Bach1 protein of the NMR usually shares the dispensable amino acids with the DMR but not the mouse, which agrees with their taxonomical position. No significant differences in Bach1 have been revealed in primates. Beyond Euarchontoglires, the Bach1 sequences of the sloth Choloepus hoffmanni, tenrec Echinops telfairi, shrew Sorex araneus, dolphin Tursiops truncatus, flying fox Pteropus vampyrus are similar to those in primates and most rodents. Conversely, other
representatives of Laurasiatheria as well as the nine-banded armadillo Dasypus novemcinctus have insertions or deletions in Bach1 in the corresponding region of the NMR. The Bach1 of the hedgehog Erinaceus europaeus, elephant Loxodonta africana, and hyrax Procavia capensis in Bach1 has a very long C-terminal deletion covering both conserved and variable regions. On the other hand, one of the Bach2 isoforms in primates Macaca mulatta, M. nemestrina, and Pan troglodytes has an N-terminal extension not observed in humans.

2) Here we consider all species represented in RefSeq and Ensembl supplemented by those in GenBank. The Bach1 protein of cartilaginous fishes lacks the Tyr-486. The Bach1 in tetrapods has conserved the functionally significant phosphorylated tyrosine and the neighboring amino acids remained largely unchanged, although D→S substitutions and proline loss in the CP motifs are observed in long-lived rodent DMR and NMR. In the Australian ghostshark Callorhinchus mili, cloudy catshark Scyliorhinus torazame, whale shark Rhincodon typus, and brownbanded bamboo shark Chiloscyllium punctatum, this tyrosine is replaced with phenylalanine. The Bach in Chona spp. has preserved this tyrosine unlike some of the neighboring amino acids. The predicted heme-binding sites in Bach1 of tetrapods insignificantly differ from those in mouse except the species described below. These sites include the 223-LCPKYR-228 (C→G in marsupials, while in the platypus Ornithorhynchus anatinus it is the same as in the mouse; or C→Y in the common wall lizard Podarcis muralis); 300-QCPAEQ-305, which considerably changed or disappeared in most mammals; 435-ECPWLQ-340 (conserved in all tetrapods); 463-NCPFIS-468 (the cysteine is conserved in tetrapods, and 1→M in placentalts including DMR, NMR and the common degu Octodon degus); 494-PCPYAC-499 (conserved in all tetrapods except the bearded dragon Pogona viticeps and platypus); and 648-DCPDSF-653 (conserved in almost all tetrapods). Bach in Chona spp. has only two conserved CP sites, and their positions differ from those involved in the heme-dependent regulation in human and mouse. A similar pattern is observed for Bach2. The 368-ACPFNK-373 heme-binding site is present in most Myophroma but not in the Upper Galilee Mountains blind mole rat (GMR) Nannospalax galili; the ACPFDK (N→D) site is found in other rodents (the North American beaver Castor canadensis, guinea pig Cavia porcellus, the GMR, and other mole-rats, DMR and NMR) and nearly all tetrapods; the ACPSDK (F→S) site is found in the Philippine tarier Carilto syrichta; ACSFDK (P→S), in flying foxes; VCPFDK (A→V), in the barbed agama; ACPLDK (F→L), in marsupials and the three-toed box turtle Terrapene carolina triunguis (with a negligible aging rate); ACPFEK, in birds, saltwater crocodile and platypus; and ACPVEK, in the western clawed frog Xenopus tropicalis; ACPFK, in the New Zealand brownbanded bamboo shark Rhincodon typus; and ACPVEK, in the flighting foxes defer from that (commonly ACPFDK) in other chiropteras. The 498-SCPVPI-503 site is strictly conserved in all species starting from cartilaginous fishes to mammals. The 505-VCPRP-510 site is also strictly conserved within the same taxonomic range excluding the L. chalumnae with the C→Y substitution. The 602-SCPQVD-607 site is conserved in the L. chalumnae and all tetrapods. This site is missing in the European cattle Bos taurus, while Bos mutus shares it with all other tetrapods. This site is missing in the bluedspotted mudskipper Boleophthalmus pectinostris as well as in other fishes. The sites notably differ in the C. mili, and whale shark Rhincodon typus from those in tetrapods. The 728-YCPVL-733 site is only found in all rodents but not in other taxa. However, most vertebrates including the European rabbit Oryctolagus cuniculus and C. mili have a different site, YCPVLR (I→R). The Chinese softshell turtle Pelodiscus sinensis demonstrate singular modifications in the YFPVLR (C→F) site; the X. tropicalis, in YCPVLQ (R→Q); the R. typus, in FCPVFR (Y→F, L→F). The evolution of the insertion in the N-hook downstream of the MSLSE motif at the N-terminus in Bach1 has been studied. The evolution of response to mitochondrial ROS will be discussed [1].

Discussion

The genomic rearrangement that gave rise to the Bach remains an open problem. The Nfe2 is most similar to the Bach in early-diverging deuterostomes. One can think that Bach emerged by duplication of Nfe2, an ortholog of Nfe2l2, in the chordate ancestor. The BTB domain essential in Bach is missing in Nfe2; however, it exists in dozens of ancient proteins (e.g., the BTB-ZF family) that could provide it for the ancestral Bach. Possible BTB sources include a gene orthologous to the European lancelet Branchiostoma lanceolatum gene BL03038_cuf1 with the domain at the N-terminus followed by a domain typical for the Zinc finger C2H2 superfamily (IPR036236). The bZIP domain specific for the C-terminal regions of Bach proteins is not found in BL03038_cuf1. Thus, Bach could emerge as a chimeric gene. Anyhow, such an evolutionary scenario is optimal for the reconstruction of the genomic structures carried out by the original algorithm. The absence of Bach proteins in Branchiostoma spp. agrees with the proposed closer phylogenetic similarity between vertebrates and tunicates rather than between vertebrates and lancelets. The absence of Bach in Oikopleura dioica, which is relatively close to Chona spp., can be attributed to its neoteny. Indeed, like other appendicularians, adult O. dioica has a discrete body and tail and preserves the notochord throughout its life, while the body structure of Chona spp. substantially changes in development. The loss or substantial change of Bach2 genes in the tuatara Sphenodon punctatus and the tortoise Cheloniaius abingdonii is accompanied by a long species-specific lifespan. The changes in the NMR Bach1 untypical for rodents can also be related to the unusually long lifespan.

ACKNOWLEDGMENT

We are very grateful to V. P. Skulachev for the original idea, as well as help, advice, and comments on this work. The study was funded by RFBR according to project no. 18-29-13037.

REFERENCES