RESEARCH ARTICLE   **Open Access**

CrossMark

Adaptive patterns in the p53 protein sequence of the hypoxia- and cancer-tolerant blind mole rat *Spalax*

Vered Domankevich[1], Yarden Opatowsky[2†], Assaf Malik[1†], Abraham B. Korol[1], Zeev Frenkel[1], Irena Manov[1], Aaron Avivi[1] and Imad Shams[1*]

Abstract

Background: The subterranean blind mole rat, *Spalax* (genus *Nannospalax*) endures extreme hypoxic conditions and fluctuations in oxygen levels that threaten DNA integrity. Nevertheless, *Spalax* is long-lived, does not develop spontaneous cancer, and exhibits an outstanding resistance to carcinogenesis in vivo, as well as anti-cancer capabilities in vitro. We hypothesized that adaptations to similar extreme environmental conditions involve common mechanisms for overcoming stress-induced DNA damage. Therefore, we aimed to identify shared features among species that are adapted to hypoxic stress in the sequence of the tumor-suppressor protein p53, a master regulator of the DNA-damage response (DDR).

Results: We found that the sequences of p53 transactivation subdomain 2 (TAD2) and tetramerization and regulatory domains (TD and RD) are more similar among hypoxia-tolerant species than expected from phylogeny. Specific positions in these domains composed patterns that are more frequent in hypoxia-tolerant species and have proven to be good predictors of species' classification into stress-related categories. Some of these positions, which are known to be involved in the interactions between p53 and critical DDR proteins, were identified as positively selected. By 3D modeling of p53 interactions with the coactivator p300 and the DNA repair protein RPA70, we demonstrated that, compared to humans, these substitutions potentially reduce the binding of these proteins to *Spalax* p53.

Conclusions: We conclude that extreme hypoxic conditions may have led to convergent evolutionary adaptations of the DDR via TAD2 and TD/RD domains of p53.

Keywords: P53, Hypoxia, Oxidative stress, RPA70, TAD2, Convergent evolution, Cancer resistance, Longevity

Background

The blind mole rat of the genus *Nannospalax* (hereafter, *Spalax*) is a solitary subterranean mammal [1] that experiences extreme and abrupt fluctuations in O_2/CO_2 levels. *Spalax* survives low-oxygen content (~7 % O_2) in its natural underground habitat, and even lower (3 % O_2) under laboratory conditions [2–4]. One of the greatest challenges faced by *Spalax* is repeated exposure to acute hypoxia followed by rapid re-oxygenation, which leads to oxidative stress [5]. Hypoxia and oxidative stress are two types of cellular stressors that place a great risk on cellular

functions and genomic stability [6–8]. Importantly, the combination of these stressors seems to further enhance genomic instability compared to the separate effects of each stress type. Hypoxia depletes dNTPs and represses DNA repair pathways that are required to overcome replication-stress, while re-oxygenation induces replication-restart that coincides with reactive oxygen species (ROS)-induced oxidative DNA damage, at a time when the repair pathways have not yet recovered [9, 10].

Genomic instability underlies both cancer and aging [11]. Nevertheless, *Spalax* does not develop spontaneous tumors and does not show clear age-related phenotypic changes, despite its relatively long lifespan (~20 years in captivity) [12]. Moreover, *Spalax* displays an outstanding tolerance to chemically induced carcinogenesis in vivo,

* Correspondence: imad.shams@univ.haifa.ac.il
†Equal contributors
[1]Institute of Evolution & Department of Evolutionary and Environmental Biology, University of Haifa, Haifa, Israel
Full list of author information is available at the end of the article

Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 2 of 15

and its fibroblasts inhibit cancer growth in vitro [13]. Another long-lived (~30 years) subterranean rodent in which spontaneous cancer is an uncommon phenomenon, relative to similar-sized rodents, is the naked mole rat, *Heterocephalus glaber* (hereafter, *H. glaber*) [14, 15]. *H. glaber* is phylogenetically distant from *Spalax* (estimated divergence time is 77.9 million years) [16]. Yet, similar to *Spalax*, it shows adaptations to hypoxic stress [17] and in vitro ability to inhibit cancer cell growth [13]. Bearing in mind the well-known strong positive correlation between aging and cancer epidemiology in many metazoans, including humans [18], common features in *Spalax* and *H. glaber* deserve special attention. Such features may relate to evolutionary adaptations to common environmental conditions that might play a compensatory role in dealing with consecutive stress-induced DNA damage.

The subterranean habitat is not the only environment in which oxygen fluctuations may occur. Diving mammals experience rapid transitions from apnea to re-oxygenation [19], while hibernating mammals face reduced metabolism and fluctuations in blood flow and oxygen consumption [20–23]. "Metabolic shutdowns" happen also in desert mammals, such as the desert mouse [24] and the jerboa [25]. Evidence for oxidative stress or antioxidant adaptations were indeed found in *Spalax* [26], diving mammals [19, 27], hibernating mammals [22], and desert mammals [28]. Bats are another example of mammals showing antioxidant adaptations [29], as might be expected from the high metabolic stress these mammals experience during flight [30, 31]. Interestingly, bats' lifespan increases with hibernation, body mass, and the occasional use of caves [32]. Their exceptional longevity is explained by multiple mechanisms for resisting oxidative damage [33] and is attributed to positively selected genes in DNA-damage-checkpoint pathways [34]. The above mentioned studies indicate that environmental conditions, which include extreme changes in oxygen supply, may require adaptations to cope with oxidative stress that are not restricted to antioxidant activity, but also include changes in the DNA-damage response (DDR) as a second line of defense.

Sophisticated DDR pathways constantly monitor genome integrity [35] and delay or stop cell-cycle progression at critical stages in response to unrepaired DNA damage, thereby preventing replication of the damaged DNA [36]. The tumor suppressor protein p53 interacts with multiple proteins participating in complex pathways of the DDR [37]. It plays an important role in cellular pathways that control genomic instability and is involved in critical defense and regulatory pathways such as cell-cycle arrest, senescence, and apoptosis [35]. The tight control by senescence and apoptosis prevents the uncontrolled proliferation of damaged cells; however, it may also deplete stem and progenitor cell pools, thus promoting tissue degeneration and aging [35]. This emphasizes the

subtle role of p53 in regulating pathways related to cancer and aging [38]. *Spalax* has acquired an R174K amino acid substitution in the DNA-binding domain (DBD) of p53. Mutation at this position in human p53 leads to impaired induction of apoptotic and cell-cycle arrest genes and is known as a frequent mutation found in many human cancers [39, 40]. The fact that *Spalax* does not develop spontaneous cancer implies that there are changes in other regions of the p53 protein and/or other related proteins to ensure integrity of the related signaling pathways and to maintain homeostasis.

Due to the critical role of p53 in the regulation of apoptosis and DNA repair, we aimed to investigate whether *Spalax* p53 protein sequence shares structural features with other phylogenetically distant mammals that are also adapted to acute-fluctuating hypoxia. Such features may indicate joint adaptations of the cellular stress response via p53 that are related to hypoxic stress resistance. Here, we have identified changes in p53 domains, which are known to include intrinsically disordered regions [41] that are highly diverse in evolution [42], and found that these domains' sequences are similar among mammals adapted to stressful hypoxic conditions, more than expected from phylogeny. These domains harbor binding sites of p53 with proteins that participate in the DDR and in certain metabolic pathways such as replication protein A 70-kDa DNA-binding subunit (RPA70), the histone acetyltransferase p300, and silent information regulator 2 (Sir2) [43–45]. Thus, changes in these domains could modulate the way in which p53 orchestrates these pathways under extreme hypoxic conditions in the subterranean habitat.

Results

Different intra-group similarities for hypoxia-tolerant and hypoxia-sensitive taxa

We hypothesized that certain functional domains in p53 have convergently evolved among hypoxia-tolerant species. If this is true, it is expected, for example, that some domains in *Spalax* p53 are more similar to the corresponding domains of other hypoxia-tolerant (HT) species, such as diving mammals, than to those of hypoxia-sensitive (HS) species, such as murine rodents, although *Spalax* is much closer to murines. To test this, we used a balanced experimental design (Fig. 1a) in which a single representative HT species and a single representative HS species is tested for different taxonomic groups (e.g., carnivores, rodents, even-toed ungulates, insectivores, etc.). In public databases we were able to find five representative HT species whose p53 genes were sequenced, belonging to five different mammalian taxonomic orders. Accordingly, five pairs of closely related species were used, where in each pair one is HT and the other is HS. In this test, amino acid sequence distances were calculated according to the Poisson correction model (see Methods) for
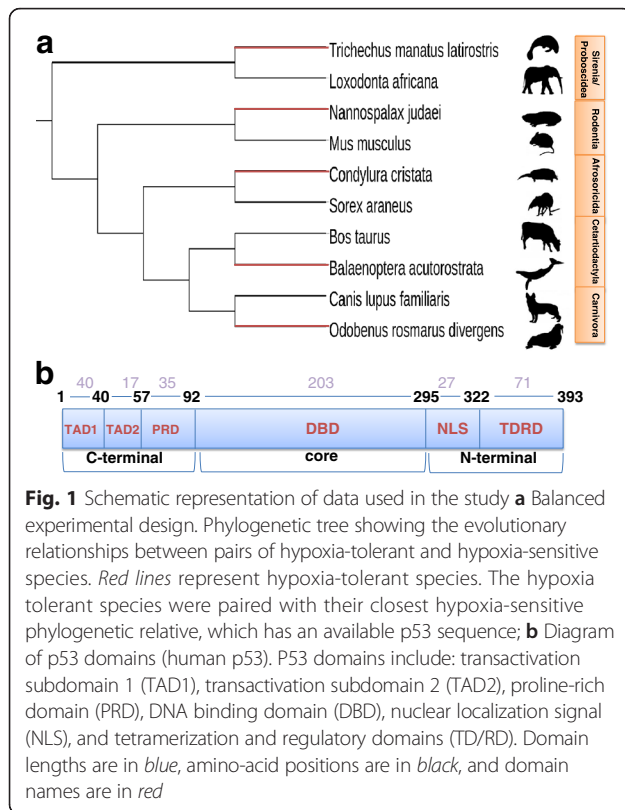
Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 3 of 15

Fig. 1 Schematic representation of data used in the study **a** Balanced experimental design. Phylogenetic tree showing the evolutionary relationships between pairs of hypoxia-tolerant and hypoxia-sensitive species. *Red lines* represent hypoxia-tolerant species. The hypoxia tolerant species were paired with their closest hypoxia-sensitive phylogenetic relative, which has an available p53 sequence; **b** Diagram of p53 domains (human p53). P53 domains include: transactivation subdomain 1 (TAD1), transactivation subdomain 2 (TAD2), proline-rich domain (PRD), DNA binding domain (DBD), nuclear localization signal (NLS), and tetramerization and regulatory domains (TD/RD). Domain lengths are in *blue*, amino-acid positions are in *black*, and domain names are in *red*

the following p53 domains (Fig. 1b): transactivation subdomain 1 (TAD1), transactivation subdomain 2 (TAD2), proline-rich domain (PRD), DNA binding domain (DBD), nuclear localization signal (NLS), and tetramerization and regulatory domains (TD/RD). For each domain in p53, it was then tested whether the pairwise distances within the group involving HT species are smaller than those of the matched group of HS species (H_1 hypothesis). In addition, the test was repeated using the estimated pairwise divergence time published in the literature [16].

Following the test described in Methods, we calculated Mann–Whitney statistics U that characterizes the difference in intra-group similarities (based on all-to-all distances in each group). *P*-values for this statistics were calculated in two ways: using U distribution (Mann-Whitney

[46]) and distribution obtained by Monte-Carlo simulations (Table 1). Monte-Carlo simulations were used to avoid biases caused by intra-group dependencies derived from the usage of each species more than once when calculating the distances of all possible pairs of species in each group (see Methods). Monte-Carlo simulations were conducted with $k_{HT} = k_{HS} = 5$ and $n_{HT} = n_{HS} = 5*4/2 = 10$, where k represents the number of species in each group (five in our test) and n represents all possible pairs of species in each group (ten in our test). U-statistic for TAD2, DBD, and TD/RD domains was significantly smaller (with FDR adjusted $p < 0.05$), allowing the rejection of the H_0 hypothesis that components of vectors S_{HT} and S_{HS} have the same joint distribution (see Methods). These observed significant differences are unlikely to be explained only by phylogeny, as no significant differences in the intra-group domain similarities were detected using the estimated pairwise divergence time.

To further verify that the obtained results remain consistent when a wider range of species is used (see Methods), we also compared *Spalax* p53 protein sequence, as a reference sequence to: (1) the entire NCBI *nr* database; (2) p53 sequences of 46 other species, which include all HT species whose p53 sequence was sequenced, as well as additional species that are adapted to other types of metabolic stress (Additional file 1: Table S1, Additional file 2: Figure S1, hypoxic- metabolic- and non-stress groups). The results of these tests (Additional file 3: Table S2) were consistent with the results based on the balanced experimental design for TAD2 and TD/RD (but not DBD), and indicated that these domains in *Spalax* are more similar to those of HT species than to those belonging to species with the highest phylogenetic relatedness to *Spalax*.

Stress-related patterns in p53 TAD2, TD, and RD

The results above may indicate that unique sequence patterns have convergently evolved in some p53 domains of hypoxia-tolerant species. Accordingly, we investigated candidate sequence patterns in p53 that could specify the similarity between species adapted to stressful hypoxic conditions. Here, "pattern" means the same combination of residues, in the same order, with a constant number of

Table 1 Domains with different intra-group similarities for hypoxia-tolerant (HT) and hypoxia-sensitive (HS) taxa

	Evolutionary tree	TAD1	TAD2	PRD	DBD	NLS	TDRD
U_{obs}	35	49	12	74	9	50	3
p_{MW}	0.2791	0.4844	0.0015	0.0313	0.0005	0.4861	0.0001
p_{MW}-FDR		0.4861	0.0029	0.0468	0.0015	0.4861	0.0003
p_{MC}	0.413	0.4895	0.0095	0.0768	0.0046	0.4882	0.0005
p_{MC}-FDR		0.4895	0.019	0.1152	0.0140	0.4895	0.0030

Statistics U and p-values were calculated for domain sequences of p53 and for evolutionary tree positions (see Methods). P_{MW} for p53 domains corresponds to p-value calculated in a standard one-tailed Mann–Whitney test [46]. P_{MC} correspond to p-values calculated in our extension on U test using Monte-Carlo simulations (see Methods). As can be seen it the table, the estimated 'divergence time' calculated based on evolutionary tree is unlikely to explain differences between the HT and HS groups, seen in the TAD2, DBD, and TDRD domains

Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 4 of 15

residues between them. We have identified two types of patterns in TAD2 and TD/RD domains of p53: (1) patterns that are more frequent in species from the hypoxic-/metabolic-stress groups relative to species from the non-stress group; (2) patterns that are characteristic solely to all "hypoxic-stress" species (Additional file 1: Table S1) and exist in none of the species included in the metabolic-/non-stress groups.

To identify the first type of patterns, the complete p53 sequences of 47 species from the three above-mentioned groups were aligned by MAFFT (sorted by input order that was determined according to the respective groups). Positions that led to incorrect matching between homologous residues were manually corrected. Common patterns in TAD2 and TD/RD were manually identified and the frequencies of these patterns in the hypoxic-/

metabolic-stress groups relative to the non-stress group were subjected to Fisher's 2×2 exact test. In addition, the significance of these patterns in all three groups was tested by Fisher's 2×3 exact test. This data set was used as it represents a wide range of species in terms of stress adaptations, so that (1) we could verify that the patterns identified in p53 protein sequence exist in most or all of hypoxia-tolerant species (2) patterns identified as different between the two extremes (hypoxic vs. non-stress groups), could be further evaluated in mammals experiencing other types of stresses (metabolic stress group). Nevertheless, as this data set is not absolutely phylogeneticaly balanced (see Methods), only the identified patterns for which the *p-value* was smaller than 0.005 were selected and retested in the mentioned balanced experimental design (see Fig. 1a).

```
a  Species                              TAD2                   TD                                        RD        b  TAD2

Spalax_judaei                        EDLLLSPEDVANWLDD---PD   TQAEKDSGESRAHSSYLKSKKGQST-SRHKK-LM------      ------EDLLLSPEDVANWLDD---PD
Eospalax_cansus                      EDLLLSAEDVANWLDD---PD   AQAEKESGESRAHSSYLKSKKGQST-SRHKK-LM------      ------EDLLLSAEDVANWLDD---PD
Eospalax_baileyi                     EDLLLSAEDVANWLDD---PD   AQAEKESGESRAHSSYLKSKKGQST-SRHKK-LM------      ------EDLLLSAEDVANWLDD---PD
Heterocephalus_glaber                DDLLLSPEDVVNWLGGN--PD   AQTEKEPGESRPHSSYLKSKKGQST-SCHKK-LM------      ------DDLLLSPEDVVNWLGGN--PD
Fukomys_damarensis                   DDLLLSPEDVVNWLGGN--PD   AQTEKEPGESRPHSSYLKSKKGQST-SSHKK-LM------      ------DDLLLSPEDVVNWLGGN--PD
Condylura_cristata                   EE-LLSSENVANWLDEA--PN   AQAGKEPGGSRAHSHLKSKKGQST-SHHKK-LM------       ------EELLSS-ENVANWLDEA--PN
Orcinus_orca                         DDLLLSPEDVANWLDER--PD   AQAGKEPGESRAHSSHLKSKKGPSP-SRHKK-LM------      ------DDLLLSPEDVANWLDER--PD
Tursiops_truncatus                   DDLLLSPEDVANWLDER--PD   AQAGKEPGESRAHSSHLKSKKGPSP-SRHKK-LM------      ------DDLLLSPEDVANWLDER--PD
Lipotes_vexillifer                   DDLLLSPEDVANWLDER--PD   AQAGKEPGESRAHSSHLKSKKGQSP-SRHKK-LM------      ------DDLLLSPEDVANWLDER--PD
Physeter_catodon                     DDLLLSPEDVANWLDER--PD   AQAGKEPGESRAHSSHLKSKKGQSP-SRHKK-LM------      ------DDLLLSPEDVANWLDER--PD
Balaenoptera_acutorostrata_scammoni  DDLLLSPEDVANWLDER--PD   AQAGRELGESRAHSSHLKSKKGQSP-SRHKK-LM------      ------DDLLLSPEDVANWLDER--PD
Delphinapterus_leucas                DDLLLSPEDVANWLDER--PD   AQAGKEPGESRAHSSHLKSKKGQSP-SRHKK-LM------      ------DDLLLSPEDVANWLDER--PD
Odobenus_rosmarus_divergens          DE-LLLSEGVANWLDAG--SD   ALSGKEPGGSRAHSSHLKSKKEQST-SRHKK-LM------      ------DELLLS-EGVANWLDAG--SD
Leptonychotes_weddellii              DE-LLLSEGVANWLDTG--SD   ALSGKEPGGSRAHSSHLKAKKGQST-SRHKK-LM------      ------DELLLS-EGVANWLDTG--SD
Trichechus_manatus_latirostris       DD-LLLTEDAATWLES---QV   AQAGKEPEGSRAHSSPPKSKKGQST-SRHKK-LM------      EVVEVADDLLLT-EDAATWLES---QV
Tupaia_belangeri                     DDLMLSPDDIEQWFTEDPGPD   AMAGKESAGSRAHSSHLKSKKGQST-SRHKK-LM------      ------DDLMLSPDDIEQWFTEDPGPD
Marmota_monax                        DDLLLSSEDVENWFDKG--PD   AQAEKEPGESRPHPSYLKSKKGQST-SRHKK-II------      ------DDLLLSSEDVENWFDKG--PD
Ictidomys_tridecemlineatus           DDLLLSSEDVENWFDKG--PD   AQAEKEPGESRPHPSYLKSKKGQST-SRHKK-II------      ------DDLLLSSEDVENWFDKG--PD
Ochotona_princeps                    -EELLST-NMTNWLEDL--GD   AQAEKEPGGSKVHSSYLKSNKGQST-SRHKK-MM------      ------EELLST-NMT-NWLEDL--GD
Chinchilla_lanigera                  DDLLLCPEDVVNWLEEN--PD   AQTEKEPGESRPHSSYLKSKKGQST-SCHKK-PM------      ------DDLLLCPEDVVNWLEEN--PD
Jaculus_jaculus                      EDLLLPPEYDENWLVEG--RD   AQAEKEPGETKAQSSYLKSKKGQST-SRHKK-LM------      ------EDLLLPPEYDENWLVEG--RD
Meriones_unguiculatus                ED-LLLPQDVTSWLGD---AD   AQAAGESGDGRAQASCLKTKKDKST-SPRKN-PM------      ------EDLLLPQDVTS-WLGD---AD
Peromyscus_maniculatus_bairdii       EDLFLS-QDVANWLEG---PD   AHAAKGSGDSRAHSSYLKSKKEQST-SRHKK-LM------      ------EDLFLSQD-VANWLEG---PD
Mesocricetus_auratus                 EELFLS-ENVANWLED---PG   AQALKASEDSGAHSSYLKSKKGQSA-SRLKK-LM------      ------EELFLS-ENVANWLED---PG
Cavia_porcellus                      DHLLLSPEEVASWLGEN--PD   AQTEKEPGESRPHSSYPKSKKGQST-SCHKK-LM------      ------DHLLLSPEEVASWLGEN--PD
Camelus_ferus                        DELLLSPE-VTNWLDEN--PD   AQAGKEPGENKAHSSHLKSKKGQSP-SRHKK-LM------      ------DELLLSPE-VTNWLDEN--PD
Vicugna_pacos                        DELLLSPE-VTNWLDEN--PD   AQAGKEPGENKAHSSHLKSKKGQSP-SRHKK-LM------      ------DELLLSPE-VTNWLDEN--PD
Bos_mutus                            DD-LLPYTDVATWLDEC--PN   ALDGREPGESRAHSSHLKSKKRPSP-SCHKK-PM------      ------DDLLPYTD-VATWLDEC--PN
Pantholops_hodgsonii                 DDLLPYSEDVATWLDEC--PN   AQAGREPGESRAHSSHLKSKKGPSP-SCHKK-PM------      ------DDLLPYSEDVATWLDEC--PN
Ursus_maritimus                      DE-LLLSEGVVNWMDEG--SD   ALSGKEPGGSRAHSSHLKAKKGQST-SRHKK-LM------      ------DELLLS-EGVVNWMDEG--SD
Myotis_davidii                       GGLDWMPFSLEDCLDNG--PN   VLAGKDPGGSKTHSSHLKPKKGQST-SRHKKRLM------      ------GGLDWMPFSLEDCLDNG--PN
Sarcophilus_harrisii                 DDIQFASPDSINWLENE--EE   AHSGKDPEGNCPHRSQLKSKRGDSTPCQGKR-LL------      ------DDIQFASPDSINWLENE--EE
Homo_sapiens                         DDLMLSPDDIEQWFTEDPGPD   AQAGKEPGGSRAHSSHLKSKKGQST-SRHKK-LM------      ------DDLMLSPDDIEQWFTEDPGPD
Galeopterus_variegatus               DNLILSPEEVADWLSED--PD   AQTGKEPGGSRAHSRHLKSKKGQST-PRHKK-IM------      ------DNLILSPEEVADWLSED--PD
Otolemur_garnettii                   DDLMLSP-DIVNWFDEG--PD   AQAEKESDGNRAHSSQPKSKKEQST-SRHKK-LM------      ------DDLMLSPDIV-NWFDEG--PD
Oryctolagus_cuniculus                DD-LLSAEDVANWLNED--PD   AQAEKEPGGSRAHSSYLKSKKGQST-SRHKK-PM------      ------DDLL-SAEDVANWLNED--PE
Rattus_norvegicus                    EDLFL-PQDVAELL-EG--PE   ARAAEESGDSRAHSSYPKTKKGQST-SRHKK-PM------      ------EDLFL-PQDVAELL-EG--PE
Mus_musculus                         DDLLL-PQDVEEFF-EG--PS   AHATEESGDSRAHSSYLKTKKGQST-SRHKK-TM------      ------DDLLL-PQDVEEFF-EG--PS
Sus_scrofa                           NDLLLSP--VTNWLDEN--PD   AQTARESGENKAHSSHLKSKKGQSP-SRHKK-PM------      ------NDLLLSP--VTNWLDEN--PD
Bos_taurus                           DD-LLPYTDVATWLDEC--PN   ALDGREPGESRAHSSHLKSKKRPSP-SCHKK-PM------      ------DDLLPYT-DVATWLDEC--PN
Ovis_aries                           DDLLPYSEDVVTWLDEC--PN   AQAGREPGESRAHSSHLKSKKGPSP-SCHKK-PM------      ------DDLLPYSEDVVTWLDEC--PN
Capra_hircus                         DDLLPYSEDVVTWLDEC--PN   AQAGREPGESRAHSSHLKSKKGPSP-SCHKK-PM------      ------DDLLPYSEDVVTWLDEC--PN
Equus_caballus                       NNLLLSP-DVVNWLDEG--PD   AQTGKEPGGSKAHSSHLKSKKGQST-SSHKK-LI------      ------NNLLLSP-DVVNWLDEG--PD
Felis_catus                          NELPLS-EDVANWLDEA--PD   AQSGKEPGGSRAHSSHLKAKKGQST-SRHKK-PM------      ------NELPLS-EDVANWLDEA--PD
Canis_lupus_familiaris               DE-LLLPESVVNWLDED--SD   AQSGKEPGGSRAHSSHLKAKKGQST-SRHKK-LM------      ------DELLLP-ESVVNWLDED--SD
Mustela_putorius_furo                DE-LLS-EGVNWLGEG--SN    ALNGKEPGGSRPHSRSI-PQPGSSL-GRSPD-PSRGEQWG      ------DELLS--EG-VNWLGEG--SN
Ceratotherium_simum_simum            DELLLSPD-VSTWLDED--PD   AQAGKEPGGSRAHSSHLKSKKGQST-SCHKK-LI------      ------DELLLSPD-VSTWLDED--PD
```

Fig. 2 Distribution of positions composing stress-related patterns in p53 TAD2 and TD/RD: **a** Patterns that are more frequent in species included in the "hypoxic" (*red*)/"metabolic" (*green*) "-stress" groups relative to those included in the "non-stress" (*blue*) group. Residues composing the patterns are marked in bold-red (LLXXE in TAD2, KXXGE and LM in TD/RD). An additional pattern was more frequent only for the "hypoxic-stress" species relative to "non-stress species" (LLXXEXXA in TAD2), marked in *bold red and black*; **b** Pattern identified in all "hypoxic-stress" species and in none of the species included in either the "metabolic-" or "non-" "stress" groups. This pattern is a combination of two motifs: [DE]LLXST (*bold red*) and [DE]XX[AV]XWL (*bold black*)

Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 5 of 15

Among the identified patterns (Fig. 2) that were found to be significantly more frequent in species included in the hypoxic-stress group relative to the non-stress group (Additional file 4: Table S3), some of the patterns were also found in the metabolic-stress group, but with a lower level of significance. In TAD2, the pattern LLXXE was significantly more frequent in the hypoxic ($p < 0.0001$) and metabolic ($p < 0.05$) groups compared to the non-stress group. Position A52 (*Spalax* p53; corresponds to E51 in human p53), and its related patterns EXXA and LLXXEXXA, were significantly more frequent only in the hypoxic-stress group ($p < 0.005$ and 0.0001, respectively). In TD/RD, the patterns KXXGE and LM were significantly more frequent in the hypoxic-stress group ($p < 0.0001$ for both patterns) and in the metabolic-stress group ($p < 0.005$ and 0.05, respectively) relative to the non-stress group. The patterns LLXXEXXA, EXXA and LM in TAD2 and RD (but not in TD) were still significantly more frequent in the hypoxic-stress group vs. the non-stress group of the balanced experimental design ($p < 0.05$), despite the small number of species included in each group (Additional file 4: Table S3, Additional file 5: Figure S2).

The next step of our analysis was to check whether the residues participating in the patterns identified in TAD2, TD, and RD (see shaded columns in Fig. 2a) are sufficiently informative to enable the robust classification of species into the three groups: hypoxic-stress, metabolic-stress, and non-stress. We used the "classification trees" algorithm to classify an extended list of mammals (66 rather than the initial 47 species, see Additional file 1: Table S1, Additional file 6: Figure S3), based on the identified positions in p53 as predictors (see Methods). The classification (Additional file 4: Table S3, Additional file 7: Figure S4) showed that all species from the hypoxic-stress group (subterranean and diving mammals) had very high probabilities of being assigned to this group. An analogous result (namely, a classification that corresponds to the subjective assignment) was obtained for the species that belonged to the non-stress group. In the metabolic-stress group, most of the species showed high probability of being assigned to stress-related groups (mostly the metabolic-stress group). Interestingly, in this category, mammals with large body mass tended to have lower probabilities of being classified into stress groups. Nineteen (19) additional species (that are not included in any of the hypoxic/metabolic/non-stress groups, see Methods) were approximately equally distributed between non-stress and stress-related groups. The most important predictor (Additional file 4: Table S3) was found to be p53 position 51 (human p53) in TAD2. These results provide further support that adaptation to stressful hypoxic conditions is predicated by changes at the identified positions.

A second type of pattern in TAD2 (Fig. 2b) domain defines unambiguously (one-to-one) the species

included in the hypoxic-stress group. This pattern can be described as a combination of two motifs: [DE]LLXST and [DE]XX[AV]XWL. The combination of these two motifs was found in all hypoxia-tolerant species, but was not found in any of the other species included in the sample (total 47 species). We then checked whether this pattern exists in the 19 additional species (out of the total 66, see Additional file 1: Table S1). The pattern was not found in any of these species, with one exception: the African elephant. In other words, among 66 species tested, this pattern was found only in hypoxia-tolerant species and in the African elephant. Nevertheless, for the African elephant, the pattern in p53-RD (LM) was not found, and thus it is distinguished from the hypoxic-stress group.

In addition, as TAD2 of p53 was shown to resemble TAD2 domain of Krueppel-like factor 1 (KLF1) [47], we also compared the sequences of KLF1-TAD2 domain of hypoxia-tolerant species to those of hypoxia-sensitive species (according to the balanced experimental design). We did not identify any hypoxia-specific pattern (Additional file 5: Figure S2) in KLF1-TAD2. This may indicate that the considered patterns are associated with unique functionality of p53 rather than with a functionality related to a global hypoxic response.

3D models of p53 in complex with other DDR proteins provide evidence of a unique DDR in *spalax*

To investigate the possible functional role of the residues identified in p53 as 'stress-related' (Fig. 2), we first allocated the available structures of p53 in complex with other proteins in which the interaction includes these residues. We subjected TAD2 and TD/RD sequences to protein BLAST, querying the Protein Data Bank (PDB) database. Only a few structures of the proteins that interact with p53 via these domains have been resolved for human proteins to date (for main examples see Additional file 8: Table S4). We then examined the literature related to these structures, and found that the residues included in the stress-related patterns are indeed directly involved in the interactions of p53 with some of the proteins (Table 2). For example, the pattern 'EXXA' (*Spalax* p53) corresponds to residues 'DXXE' (positions 48–51, human p53), which are responsible for electrostatic interactions between p53 and other proteins such as RPA70 and p62.

In addition, we observed large differences between *Spalax* and humans with respect to p53 interactions with proteins involved in DDR pathways via TAD2, as demonstrated in Additional file 9: Figure S5. Therefore, we hypothesized that p53's interactions with these proteins are different in *Spalax* compared to humans. We then specifically examined p53's interaction with RPA70 by model replacement of the human p53 residues participating in the direct interface with RPA70, with the

Table 2 Residues included in the stress-related patterns participate in the binding of p53 with DDR proteins

p53 interaction with protein (PDBID)/pattern residues sp53 (hp53)	*LL (ML)*	*EXXA (DXXE)*	*LM (LM)*
P300 (2MZD)	L45 binds to nonpolar patches on the surface of Taz2 domain of p300	Salt bridges and hydrogen bonds between D48 and E51 and Taz2 domain of p300	
RPA70 (2B3G)		Electrostatic interactions of D48 and E51 with RPA70N	
CBP (2 L14, 1JSP)	Hydrophobic interactions of M44,L45 with the NCBD domain of CBP		Hydrophobic interactions of L383 and M384 with the bromodomain of CBP
P62 (2RUK)		Electrostatic interaction of D48 and E51 with the PH domain of p62	
Tfb1 (2GS0)		Electrostatic interaction of E51 with Tfb1	
HMGB1 (2LY4)	Hydrophobic interaction of M44,L45 with HMGB1		
Sir2 (1MA3)			Hydrogen bond of L383 with sir2
S100B (1DT7)			Hydrophobic interaction of L383 with S100B

Some of the residues composing the stress-related patterns in TAD2 and RD participate in *Spalax* (s) p53 direct interface with DDR proteins. Specific interactions via these residues between DDR proteins and human (h) p53 are described for each protein. The data for each protein was obtained from manuscripts related to the resolved structures of the protein complex with p53 in humans (PDBID is mentioned in the table for each protein)

corresponding residues in *Spalax*. From this analysis, it seems that the human p53–RPA70 interaction interface involves more hydrophobic and ionic interactions relative to *Spalax* and therefore, the interaction in *Spalax* is less stable (Fig. 3a). Importantly, the residues that mediate the interaction with p53 are identical in humans and *Spalax* in the RPA70 protein (data not shown), unlike the interface-residues in the p53 protein, which are different between the two species. This suggests that p53 and RPA70 did not co-evolve and strengthens the conclusion that the interaction between them is different

in *Spalax* compared to humans. Similar results were obtained when we modeled *Spalax*-p53's interaction with p300 via TAD2. The interface in the p300 side of the interaction was conserved (data not shown); however, we have identified reduced electrostatic interactions in *Spalax* compared to humans due to different interface-residues in p53 (Fig. 3b). We then used MutaBind server to estimate the putative changes in the binding affinity of *Spalax* p53 interactions with other DDR proteins (Additional file 8: Table S4). This analysis supported the possibility that *Spalax*'s stress-related substitutions in TAD2 reduce
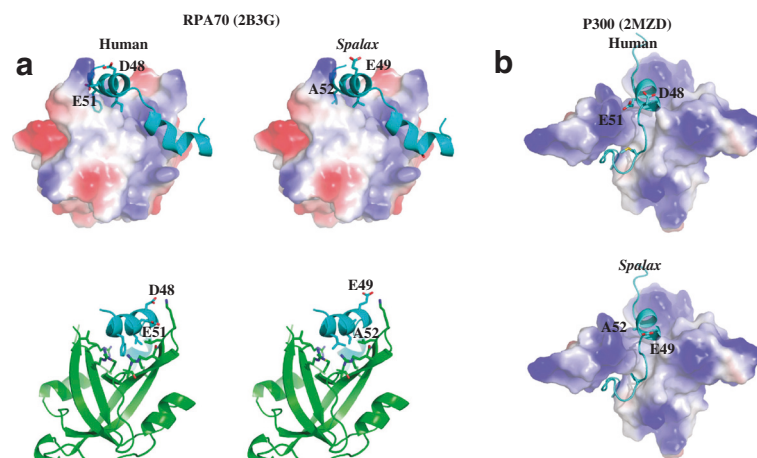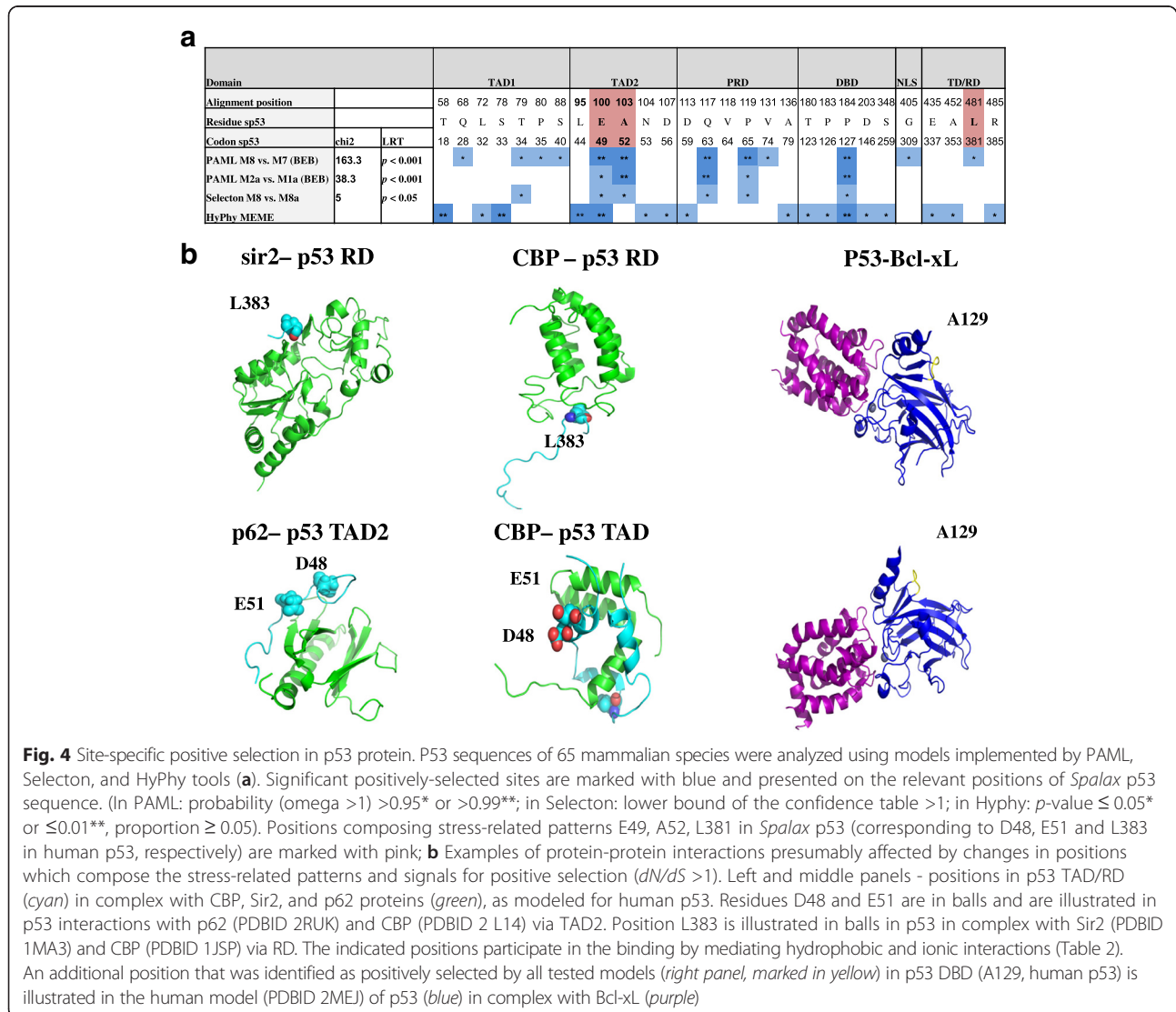

Fig. 3 Model of *Spalax* p53 N-terminal (p53N) interaction with: **a** RPA70 repair protein. *Left* – Crystal structure of human p53N fragment (33–60) bound to RPA70N (PDBID 2B3G). The p53 fragment is depicted as a cyan helix and RPA70N as an electrostatic surface (*upper panel*) or as green ribbons (*lower panel*). *Right* – Model of the *Spalax* p53–RPA70 complex, based on the human complex. **b** P300 co-activator. *Upper panel* - The structure of human p53N bound to p300 (PDBID 2MZD). *Lower panel* - Model of the *Spalax* p53–p300 complex, based on the human complex. From this analysis, it seems that the human p53 interface with RPA70 and P300 involves more hydrophobic and ionic interactions than the *Spalax* interface, making the complexes more stable

the binding affinity of p53 to other DDR proteins such as RPA70 and p62, compared to humans. For the stress-related pattern in TD/RD (owned by both *Spalax* and humans) it was demonstrated that this pattern potentially increases the binding affinity of p53 to proteins such as sir2, compared to other species that do not own the pattern.

Site-specific positive selection in p53

The adaptive role of the sites in which residues composing the above-mentioned stress-related patterns are located was further investigated. We hypothesized that specific p53 amino acid sites in TAD2 and TD/RD were positively selected during evolution. By 'positive selection', we mean amino acid substitutions that conferred adaptive advantages under selective pressure. PAML package and Selecton online-server were used to estimate

codons' non-synonymous vs. synonymous substitution rates (dN/dS). A likelihood ratio test (LRT) was conducted comparing two nested models: a null model that assumes no positive selection and an alternative model that assumes positive selection. The same 65 sequences were also analyzed using the HyPhy model MEME (Mixed Effects Model of Episodic Diversifying Selection), which can identify sites that were positively selected only in a small fraction of the tree branches. Positions E49, A52 in *Spalax* p53-TAD2 (corresponding to D48, E51 in the human p53) and position L381 in *Spalax* p53-RD (corresponding to L383 in the human p53), which participate in the stress-related patterns, were significantly identified as positively selected by 4, 3, and 1 models out of 4, respectively (see Fig. 4a, Additional file 10: Table S5). The putative role of these positions is illustrated in the 3D structures of p53's interactions with several proteins via TAD2 and

Fig. 4 Site-specific positive selection in p53 protein. P53 sequences of 65 mammalian species were analyzed using models implemented by PAML, Selecton, and HyPhy tools (**a**). Significant positively-selected sites are marked with blue and presented on the relevant positions of *Spalax* p53 sequence. (In PAML: probability (omega >1) >0.95* or >0.99**; in Selecton: lower bound of the confidence table >1; in Hiphy: *p*-value ≤ 0.05* or ≤0.01**, proportion ≥ 0.05). Positions composing stress-related patterns E49, A52, L381 in *Spalax* p53 (corresponding to D48, E51 and L383 in human p53, respectively) are marked with pink; **b** Examples of protein-protein interactions presumably affected by changes in positions which compose the stress-related patterns and signals for positive selection (dN/dS >1). Left and middle panels - positions in p53 TAD/RD (*cyan*) in complex with CBP, Sir2, and p62 proteins (*green*), as modeled for human p53. Residues D48 and E51 are in balls and are illustrated in p53 interactions with p62 (PDBID 2RUK) and CBP (PDBID 2 L14) via TAD2. Position L383 is illustrated in balls in p53 in complex with Sir2 (PDBID 1MA3) and CBP (PDBID 1JSP) via RD. The indicated positions participate in the binding by mediating hydrophobic and ionic interactions (Table 2). An additional position that was identified as positively selected by all tested models (*right panel, marked in yellow*) in p53 DBD (A129, human p53) is illustrated in the human model (PDBID 2MEJ) of p53 (*blue*) in complex with Bcl-xL (*purple*)

Domankevich et al. BMC Evolutionary Biology (2016) 16:177

Page 8 of 15

RD (Fig. 4b). Accordingly, since TAD2 seems to have a unique stress-related structure and function, positions E49 and A52 (Spalax p53) inside this domain are promising candidates for further analysis. Notably, all tests implemented in Selecton, PAML, and Hyphy identified another residue in the DBD of p53 as a significantly positively selected site: P127 in Spalax p53, corresponding to A129 in human p53. This position is located in a loop that fine-tunes the angle between two stapled beta-strands (Fig. 4b). The importance of this position could be related to the role of the two beta-strands' edges. For example, these edges include positions such as K120 (human p53), which is directly involved in p53 interactions with the DNA and the anti-apoptotic protein Bcl-xL [48].

Discussion

In this study, we identified similarities in p53 functional domains that are shared between mammals adapted to stressful hypoxic conditions. TAD2 and TD/RD of p53 are rapidly evolving, compared to other domains in p53, such as the DBD [49]. Nevertheless, they are similar between mammals that are not phylogenetically closely related, but share the ability to endure extreme changes in oxygen levels. We suggest that this similarity is an indication of convergent evolution, resulting from selective pressures that induced common molecular adaptations via these p53 domains. We further identified sequence patterns in these domains that are more frequent in mammals adapted to acute-fluctuating hypoxia and demonstrated that the residues participating in the identified patterns can predict species classification according to stress characteristics. An additional pattern in TAD2 was observed in all hypoxia tolerant-species and in none of the other species participating in this study (a total of 66 tested), with one exception: the African elephant. Interestingly, this mammal is cancer-resistant and expresses a unique DNA damage response that is attributed to mechanisms related to p53 [50]. The latter pattern is a combination of two motifs: [DE]LLXST, which includes Ser46 and may relate to the binding of proteins that induces its phosphorylation, and [DE]XX[AV]XWL, which includes a motif that is parallel to a motif found in p53-TAD1, which is attributed to the binding of p53 with MDM2 [51] and thus may relate to the competitive binding of MDM2 to p53-TAD2 and p53-TAD1 [52].

Protein sequence motifs, which are located in rapidly evolving intrinsically disordered regions of a hub protein and include segments that mediate its binding with other proteins, fits well with the definition of Short Linear Motifs (SLiMs) [53–55]. These motifs represent a functional change in a flexible low-affinity binding region, which may divert the regulation of a complex signaling network. We suggest that the patterns identified in this study may potentially serve as SLiMs that have evolved among hypoxia-tolerant species. The evolution of SLiMs seems to be very complex and is not yet well understood. Nevertheless, recent theories based on evidence from the past decade [56] suggest that the appearance of SLiMs could be random and "evolutionarily transient" (i.e., ex nihilo lost and gained along lineages). This unique evolutionary process may minimize the dependencies between phylogenetically related species and between species and their ancestors; however, it does not contradict events of positive and negative selection along their evolution. In this study, we also showed that positions of residues that make up the identified patterns (E49 and A52 in Spalax p53, corresponding to D48 and E51 in human p53), can be considered as positively selected. Importantly, these residues stabilize electrostatic interactions between p53 and proteins such as the cofactor and histone acetyltransferase p300 and the DNA-repair proteins RPA70 and Tfb1/p62 [57–61].

The intrinsically disordered nature of TAD2 and RD enables them to interact with a range of proteins that function as a complex network. Among them are DNA-repair proteins (Additional file 8: Table S4) that when defected cause susceptibility to a variety of congenital cancers and progeroid syndromes (genetic disorders that mimic physiological aging). These syndromes are attributed to the failure of genome-maintenance mechanisms [62], suggesting the involvement of TAD2 and TD/RD in p53-mediated regulation of genomic stability [35]. Only a few 3-D structures of the proteins in complex with p53 via TAD2 and TD/RD have been solved to date, while other interactions are still unknown. Thus, the possible role of residues making up the patterns identified in this study has yet to be elucidated. Nevertheless, to get indications of the possible adaptive importance of the specific changes in these domains in Spalax p53, we applied these changes in silico in human p53 and showed, albeit indirectly, some evidence for decreased affinity between the modified p53 (according to the hypoxia-tolerant pattern) and RPA70 compared to human p53. This complies with a previous study [63] demonstrating that a few combinations of p53 mutations: [L43A, **M44A**, **L45A**], or [**D48H**, D49H] or [I50A, **E51A**, Q52A], have disrupted the binding of p53 to RPA70. Each of these combinations includes residues that participate in the pattern identified in the current study as stress-related in TAD2 (highlighted in bold). Other studies [64, 65] have demonstrated that the mutations **D48H** and D49H in p53-TAD2 preserved the transactivation function of p53, but disrupted p53 interaction with RPA70 and the subsequent inhibition of homologous recombination (HR) by p53 [65]. This finding suggests the involvement of these residues in the sequestration of RPA70, which is required for HR suppression, and therefore it may be concluded that an

Domankevich et al. BMC Evolutionary Biology (2016) 16:177

Page 9 of 15

inhibition of HR suppression (i.e., HR promotion) might occur in *Spalax*.

TAD2 of p53 functions as a single-stranded (ss) DNA mimic and interacts with proteins that also interact with ssDNA [58, 59]. This DNA mimicking ability may have a role in the dynamics of these proteins' translocation to the damage site. For example, RPA70 has higher affinity to ssDNA than to p53, and upon DNA damage, it detaches from p53 and integrates with ssDNA at the damage site [58]. Thus, the potential decrease in the binding affinity of p53 to RPA70 in *Spalax* may lead to faster recruitment of RPA70 in the case of DNA damage. Another protein that interacts with p53 via TAD2 and shows similar structural competition with ssDNA is breast cancer 2 protein (BRCA2), which is also involved in double-strand break repair and HR [66, 67]. There is evidence that residues in TAD2, including the above-mentioned D48 and E51 (human p53), play a role in the binding of p53 to BRCA2 [67], implying a similar role in the recruitment of BRCA2 to the damage site. In this interaction, it was suggested that charged residues, such as D48 and E51 (corresponding to E49 and A52 in *Spalax* p53) attract the partner domains of BRCA2 to p53 at the initial phase of binding, whereas hydrophobic residues, such as L43 and L45 (human p53), which is also included in the identified patterns, are associated with the folding of the participating activation domain in BRCA2.

P53 transactivation is regulated by p300, which also affects p53 turnover, depending on the cellular context and the environmental stimuli, e.g., those that induce DNA damage [68]. A recent study investigating the structure of p53-TAD2 in complex with p300 [61] indicates the involvement of almost all the residues identified in the TAD2-stress-related pattern in this interaction. As in the case of RPA70, mutations in positions D48 and E51 (human p53) were shown to induce a decrease in the binding affinity of p53 to p300 [61]. This fits our modeling of the p53-p300 interaction via TAD2 indicating a potentially weaker binding in *Spalax* relative to humans. Importantly, under hypoxia, p53 crosstalks with the hypoxic response transcription factor hypoxia inducible factor 1 (HIF-1) and both compete for p300 as a co-activator to enhance transcription of their target genes [69]. This competition may suggest that the possible decrease in the binding affinity of *Spalax* p53 to p300 indirectly increases p300 availability to HIF-1 and thus enhances HIF-1-mediated transcription (hypoxic response) on the account of p53-mediated-transcription (apoptotic response). This could serve *Spalax* as a complementary mechanism to its constitutively high level of HIF-1 expression [70], as the upregulation of HIF-1 alone is not sufficient to enhance its target-genes transcription, and a compatible amount of available p300 is needed. The potential enhancement of HIF-1-mediated transcription on the account of p53 transcription may shift the balance in *Spalax* cells between survival and apoptosis, in favor of survival.

Most interesting is to examine proteins that interact with both TAD2 and TD/RD of p53, as such interactions can point to a functional link between these two domains. A functional link between TAD2 and RD of p53 was previously demonstrated in the case of the homeodomain-interacting protein kinase-2 (HIPK2), which its knockdown inhibited S46 phosphorylation in TAD2 and K382 acetylation in RD; both S46 phosphorylation and K382 acetylation are required for p53–mediated apoptosis [71]. CREB-binding protein (CBP) belong to the same protein family [72] as p300 and share a similar structure and function as co-activators. Both interact with p53 via TAD2 and acetylate the same residue (K382, human p53-RD) [73, 74]. This residue is also a substrate for the NAD-dependent protein deacetylase SIRT1 (human sirtuin) [75]. Hence, CBP/p300 and SIRT1 may compete for binding p53 in this location in RD. It was shown that the adjacent L383 (human p53), which is included in the stress-related patterns identified in this study, participates in p53 interaction with CBP and SIRT1 via RD, which is required for the acetylation/deacetylation of K382 in p53 (Table 2, Fig. 4). Acetylation of p53 by p300 positively regulates p53 activity, while deacetylation by SIRT1 negatively regulates p53 activity, suppresses apoptosis, and prolongs cellular survival in response to DNA damage [75]. It can be suggested that the patterns we found in p53, may promote p53 deacetylation by SIRT1 on the account of p53 acetylation by p300/CBP, due to the potential decrease in p300/CBP binding to p53 (via TAD2) that may indirectly increase the availability of K382 in RD to SIRT1. SIRT1 homologs have been shown to induce slow-aging in *Caenorhabditis elegans*, *Drosophila*, and mice [76]. It was suggested that this effect is achieved by delaying apoptosis via p53 and giving cells more time to repair the damage [77]. We see the similarities in p53 protein sequence found between *Spalax* and other stress-tolerant and phylogenetically distant mammals as a gateway for investigating potential evolutionary-convergent interconnected mechanisms that may relate its tolerance to extreme changes in oxygen levels to longevity and cancer resistance.

Conclusions

In this study we showed that the p53 domains TAD2 and TD/RD are more similar between species adapted to hypoxic stress than expected from phylogeny. Patterns shared in these domains among hypoxia-tolerant species include residues that mediate the binding of p53 with critical DDR proteins. Part of the residues are located in positions that were identified as positively selected and potentially play a role in reducing the binding of p53

Domankevich et al. BMC Evolutionary Biology (2016) 16:177

Page 10 of 15

with p300 and RPA70, suggesting a possible change in p53-mediated regulation of the DDR among species adapted to extreme changes in oxygen levels. Further investigation of TAD2 and TD/RD in the context of hypoxia and DNA repair may lead to progress in understanding cellular hypoxia tolerance, cancer resistance, and aging.

Methods

Data set and experimental design

In the current study, p53 sequences of 66 placental mammals were used (Additional file 1: Table S1, Additional file 6: Figure S3). From this group of 66 species, 47 species (Additional file 2: Figure S1) were classified into 3 groups: (1) "hypoxic-stress" group, which includes 15 species adapted to acute and transient environmental hypoxia (e.g., subterranean and diving mammals). Species in this group acquired the ability to inhabit environments that are characterized by extreme and frequent oxygen fluctuations and oxygen deprivation. In this group, we included all known-to-date mammals that, to the best of our knowledge, fit this criterion and have the full length of the p53 protein sequence published. The information regarding mammalian adaptations to different types of hypoxic stresses is available in a review by Ramirez et al. [78], and in our previous study on *Spalax* [4]; (2) "non-stress" group, which includes 16 species that do not endure any type of hypoxic stress (organismal/external/chronic/acute) or other types of stresses affecting cellular metabolism and extreme changes in availability of oxygen to tissues and cells; (3) "metabolic stress" group, which includes 16 species that fit neither the acute-hypoxic nor the non-stress criteria, and are known to endure other types of metabolic stresses such as hibernation, dehydration, extreme temperatures, chronic hypoxia, flying, or different combinations of these stresses (Additional file 1: Table S1). Extra caution should be taken when interpreting the results of the statistical analysis related to this group due to its heterogeneity. The remaining other 19 mammals could not be classified into any of the groups due to insufficient or inconclusive information in the literature or because including them in one of the groups strongly disturb the balance between groups. These mammals were termed as "additional" group and their p53 sequences were mostly used in analyses for which prior classification is not required, such as positive selection.

Species classification into the three first groups specified above was based on the summarized literature regarding each species from the viewpoint of stress adaptations. The classification was optimized as explained below, so that stress adaptations related to p53 could be investigated. For example, some mammals were classified as "non-stress" despite the fact they have semi-aquatic family members (e.g., fishing cat and mink are relatives of cat and ferret, respectively) or evidence of semi-aquatic adaptations (e.g.,

rhinoceros and humans) [79]. In addition, in this experimental design, the ability to obtain balanced proportions of species from different taxonomic orders and families in each group was also limited by the number of species for which the p53 sequence had been published and by the actual adaptations of the species. Thus, for example, in the metabolic-stress group, there are more rodents than in the other groups. Due to the mentioned limitations, an additional experimental design was used in which each hypoxia-tolerant species was paired with its closest hypoxia-sensitive phylogenetic relative that has an available p53 sequence (see Fig. 1a). Specifically, we selected all pairs of species so that in each pair: (1) there is a unique last common ancestor, namely: not the common ancestor of any other pair of species in the test; (2) one species is hypoxia-tolerant and the other is hypoxia-sensitive. It should be noted that two species from the "additional group" (shrew and elephant) were used in the balanced experimental design to represent hypoxia-sensitive species. This is because these species are the most closely related species (for which p53 sequence was published) to the stare-nosed-mole and manatee, respectively, and despite some stress-related inconclusive data (for example, it is not clear whether shrew faces metabolic stress due to its small body mass).

Protein and mRNA sequences were retrieved from NCBI protein and nucleotide databases, respectively [http://www.ncbi.nlm.nih.gov/]. Habitat and behavioral information were obtained from the Animal Diversity Web (ADW) online database [80]. Taxonomy information was collected from the NCBI Taxonomy browser [81]. Maximal lifespan and adult weight information were obtained from AnAge database [12]. Accession numbers, taxonomy information, lifespan, weight, and additional information related to stress adaptations are available in Additional file 1: Table S1.

Analysis of the p53 protein sequence and its functional domains

Forty-seven p53 sequences of species included in the hypoxic-, metabolic- and non-stress groups were aligned using MAFFT (multiple alignment using fast Fourier transform) L-INS-i (v7), which applies an iterative refinement method for local pairwise alignment, based on the BLOSUM62 model [82]. To obtain p53 fragments for each functional domain in p53, the alignment result was trimmed using AliView alignment editor [83] to the following domains: TAD1, TAD2, PRD, DBD, NLS, and TD/RD, according to relative positions in human p53 (Fig. 1b). Gaps were removed from the sub-alignments, and the resulting fragments were used for further analysis as described below. The fragments of each p53 functional domain were aligned using MAFFT (as described above). These alignments were

Domankevich et al. BMC Evolutionary Biology (2016) 16:177

Page 11 of 15

used to create a distance matrix according to a Poisson correction model [84], using MEGA6 [85].

Domain specific intra-group similarity

In order to test whether pairwise distances within a group involving hypoxia-tolerant species and a matched group of hypoxia-sensitive species are similar, the following test was performed. We characterized domain specific intra-group similarity for group G of k taxa by a vector S_G of $n = k(k-1)/2$ distances, scored based on the selected domain for all pairs of taxa from this group. We then tested the following hypothesis H_0: for the selected domain, components of vectors S_{HT} and S_{HS} have the same intra-group joint distribution. Testing this hypothesis by a standard one-tailed Mann–Whitney U test is problematic as the components of each of these two vectors are not independent. For example, information on d_{ij} and d_{jk} provides some information on d_{ik}. To overcome this difficulty, we extended the standard U test by taking into account this type of dependency in row data as described below. Analogously to a standard U test, we used test statistic U equal to the number of pairs (out of all $n_{HT}*n_{HS}$) of components of vectors S_{HT} and S_{HS} such that $S_{HT} > S_{HS}$. To conduct the test on real data, we estimate the distribution of U values under hypothesis H_0 by the following Monte-Carlo [86] simulations: In each of $N = 100,000$ runs we simulated $k_{HT} + k_{HS}$ random points in Euclidean space. Coordinates of these points were simulated as realization of independent random values with uniform distribution from 0 to 1. For groups of the first k_{HT} and of the next k_{HS} points, we calculated all n_{HT} and n_{HS} intra-group (dependent) distances as well as U-value. We further estimate p-value for the observed real vectors S_{HT} and S_{HS} as the proportion of simulated runs (out of $N = 100,000$) with $U-EU \geq U_{obs}-EU$, where $EU = (n_{HT}*n_{HS})/2$ is an expectation of statistic U under hypothesis H_0. If this p-value is less than some selected cutoff (e.g., 0.05) then the hypothesis H_0 should be rejected.

Comparing Spalax p53 domains to the domains of other species

Spalax p53 domains were compared to the p53 domains of a range of species: (1) species included in the entire NCBI non-redundant (*nr*) database (2) selected 46 species, as described below. We used two types of similarity measures (i) *Spalax* p53 domains were used as Blastp queries against the *nr* database, which includes most of the non-redundant protein sequences. Blastp hits were then sorted by E-value; (ii) The distances of p53 domains belonging to selected 46 species, from *Spalax* were calculated using Poisson correction model. Fisher's exact was used to test whether the domain-distance categories (namely, 'close' and 'distant' to *Spalax*) are significantly associated with the stress-tolerance categories. The

categories 'distant' and 'close' were measured by: (a) expected distance by phylogeny; (b) observed domain sequence distance. An 'expected' 2×2 contingency table was built with the counts of 'hypoxia-tolerant'/'hypoxia-sensitive' species that are 'close'/'distant' from *Spalax*, where 'close' species were classified as those belonging to the *Rodentia* or *Lagomorpha* clades (with less than 87 million years divergence time from *Spalax*). In this test, 'hypoxia-tolerant' refers to 15 species included in the hypoxic-stress group and 'hypoxia-sensitive' refers to 32 species included in the metabolic- and non-stress groups (see Additional file 1: Table S1). The 'observed' 2×2 contingency table included the same total counts of 'distant'/'close' species as in the 'expected' table, but the distance measure was domain distance rather than phylogenetic distance. Fisher's exact test was conducted using the R function 'fisher.test'. The data set tested included a higher proportion of species categorized as phylogenetically "distant" than phylogenetically "close" among the HT group (15 species), and a lower proportion for the HS group (32 species). Thus, Fisher's exact test odds ratio (OR) parameter was adjusted and defined as $(H_{close}/H_{dist})/(S_{close}/S_{dist})$; where: H_{close} and H_{dist} are the counts of HT species that are either closely-related or distantly-related to *Spalax*, respectively; S_{close} and S_{dist} are the counts of HS species that are either closely related or distantly related to *Spalax*, respectively. The parameter '*or*' (odds ratio) of the R function 'fisher.test' was calculated from the 'expected' table, to reflect the expected proportions of HT species in the 'close' vs. 'distant' group (see Additional file 3: Table S2, TAD2 example). In addition, the test was repeated for a subset of 43 species and of 41 representative species (for groups of species separated by <10 million years, only a single species was selected, and the rest were excluded). Furthermore, the test was repeated 100 times with random sub-samples of 38 and 33 species, out of the 43 species (jackknife resampling method) and the mean p-value was calculated. In the above-described test we used a data set that included all known hypoxia-tolerant species for which p53 sequence was published. This comprises species within clades that consist of only hypoxia-tolerant species (Additional file 2: Figure S1), and thus it may introduce potential biases due to dependent samples. Therefore, the results were tested also in a balanced data set (as explained in the previous Methods sections).

Classification analysis

For classification analysis, we used the data-mining method "Boosting Trees for Classification" (http://www.statsoft.com/Textbook/Boosting-Trees-Regression-Classification) [87]. This method predicts the membership of objects in different groups based on the values of one or more predictor variables. The p53 protein sequences of 66 species (see Additional file 1: Table S1) were

Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 12 of 15

aligned using CLUSTALW2 Omega (1.2.1) [88], where each alignment position was considered as a potential predictor variable in the analysis. The actual predictors were set to the positions that were found to be more frequent in species adapted to hypoxic/metabolic stress (see Results, Fig. 2a). The species' classification into three categories (hypoxic-, metabolic-, and non-stress) was set to be the dependent variable, and the algorithm classified the species based on the defined predictors. The "Training" data was set to be the 47 species that we subjectively assigned to the "hypoxic-stress", "metabolic-stress", and "non-stress" groups. This classification was determined according to information we collected from the literature related to the ability of the species to endure different types of hypoxic/metabolic stresses (see Additional file 1: Table S1). Thus, these species were used by the boosted-tree-algorithm to build the model, and their subjective classification was set to be the initial observed values. The remaining 19 species ("Additional" group) comprised the "Test" data that was used to evaluate the fitting of the model over successive iterations and did not participate in the building of the model. The initial observed values of these species were tentatively set to the intermediate stress type, namely: "metabolic stress".

Structural analysis of *Spalax* p53

PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC was used for molecular visualization of proteins' interaction with p53. For structures involving p53 TAD2 (Additional file 9: Figure S5), all residues that differ between human and *Spalax* p53 sequences were marked and displayed in balls. For visualizing 'Selecton' positive-selection results (Fig. 4), the residues in TAD2 and RD for which dN/dS >1 were marked by the Selecton software in the relevant PDB file and then displayed by PyMOL, accordingly. For modeling the structure of *Spalax* p53 in complex with RPA70 according to the human complex, only the residues participating in the direct interface between p53 and RPA70 were replaced by *Spalax* residues and displayed. For modeling the structure of *Spalax* p53 in complex with p300, only residues that differ between *Spalax* and humans and participate in the stress-related patterns were replaced and displayed. PDB files were retrieved from the PDB database [89]. MutaBind Server [90] was used to estimate the putative changes in the binding affinity of *Spalax* p53 to other DDR proteins, compared to other species.

Site-specific positive-selection analysis

A phylogenetic tree of 65 species was built according to the phylogenetic tree topologies in TimeTree of Life [16], a resource for estimations of divergence-time between species. Then, the tree-branch lengths

were optimized using Rate4Site [91] according to the p53 protein sequences of the 65 species, aligned by MAFFT (Additional file 11: Table S6). Since site-specific positive-selection analysis tools are extremely sensitive to alignment errors, when aligning p53 codon sequences, we manually inspected, using Ali-View alignment editor [83], the results of two codon-alignment methods: (1) PAGAN v.0.47 [92] and (2) the method implemented by the Selecton server [93], which is based on MUSCLE multiple alignment [94]. Due to an ambiguous result for amino acid residues in the disordered TAD2 when using PAGAN alignment, the aligned codon sequences produced by Selecton were chosen for use as input data in the analytical procedures described below, in addition to the phylogenetic tree of 65 species for which the branches were optimized.

Positive-selection analysis was conducted using: (1) The HyPhy module MEME [95]. This method reports site-specific dN/dS (i.e., non-synonymous vs. synonymous substitution rate), allowing the estimated dN/dS rate to change between both sites and branches, as expected in the case of p53, where it is likely that at different sites, only some branches were positively selected. For this analysis, following the HyPhy manual: SLAC analysis was run first, followed by MEME analysis, both analyses with HyPhy QuickSelectionDetection.bf module; (2) Maximum likelihood estimation of site-specific dN/dS using CODEML program in PAML version 4.8 [96, 97]. This program allows comparing the models M1a vs. M2a, and M7 vs. M8. The null-models M1a and M7 estimate the probabilities of 2 and 10 classes of dN/dS for each site, all with dN/dS between 0 and 1. The alternative models M2a and M8 include an additional site class with dN/dS >1 estimate. Using a likelihood ratio test (LRT), p-values were calculated from χ^2 distribution. If a specific alternative model was found to be significant (LRT p-value < 0.05), then positive selection was tested for each site (p > 0.95 for the dN/dS >1 site class). Site class probability was tested using Bayes Empirical Bayes (BEB) [98]. (3) The Selecton online server [93] [http://selecton.tau.ac.il/]. This server implements an empirical Bayesian method for dN/dS calculation and allows to compare the M8 model (that assumes selection) vs. the M8a null model (that assumes no selection) by LRT as described above, and to identify site-specific selection, accordingly. It is important to clarify that site-specific positive selection models do not provide information regarding the lineage in which the positive selection might occur. Methods that provide information regarding the site and the branch (e.g., branch-site model, PAML) were not used since these methods are not reliable for regions containing alignment gaps, which were tested here [99].

Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 13 of 15

Additional files

Availability of data and material
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' contributions
Wrote the paper: VD, AK, IS, AM. Analyzed the data: VD, AM, AK, ZF. Performed structural analyses: YO. Designed the experiments: AK, VD, AM, IS, AA, IM. Collected the data: VD. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Author details
[1]Institute of Evolution & Department of Evolutionary and Environmental Biology, University of Haifa, Haifa, Israel. [2]Mina and Everard Goodman Faculty of Life Sciences and Advanced Materials and Nanotechnology Institute, Bar-Ilan University, Ramat-Gan, Israel.

References
1. Nevo E, Ivanitskaya E, Beiles A. Adaptive radiation of blind subterranean mole rats: naming and revisiting the four sibling species of the Spalax ehrenbergi superspecies in Israel: Spalax galili (2n = 52), S. golani (2n = 54), S. carmeli (2n = 58) and S. judaei (2n = 60). Leiden: Backhuys Publishers; 2001.

2. Avivi A, Resnick MB, Nevo E, Joel A, Levy AP. Adaptive hypoxic tolerance in the subterranean mole rat Spalax ehrenbergi: the role of vascular endothelial growth factor. FEBS Lett. 1999;452(3):133–40.

3. Avivi A, Shams I, Joel A, Lache O, Levy AP, Nevo E. Increased blood vessel density provides the mole rat physiological tolerance to its hypoxic subterranean habitat. FASEB J. 2005;19(10):1314–6.

4. Shams I, Avivi A, Nevo E. Oxygen and carbon dioxide fluctuations in burrows of subterranean blind mole rats indicate tolerance to hypoxic-hypercapnic stresses. Comp Biochem Physiol A Mol Integr Physiol. 2005;142(3):376–82.

5. Lanoix D, Lacasse AA, Reiter RJ, Vaillancourt C. Melatonin: the watchdog of villous trophoblast homeostasis against hypoxia/reoxygenation-induced oxidative stress and apoptosis. Mol Cell Endocrinol. 2013;381(1-2):35–45.

6. Aguilera A, Garcia-Muse T. Causes of genome instability. Annu Rev Genet. 2013;47:1–32.

7. Bristow RG, Hill RP. Hypoxia and metabolism. Hypoxia, DNA repair and genetic instability. Nat Rev Cancer. 2008;8(3):180–92.

8. Dizdaroglu M. Oxidatively induced DNA damage: mechanisms, repair and disease. Cancer Lett. 2012;327(1-2):26–47.

9. Klein TJ, Glazer PM. The tumor microenvironment and DNA repair. Semin Radiat Oncol. 2010;20(4):282–7.

10. Pires IM, Bencokova Z, Milani M, Folkes LK, Li JL, Stratford MR, Harris AL, Hammond EM. Effects of acute versus chronic hypoxia on DNA damage responses and genomic instability. Cancer Res. 2010;70(3):925–35.

11. Maslov AY, Vijg J. Genome instability, cancer and aging. Biochim Biophys Acta. 2009;1790(10):963–9.

12. Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D, Costa J, Fraifeld VE, de Magalhaes JP. Human ageing genomic resources: integrated databases and tools for the biology and genetics of ageing. Nucleic Acids Res. 2013;41(Database issue):D1027–1033.

13. Manov I, Hirsh M, Iancu TC, Malik A, Sotnichenko N, Band M, Avivi A, Shams I. Pronounced cancer resistance in a subterranean rodent, the blind mole-rat, Spalax: in vivo and in vitro evidence. BMC Biol. 2013;11:91.

14. Buffenstein R. The naked mole-rat: a new long-living model for human aging research. J Gerontol Ser A Biol Med Sci. 2005;60(11):1369–77.

15. Delaney MA, Ward JM, Walsh TF, Chinnadurai SK, Kerns K, Kinsel MJ, Treuting PM. Initial case reports of cancer in naked mole-rats (heterocephalus glaber). Vet Pathol. 2016;53(3):691–6.

16. Hedges SB, Dudley J, Kumar S. TimeTree: a public knowledge-base of divergence times among organisms. Bioinformatics. 2006;22(23):2971–2.

17. Larson J, Park TJ. Extreme hypoxia tolerance of naked mole-rat brain. Neuroreport. 2009;20(18):1634–7.

18. de Magalhães JP. How ageing processes influence cancer. Nat Rev Cancer. 2013;13(5):357–65.

19. Wilhelm Filho D, Sell F, Ribeiro L, Ghislandi M, Carrasquedo F, Fraga CG, Wallauer JP, Simoes-Lopes PC, Uhart MM. Comparison between the antioxidant status of terrestrial and diving mammals. Comp Biochem Physiol A Mol Integr Physiol. 2002;133(3):885–92.

20. Carey HV, Frank CL, Seifert JP. Hibernation induces oxidative stress and activation of NK-kappaB in ground squirrel intestine. J Comp Physiol B. 2000;170(7):551–9.

21. Carey HV, Andrews MT, Martin SL. Mammalian hibernation: cellular and molecular responses to depressed metabolism and low temperature. Physiol Rev. 2003;83(4):1153–81.

22. Orr AL, Lohse LA, Drew KL, Hermes-Lima M. Physiological oxidative stress after arousal from hibernation in Arctic ground squirrel. Comp Biochem Physiol A Mol Integr Physiol. 2009;153(2):213–21.

23. Toien O, Drew KL, Chao ML, Rice ME. Ascorbate dynamics and oxygen consumption during arousal from hibernation in Arctic ground squirrels. Am J Physiol Regul Integr Comp Physiol. 2001;281(2):R572–583.

24. Merkt JR, Taylor CR. "Metabolic switch" for desert survival. Proc Natl Acad Sci U S A. 1994;91(25):12313–6.

25. el Hachimi Z, Tijane M, Boissonnet G, Benjouad A, Desmadril M, Yon JM. Regulation of the skeletal muscle metabolism during hibernation of Jaculus orientalis. Comp Biochem Physiol B. 1990;96(3):457–9.

26. Schulke S, Dreidax D, Malik A, Burmester T, Nevo E, Band M, Avivi A, Hankeln T. Living with stress: regulation of antioxidant defense genes in the subterranean, hypoxia-tolerant mole rat, Spalax. Gene. 2012;500(2):199–206.

27. Vazquez-Medina JP, Zenteno-Savin T, Elsner R, Ortiz RM. Coping with physiological oxidative stress: a review of antioxidant strategies in seals. J Comp Physiol B. 2012;182(6):741–50.

Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 14 of 15

28. Ali MA, Kazzam E, Amir N, Nyberg F, Adem A. Effects of dehydration and blockade of angiotensin II AT1 receptor on stress hormones and anti-oxidants in the one-humped camel. BMC Vet Res. 2013;9:232.

29. Seim I, Fang X, Xiong Z, Lobanov AV, Huang Z, Ma S, Feng Y, Turanov AA, Zhu Y, Lenz TL, et al. Genome analysis reveals insights into physiology and longevity of the Brandt's bat Myotis brandtii. Nat Commun. 2013;4:2212.

30. Maina JN. What it takes to fly: the structural and functional respiratory refinements in birds and bats. J Exp Biol. 2000;203(Pt 20):3045–64.

31. Salmon AB, Leonard S, Masamsetti V, Pierce A, Podlutsky AJ, Podlutskaya N, Richardson A, Austad SN, Chaudhuri AR. The long lifespan of two bat species is correlated with resistance to protein oxidation and enhanced protein homeostasis. FASEB J. 2009;23(7):2317–26.

32. Wilkinson GS, South JM. Life history, ecology and longevity in bats. Aging Cell. 2002;1(2):124–31.

33. Munshi-South J, Wilkinson GS. Bats and birds: Exceptional longevity despite high metabolic rates. Ageing Res Rev. 2010;9(1):12–9.

34. Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, et al. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science. 2013; 339(6118):456–60.

35. Reinhardt HC, Schumacher B. The p53 network: cellular and systemic DNA damage responses in aging and cancer. Trends Genet. 2012;28(3):128–36.

36. Polo SE, Jackson SP. Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. Genes Dev. 2011;25(5):409–33.

37. Lim YP, Lim TT, Chan YL, Song AC, Yeo BH, Vojtesek B, Coomber D, Rajagopal G, Lane D. The p53 knowledgebase: an integrated information resource for p53 research. Oncogene. 2007;26(11):1517–21.

38. Rodier F, Campisi J, Bhaumik D. Two faces of p53: aging and tumor suppression. Nucleic Acids Res. 2007;35(22):7475–84.

39. Avivi A, Ashur-Fabian O, Joel A, Trakhtenbrot L, Adamsky K, Goldstein I, Amariglio N, Rechavi G, Nevo E. P53 in blind subterranean mole rats–loss-of-function versus gain-of-function activities on newly cloned Spalax target genes. Oncogene. 2007;26(17):2507–12.

40. Band M, Ashur-Fabian O, Avivi A. The expression of p53-target genes in the hypoxia-tolerant subterranean mole-rat is hypoxia-dependent and similar to expression patterns in solid tumors. Cell Cycle. 2010;9(16):3347–52.

41. Uversky VN, Dave V, Iakoucheva LM, Malaney P, Metallo SJ, Pathak RR, Joerger AC. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. Chem Rev. 2014;114(13):6844–79.

42. Xue B, Brown CJ, Dunker AK, Uversky VN. Intrinsically disordered regions of p53 family are highly diversified in evolution. Biochim Biophys Acta. 2013; 1834(4):725–38.

43. Zou Y, Liu Y, Wu X, Shell SM. Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. J Cell Physiol. 2006;208(2):267–73.

44. Luo J, Nikolaev AY, Imai S, Chen D, Su F, Shiloh A, Guarente L, Gu W. Negative control of p53 by Sir2alpha promotes cell survival under stress. Cell. 2001;107(2):137–48.

45. Zhou XY, Shibusawa N, Naik K, Porras D, Temple K, Ou H, Kaihara K, Roe MW, Brady MJ, Wondisford FE. Insulin regulation of hepatic gluconeogenesis through phosphorylation of CREB-binding protein. Nat Med. 2004;10(6):633–7.

46. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics. 1947;18(1):50–60.

47. Mas C, Lussier-Price M, Soni S, Morse T, Arseneault G, Di Lello P, Lafrance-Vanasse J, Bieker JJ, Omichinski JG. Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF). Proc Natl Acad Sci U S A. 2011;108(26):10484–9.

48. Follis AV, Llambi F, Ou L, Baran K, Green DR, Kriwacki RW. The DNA-binding domain mediates both nuclear and cytosolic functions of p53. Nat Struct Mol Biol. 2014;21(6):535–43.

49. Walker DR, Bond JP, Tarone RE, Harris CC, Makalowski W, Boguski MS, Greenblatt MS. Evolutionary conservation and somatic mutation hotspot maps of p53: correlation with p53 protein structural and functional features. Oncogene. 1999;18(1):211–8.

50. Abegglen LM, Caulin AF, Chan A, Lee K, Robinson R, Campbell MS, Kiso WK, Schmitt DL, Waddell PJ, Bhaskara S, et al. Potential Mechanisms for Cancer Resistance in Elephants and Comparative Cellular Response to DNA Damage in Humans. JAMA. 2015;314(17):1850–60.

51. Jenkins LM, Yamaguchi H, Hayashi R, Cherry S, Tropea JE, Miller M, Wlodawer A, Appella E, Mazur SJ. Two distinct motifs within the p53 transactivation domain bind to the Taz2 domain of p300 and are differentially affected by phosphorylation. Biochemistry. 2009;48(6):1244–55.

52. Shan B, Li DW, Bruschweiler-Li L, Bruschweiler R. Competitive binding between dynamic p53 transactivation subdomains to human MDM2 protein: implications for regulating the p53.MDM2/MDMX interaction. J Biol Chem. 2012;287(36):30376–84.

53. Edwards RJ, Palopoli N. Computational prediction of short linear motifs from protein sequences. Methods Mol Biol. 2015;1268:89–141.

54. Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ, Davey NE. Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. Chem Rev. 2014;114(13):6733–78.

55. Dinkel H, Michael S, Weatheritt RJ, Davey NE, Van Roey K, Altenberg B, Toedt G, Uyar B, Seiler M, Budd A, et al. ELM–the database of eukaryotic linear motifs. Nucleic Acids Res. 2012;40(Database issue):D242–251.

56. Davey NE, Cyert MS, Moses AM. Short linear motifs - ex nihilo evolution of protein regulation. Cell Commun Signal. 2015;13:43.

57. Di Lello P, Jenkins LM, Jones TN, Nguyen BD, Hara T, Yamaguchi H, Dikeakos JD, Appella E, Legault P, Omichinski JG. Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. Mol Cell. 2006;22(6):731–40.

58. Bochkareva E, Kaustov L, Ayed A, Yi GS, Lu Y, Pineda-Lucena A, Liao JC, Okorokov AL, Milner J, Arrowsmith CH, et al. Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. Proc Natl Acad Sci U S A. 2005;102(43):15412–7.

59. Rowell JP, Simpson KL, Stott K, Watson M, Thomas JO. HMGB1-facilitated p53 DNA binding occurs via HMG-Box/p53 transactivation domain interaction, regulated by the acidic tail. Structure. 2012;20(12): 2014–24.

60. Okuda M, Nishimura Y. Extended string binding mode of the phosphorylated transactivation domain of tumor suppressor p53. J Am Chem Soc. 2014;136(40):14143–52.

61. Miller Jenkins LM, Feng H, Durell SR, Tagad HD, Mazur SJ, Tropea JE, Bai Y, Appella E. Characterization of the p300 Taz2-p53 TAD2 complex and comparison with the p300 Taz2-p53 TAD1 complex. Biochemistry. 2015;54(11):2001–10.

62. Rappaport N, Twik M, Nativ N, Stelzer G, Bahir I, Stein TI, Safran M, Lancet D. MalaCards: A Comprehensive Automatically-Mined Database of Human Diseases. Curr Protoc Bioinformatics. 2014;47:1 24. 21-21 24 19.

63. Abramova NA, Russell J, Botchan M, Li R. Interaction between replication protein A and p53 is disrupted after UV damage in a DNA repair-dependent manner. Proc Natl Acad Sci U S A. 1997;94(14):7186–91.

64. Leiter LM, Chen J, Marathe T, Tanaka M, Dutta A. Loss of transactivation and transrepression function, and not RPA binding, alters growth suppression by p53. Oncogene. 1996;12(12):2661–8.

65. Romanova LY, Willers H, Blagosklonny MV, Powell SN. The interaction of p53 with replication protein A mediates suppression of homologous recombination. Oncogene. 2004;23(56):9025–33.

66. Brandsma I, Gent DC. Pathway choice in DNA double strand break repair: observations of a balancing act. Genome integrity. 2012;3(1):9.

67. Rajagopalan S, Andreeva A, Rutherford TJ, Fersht AR. Mapping the physical and functional interactions between the tumor suppressors p53 and BRCA2. Proc Natl Acad Sci U S A. 2010;107(19):8587–92.

68. Grossman SR. p300/CBP/p53 interaction and regulation of the p53 response. Eur J Biochem. 2001;268(10):2773–8.

69. Sermeus A, Michiels C. Reciprocal influence of the p53 and the hypoxic pathways. Cell Death Dis. 2011;2:e164.

70. Shams I, Avivi A, Nevo E. Hypoxic stress tolerance of the blind subterranean mole rat: expression of erythropoietin and hypoxia-inducible factor 1 alpha. Proc Natl Acad Sci U S A. 2004;101(26):9698–703.

71. Puca R, Nardinocchi L, Sacchi A, Rechavi G, Givol D, D'Orazi G. HIPK2 modulates p53 activity towards pro-apoptotic transcription. Mol Cancer. 2009;8:85.

72. Teufel DP, Freund SM, Bycroft M, Fersht AR. Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. Proc Natl Acad Sci U S A. 2007;104(17):7009–14.

73. Sakaguchi K, Herrera JE, Saito S, Miki T, Bustin M, Vassilev A, Anderson CW, Appella E. DNA damage activates p53 through a phosphorylation-acetylation cascade. Genes Dev. 1998;12(18):2831–41.

Domankevich *et al. BMC Evolutionary Biology* (2016) 16:177

Page 15 of 15

74. Liu L, Scolnick DM, Trievel RC, Zhang HB, Marmorstein R, Halazonetis TD, Berger SL. p53 sites acetylated in vitro by PCAF and p300 are acetylated in vivo in response to DNA damage. Mol Cell Biol. 1999;19(2):1202–9.

75. North BJ, Verdin E. Sirtuins: Sir2-related NAD-dependent protein deacetylases. Genome Biol. 2004;5(5):224.

76. Guarente L. Introduction: sirtuins in aging and diseases. Methods Mol Biol. 2013;1077:3–10.

77. Howitz KT, Bitterman KJ, Cohen HY, Lamming DW, Lavu S, Wood JG, Zipkin RE, Chung P, Kisielewski A, Zhang LL, et al. Small molecule activators of sirtuins extend Saccharomyces cerevisiae lifespan. Nature. 2003;425(6954):191–6.

78. Ramirez JM, Folkow LP, Blix AS. Hypoxia tolerance in mammals and birds: from the wilderness to the clinic. Annu Rev Physiol. 2007;69:113–43.

79. Williams MF. Morphological evidence of marine adaptations in human kidneys. Med Hypotheses. 2006;66(2):247–57.

80. Myers P, Espinosa R, Parr CS, Jones T, Hammond GS, Dewey TA. The Animal Diversity Web (online). 2015. Accessed at http://animaldiversity.org.

81. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2009;37(Database issue):D5–15.

82. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.

83. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276–8.

84. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. Evolving genes and proteins. 1965;97:97–166.

85. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9.

86. Fishman G. Monte Carlo: concepts, algorithms, and applications. New York: Springer; 1996.

87. Friedman J, Hastie T, Tibshirani R. The elements of statistical learning, vol. 1. Berlin: Springer series in statistics Springer; 2001.

88. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.

89. Bernstein FC, Koetzle TF, Williams GJ, Meyer Jr EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol. 1977;112(3):535–42.

90. Li M, Simonetti FL, Goncearenco A, Panchenko AR. MutaBind estimates and interprets the effects of sequence variants on protein-protein interactions. Nucleic Acids Res. 2016;44(W1):W494–501.

91. Pupko T, Bell RE, Mayrose I, Glaser F, Ben-Tal N. Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics. 2002;18 Suppl 1:S71–77.

92. Loytynoja A, Vilella AJ, Goldman N. Accurate extension of multiple sequence alignments using a phylogeny-aware graph algorithm. Bioinformatics. 2012;28(13):1684–91.

93. Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. Nucleic Acids Res. 2007;35(Web Server issue):W506–511.

94. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics. 2004;5:113.

95. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 2012;8(7):e1002764.

96. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24(8):1586–91.

97. Yang Z, Nielsen R. Mutation-selection models of codon substitution and their use to estimate selective strengths on codon usage. Mol Biol Evol. 2008;25(3):568–79.

98. Yang Z, Wong WS, Nielsen R. Bayes empirical bayes inference of amino acid sites under positive selection. Mol Biol Evol. 2005;22(4):1107–18.

99. Fletcher W, Yang Z. The effect of insertions, deletions, and alignment errors on the branch-site test of positive selection. Mol Biol Evol. 2010;27(10):2257–67.