Gene Reports 17 (2019) 100508

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Genome distance between conserved elements in neighborhoods of growth-regulating genes is correlated with morpho-physiological traits in mammals

Dmitriy Romanov*, Elena Butenko, Georgy Bakhtadze, Tatiana Shkurat

Department of Genetics, Academy of Biology and Biotechnologies, Southern Federal University, Stachki av. 194/1, Rostov-on-Don 344090, Russia

ARTICLE INFO

Keywords:
Growth-regulating genes
Conserved elements
Genome distance
Morpho-physiological traits
Mammals

ABSTRACT

The problem of growth regulation in mammals is one of the long-standing mysteries in biology. The growth rate is high in early postnatal life, but then decreases and eventually ceases in multiple tissues, determining adult body size, in all mammalian species. It has been shown that age-related changes in growth rate in mammals are associated with declining in expression of large set of growth-promoting genes. On the other hand, the somatotropic axis genes are known to be involved in the systemic growth regulation.

We have found very strong correlation between some morpho-physiological traits of mammals and genome distances between conserved elements in neighborhoods of these genes among mammals. We believe that genome distance between some regulatory sites of the gene may determine gene expression and eventually affect the phenotype. We first suggest that the difference of these distances among mammals can cause evolutionary variation of morpho-physiological traits. We have also proposed the particular molecular mechanisms of regulation of these genes.

1. Introduction

Body growth in animals is rapid in early life but then progressively slows down, imposing a limit on adult body size (Lui and Baron, 2011). Moreover, other morpho-physiological traits, such as lifespan or age of maturity, are known to be also evolutionary conserved among species: each species has its own value for each trait.

Growth deceleration in mammals is mostly governed by local rather than systemic mechanisms. Whole-genome gene expression assays in different tissues of various mammals showed that growth deceleration is closely related with downregulation of the following 10 genes: *Ezh2*, *Gpc3*, *Mdk*, *Mest*, *Mycn*, *Peg3*, *Plagl1*, *Smo*, *Igf2* and *E2f3* (Delaney et al., 2014; Lui et al., 2010a, 2008, 2010b).

It is still an issue what mechanisms coordinate downregulation of these genes and what are the factors of the evolutionary modulation of the pace of respective genetic program among mammals (Delaney et al., 2014; Finkielstain et al., 2009; Lui and Baron, 2011). The difference in age-related downregulation in various mammals may be due to different arrangement of *cis*-regulatory elements in the neighborhoods of these genes.

On the other hand, the somatotropic axis genes, which are mainly represented by 7 genes: *Gh1*, *Ghrh*, *Ghrl*, *Sst*, *Igfbp3*, *Igfbp1* and *Igf1*, are known to be involved in the systemic growth regulation in mammals (Lui et al., 2015). Previously, we have found strong correlation between lifespan, age of maturity and gene-telomere distances for the *Ghrh* and *Sst* genes (Romanov et al., 2019). We have also shown that some of the morpho-physiological traits in mammals correlate with the size of the non-coding regions of some of these genes (Shkurat et al., 2015).

Many *cis*-regulatory elements being conserved elements as well (Liu et al., 2008), the aim of this research was to study the relationships between morpho-physiological traits of mammals and genome distances between the conserved elements in the neighborhoods of the aforementioned 17 growth-regulating genes.

2. Materials and methods

We studied 36 species of mammals, listed in the Table 1. The adult body mass, adult body length, age of maturity and lifespan of animals were assessed. The trait values were obtained from the PanTHERIA database (Jones et al., 2009).

The next 17 genes were chosen for the study: 10 growth-promoting genes (*Ezh2*, *Gpc3*, *Mdk*, *Mest*, *Mycn*, *Peg3*, *Plagl1*, *Smo*, *Igf2* and *E2f3*) and 7 somatotropic axis genes (*Gh1*, *Ghrh*, *Ghrl*, *Igf1*, *Sst*, *Igfbp3* and *Igfbp1*).

abbreviations: b.p., base pair; DHS, DNase I hypersensitive sites; TFBS, Transcription factor binding site; lncRNA, Long non-coding RNA; miRNA, micro RNA; TPE-OLD, Telomere position effect over long distances; ICR, imprinting control region

* Corresponding author.
E-mail addresses: rdme@yandex.ru (D. Romanov), krupskaya_nk@mail.ru (E. Butenko), g-bakhtadze@yandex.ru (G. Bakhtadze), tshkurat@yandex.ru (T. Shkurat).

https://doi.org/10.1016/j.genrep.2019.100508
Received 27 June 2019; Received in revised form 9 September 2019; Accepted 12 September 2019
Available online 30 October 2019

Table 1
Morpho-physiological traits of 36 studied mammals. Values for traits were obtained from PanTHERIA database.

Species	Adult body mass (g)	Adult body length (mm)	Age of maturity (days)	Lifespan (months)
Mus musculus	19.3	100	60	24
Sorex araneus	9.18	72	290	20
Cricetulus griseus	37	150	40	28
Microtus ochrogaster	42.5	121	34	18
Tupaia chinensis	180	180	90	120
Rattus norvegicus	282.89	200	50	30
Callithrix jacchus	290.21	210	430	150
Pteropus alecto	610.13	240	540	235
Otolemur garnettii	811.17	250	590	200
Oryctolagus cuniculus	1590.57	400	180	145
Felis catus	2884.8	540	280	168
Macaca mulatta	6455.19	550	1100	204
Macaca fascicularis	4569.32	520	1200	216
Chlorocebus sabaeus	4300	560	1700	240
Nomascus leucogenys	7320	600	2800	336
Canis lupus familiaris	31,756.51	900	600	220
Papio anubis	17,728.56	700	3000	310
Capra hircus	47,386.47	1000	450	190
Pan paniscus	35,119.95	750	5000	490
Acinonyx jubatus	50,577.92	1420	680	225
Pan troglodytes	45,000	1500	3600	630
Homo sapiens	58,540.63	1750	5300	840
Ovis aries	39,097.89	1300	750	160
Sus scrofa	84,471.54	1350	330	175
Pongo abelii	39,696.12	1400	5500	360
Lipotes vexillifer	112,138.32	1700	3100	300
Gorilla gorilla	112,588.99	1600	3050	500
Panthera tigris altaica	161,914.66	1830	1460	215
Ursus maritimus	371,703.81	2000	1720	320
Equus caballus	403,598.53	2050	760	330
Tursiops truncatus	281,040.55	3000	3200	250
Bos taurus	618,642.42	2500	700	180
Loxodonta africana	3,824,539.93	5000	4600	800
Balaenoptera acutorostrata	5,587,093.59	5500	2500	510
Orcinus orca	5,628,758.66	8600	5700	600
Physeter catodon	14,540,959.82	12,000	3700	850

The information on each gene across the selected mammals was taken by full-text search in the NCBI Gene database. Sequences of neighborhoods of the genes studied were obtained in the GenBank format from the NCBI Nucleotide database with the help of the E-utilities (Kans, 2013). Gene neighborhood was defined as a gene locus flanked by 50,000 b.p.

Dot plot alignments of each gene neighborhood with the corresponding human gene neighborhood were performed. Sequence elements, preserved in each gene neighborhood of the gene, were considered to be conserved. Each conserved element was assigned a unique name, for instance the MYCN(-6893), where the MYCN denotes the name of the gene and the number denotes the distance from the start of the gene to the center of the element in *human*.

Dot plot alignments were generated and processed with the Java developed *dotolog* program.

Correlation analysis of the relationships between morpho-physiological traits of mammals and genome distances between the conserved elements in the neighborhoods of the genes studied was carried out. The relationships with the Spearman's correlation coefficient greater than 0.7 and the p-value less than 0.05 were considered to be statistically significant. The problem of multiple comparisons was addressed by multiplying the resulting p-value by the number of comparisons according to the Bonferroni correction.

Correlation analysis was performed by means of implementation of the Java subroutines with the help of the Java library Apache Commons Math (The Apache Software Foundation, 2006).

Search of the known genomic elements in the neighborhoods of the genes studied was performed with the developed BLAST-based searching pipeline *mblast*. Databases of known genomic elements are given in Table 2. A list of overlapping genomic elements for each conserved element was obtained.

3. Results

We examined whether the inter-conserved element distances in the neighborhoods of 17 growth-regulating genes correlated with morpho-physiological traits in mammals.

Totally we performed 44,336 comparisons between the inter-conserved element distances and morpho-physiological traits of mammals. To counteract the problem of multiple comparisons the Bonferroni correction was made with the multiplier that equaled 50,000.

The genome distances between some conserved elements in the neighborhoods of the *Mycn* and *Plagl1* genes were found to correlate significantly with the adult body mass and length of mammals and some inter-conserved element distances in the neighborhood of the *Ezh2* gene were found to correlate significantly with lifespan of mammals (Tables 3 and 4).

Negative correlation was observed for the inter-conserved element distances in the neighborhoods of the *Mycn* and *Ezh2* genes; the only pair of significantly correlated conserved elements in the neighborhood of the *Plagl1* gene showed positive correlation.

As expected, correlations were almost the same for the adult body mass and adult body length since these traits are known to be tightly correlated with each other (Jones et al., 2009).

Age of maturity did not significantly correlate with any inter-conserved element distance in the neighborhoods of genes studied.

Strikingly, for both the *Mycn* and *Ezh2* gene the pair of the topmost

Table 2
Databases of known study. The sequences of the genomic elements were downloaded from the corresponding database. If only genome coordinates of the elements were available in the database, the corresponding sequences were obtained with the help of *bedtools* software suit (Quinlan, 2014).

Database	Description	URL
EPD	The Eukaryotic Promoter Database	http://epd.vital-it.ch/
JASPAR	The high-quality transcription factor binding profile database	http://jaspar.genereg.net/
ORegAnno	A community-driven resource for curated regulatory annotation	http://www.oreganno.org/
LNCipedia	The long non-coding RNA database	http://lncipedia.org/
miRBase	The microRNA database	http://www.mirbase.org/
Pseudogene	A database of pseudogenes	http://www.pseudogene.org/
dbSUPER	A database of super-enhancers	http://bioinfo.au.tsinghua.edu.cn/dbsuper/
DENdb	A repository of predicted enhancers	http://www.cbrc.kaust.edu.sa/dendb/
VISTA	A database of tissue-specific human enhancers	https://enhancer.lbl.gov/
ENCODE	Transcription Factor ChIP-seq from ENCODE with Factorbook Motifs	http://genome.ucsc.edu/
NCBI dbSNP	NCBI database for nucleotide variations	https://www.ncbi.nlm.nih.gov/SNP/
NCBI dbVar	NCBI database of large scale genomic variants	https://www.ncbi.nlm.nih.gov/dbvar/
NCBI BLAST	NCBI BLAST databases	ftp://ftp.ncbi.nih.gov/blast/db/FASTA/

Table 3
Correlation between the inter-conserved element distance and morpho-physiological traits of mammals. Elements, which overlap gene promoters, are highlighted in bold.

Corr. coeff.	P-value	Bonferroni corrected p-value	Element 1	Element 2	Trait
−0.81	1.45×10^{-8}	0.0007	MYCN(−6893)	**MYCN(74)**	Adult
−0.77	1.32×10^{-7}	0.007	MYCN(−6893)	**MYCN(1273)**	body
−0.79	1.5×10^{-7}	0.008	MYCN(−6893)	MYCN(−1587)	mass
−0.78	1.58×10^{-7}	0.008	MYCN(−6893)	MYCN(3711)	
−0.75	6.51×10^{-7}	0.03	MYCN(−25368)	MYCN(−4493)	
−0.75	7.31×10^{-7}	0.04	MYCN(−25368)	**MYCN(74)**	
0.73	6.48×10^{-7}	0.03	**PLAGL1(79389)**	PLAGL1(112171)	
−0.8	4.71×10^{-8}	0.002	MYCN(−6893)	**MYCN(74)**	Adult
−0.77	2.12×10^{-7}	0.01	MYCN(−6893)	**MYCN(1273)**	body
−0.77	3.08×10^{-7}	0.02	MYCN(−6893)	MYCN(3711)	length
−0.77	3.79×10^{-7}	0.02	MYCN(−6893)	MYCN(−1587)	
−0.76	5.9×10^{-7}	0.03	MYCN(−25368)	MYCN(−4493)	
0.74	3.37×10^{-7}	0.02	**PLAGL1(79389)**	PLAGL1(112171)	
−0.79	4.07×10^{-7}	0.02	EZH2(−8314)	EZH2(2278)	Lifespan
−0.78	5.37×10^{-7}	0.03	EZH2(−8314)	**EZH2(194)**	
−0.78	5.99×10^{-7}	0.03	EZH2(−8314)	**EZH2(1627)**	

correlated conserved elements contained the element overlapped with the *promoter* of the same gene. For the *Plagl1* gene, such topmost pair did not contain element overlapped with the promoter of the gene, though it contained the element overlapped with the gene start site in some mammals (*Tupaia chinensis*, *Pan paniscus*, *Equus caballus*) and presumably contained the promoter sequences.

Semilog plots of the inter-element distance and morpho-physiological trait for the pairs of the topmost correlated conserved elements are given in Fig. 1. In each plot dots appear to line up along the straight line suggesting that morpho-physiological traits depend *exponentially* on the inter-element distance.

DNA sequences of the topmost correlated conserved elements are given in Table 5.

The list of the known genomic elements which overlapped significantly the correlated conserved elements is given in Table 6. None of the significantly correlated conserved elements overlapped with the Alu elements, lncRNAs, pseudogenes, VISTA enhancers, pre-miRNAs or mature miRNAs.

As expected, most of the gene start associated elements overlapped the DHSs (DNase I hypersensitivity sites) and contained many TFBSs (transcription factor binding sites).

The genomic elements, which overlap the topmost correlated conserved elements, may be involved in regulation of the corresponding gene. The next pairs of the conserved elements were considered (overlapping genomic elements are given in square brackets; conserved elements, overlapped gene promoters, are highlighted in bold):

- MYCN(-6893) [STAT1, GATA2] and **MYCN(74)**;
- **PLAGL1(79389)** [CDX1] and PLAGL1(112171) [CDX1, NFIC];
- EZH2(-8314) [E2F1] and **EZH2(194)**.

Remarkably, the GATA2 appears to induce growth via downstream activation of the *Mycn* (Nandakumar et al., 2015). Another study has demonstrated that the STAT1 might indirectly negatively regulate the *Mycn* (Hsu et al., 2016). There are little evidences in literature that either the CDX1 or the NFIC are involved in regulation of the *Plagl1*. It has been demonstrated that the E2F1 directly regulates the transcription of the *Ezh2* (Lee et al., 2015).

It is noteworthy that the genes studied appear to constitute some part of the growth-regulating network with only a few hubs. For example, the *Igf2* is known to be regulated by the *E2f3* (Lui and Baron, 2013). Hence, the expression of only a few master genes needs to be modulated in order to modulate the behavior of the whole network. Indeed, the *Mycn*, *Plagl1* and *Ezh2* genes may be considered as such master genes.

The genome variations between the significantly correlated conserved elements may also be associated with phenotype.

The list of the known variations between the topmost correlated conserved elements was obtained.

None of the intervals between the significantly correlated pairs of the conserved elements *contained* any clinically significant variation, though the list of clinically significant variations *overlapped* inter-conserved element intervals is given in Table 7.

Both clinically significant variations nsv997068 and nsv997222 affected the whole *Plagl1* and *Mycn* genes respectively and were considered as large gene rearrangements. The nsv510056 variation inside the *Plagl1* gene is associated with hydatidiform mole (MeSH D006828), which is known to be a severe developmental abnormality.

4. Discussion

4.1. Modulation of the growth-promoting gene expression

In early postnatal life, the extensive changes in gene expression occur concomitantly in multiple major organs, indicating the existence of a common core developmental genetic program. This program includes hundreds of growth-promoting genes that are downregulated with age in liver, kidney, lung, and heart, and there is evidence that this component of the program drives the widespread decline in cell proliferation that occurs in juvenile life, as organs approach adult sizes (Lui et al., 2014).

It is widely accepted that growth velocity curves of mammals are similar to each other (Bogin, 1999). On the other hand, temporal patterns of growth deceleration were shown to be similar to temporal patterns of declining gene expression (Lui et al., 2008). Thus, growth velocity appears to be *proportional* to the expression of the growth-promoting genes.

Obviously, the adult body mass is an integral of growth velocity over the lifespan. Operation of integration being a linear function of the integrand, any modulation of the growth-promoting gene expression will lead to conformable changes in the adult body mass and, conversely, interspecies differences in the adult body mass and length may point out on the type of modulation of the growth-promoting gene expression. It should be emphasized that adult body mass depends mainly on growth velocity and to a very little degree on lifespan, since maximal growth occurs before the age of maturity and almost ceases in later life.

Thus, the next assumptions are proposed:

- The growth in mammals is regulated by the same evolutionary conserved gene network.
- The same *cis*-elements control these gene in various mammals, each gene having its own set of *cis*-elements.
- Patterns of growth deceleration are similar in all mammals.
- Patterns of declining of gene expression are similar in all mammals.

Table 4
The inter-conserved element distances for the topmost correlated pairs of elements in the neighbourhood of the *Mycn*, *Plagl1* and *Ezh2* genes. Bonferroni corrected p-values × 50,000 are provided. Elements, which overlap gene promoters, are highlighted in bold.

Species	Inter-element distance (b.p.)		
	MYCN(−6893) – **MYCN(74)**	**PLAGL1(79389)** – PLAGL1(112171)	EZH2(−8314) – **EZH2(194)**
Mus musculus	9813	16,700	—
Sorex araneus	8653	20,565	—
Cricetulus griseus	9513	20,016	—
Microtus ochrogaster	9270	19,221	—
Tupaia chinensis	8813	34,230	—
Rattus norvegicus	9732	16,954	—
Callithrix jacchus	8686	35,168	10,860
Pteropus alecto	7772	29,476	7823
Otolemur garnettii	8269	28,891	11,345
Oryctolagus cuniculus	7160	30,031	—
Felis catus	5501	36,567	9551
Macaca mulatta	6956	30,704	9128
Macaca fascicularis	6939	30,693	9267
Chlorocebus sabaeus	8202	31,018	8559
Nomascus leucogenys	6999	33,339	8767
Canis lupus familiaris	—	37,555	9167
Papio anubis	6972	30,711	9329
Capra hircus	5600	34,969	11,560
Pan paniscus	7061	32,755	8524
Acinonyx jubatus	5356	36,016	9348
Pan troglodytes	6971	32,774	8519
Homo sapiens	6967	32,782	8508
Ovis aries	5601	34,013	11,277
Sus scrofa	6295	33,227	10,880
Pongo abelii	6958	—	8555
Lipotes vexillifer	6150	35,416	8493
Gorilla gorilla	6990	32,822	9607
Panthera tigris altaica	5410	36,125	9688
Ursus maritimus	5249	37,651	9319
Equus caballus	5342	35,752	9313
Tursiops truncatus	—	36,951	9023
Bos taurus	—	33,986	10,410
Loxodonta africana	6041	38,866	6780
Balaenoptera acutorostrata	5760	35,595	8381
Orcinus orca	5996	35,185	8297
Physeter catodon	—	36,767	8040
Correlation with adult body mass	$r_s = -0.81$, $p = 0.0007$	$r_s = 0.73$, $p = 0.03$	
Correlation with adult body height	$r_s = -0.8$, $p = 0.002$	$r_s = 0.74$, $p = 0.02$	
Correlation with lifespan			$r_s = -0.78$, $p = 0.03$

- Growth rate is proportional to the gene expression.
- Modulation of gene expression results in proportional change of the adult body mass and length.

In such assumptions, the difference in the adult body mass of mammals can be caused by the next reasons.

On the one hand, *local* environment of the gene, represented by its neighborhood with the *cis*-elements, can vary between species. In such case, the main difference will be the *placement* of the *cis*-elements with varying genome distances between these elements.

On the other hand, *global* environment of the gene can vary, that is gene location in the genome, particularly, location on the chromosome and gene-telomere distance. We have previously demonstrated that the gene-telomere distance for the *Ghrh* and *Sst* somatotropic axis genes, as well as the TPE-OLD (telomere position effect over long distances) regulated the *C1s* and *Notch1* genes and significantly correlated with the lifespan and age of maturity in mammals (Romanov et al., 2019).

Thus, the genome distance may be considered as a modulating factor of the growth-regulating gene expression that eventually determines phenotype.

In the present study, only local environment of the genes was examined.

4.2. Age-related declining of the growth-promoting gene expression

It has been suggested that epigenetic mechanisms can be involved in the declining of expression of growth-regulating genes (Lui and Baron, 2011; Lui et al., 2008, 2010b).

On the one hand, epigenetic modifications of the *promoter* can silence gene expression. Broad shifts in gene expression during early postnatal life were shown to be associated with shifts in the H3K4 trimethylation patterns of the promoter regions of many growth-regulating genes (Lui et al., 2014). It has been also hypothesized that decline in the expression is caused by altered DNA methylation of the promoter, since it is a widely known silencing mechanism, transferring throughout the cell division. However, contrary to this hypothesis, it has been shown that the methylation status of the promoter regions of the *Mest*, *Peg3*, and *Plagl1* genes does not change with age (Lui et al., 2008).

On the other hand, epigenetic modifications of the *distal regulatory regions* may be involved in gene silencing. It has been shown that DNA methylation of some distal regulatory sites may be more closely related to changes in gene expression as compared with promoter methylation (Aran and Hellman, 2013; Aran et al., 2013).

Remarkably, it has been recently shown that average DNA methylation level of certain sites in the genome (i) is *close to zero* for embryonic cells, (ii) shows *logarithmic* dependence on age until adulthood and slows to a *linear* dependence later in life and (iii) demonstrates similar behavior in *chimpanzee* tissues (Horvath, 2013). The link between age-associated changes to the methylome and functional changes in gene expression patterns has also been demonstrated (Hannum et al., 2013). It has been hypothesized that acquisition of promoter methylation with age increases the chance of methylation induced silencing at that promoter or other important regions such as enhancers (Christensen et al., 2009). Some promoters have been shown to change methylation progressively and linearly with age (Issa, 2014).

It should be emphasized that DNA methylation and chromatin modifications appear to be connected and they both operate along a common pathway to repress transcription (Curradi et al., 2002; Lewin et al., 2011).

4.3. Long-distance regulation of gene expression

Three classical modes of action-at-a-distance have been reported in eukaryotes: looping, sliding and oozing (Talbert and Henikoff, 2006). It is important to note that exactly three distinct mechanisms of long-distance modulation of gene expression have been recently confirmed in yeast: Sir-dependent silencing and transcriptional interference, which are both repressive mechanisms, and 3D clustering with a subset of co-regulated genes, which enhances transcription, transcription in yeast having been pointed out to be similar to those in higher eukaryotes (Du et al., 2017).

Enhancer-promoter interaction is a widely known long-distance chromatin looping interaction that activates expression (Harmston and Lenhard, 2013; Krivega and Dean, 2012). It has been reported that enhancer's influence can depend on genome distance between enhancer and promoter (MacIsaac et al., 2010; Quintero-Cadena and Sternberg, 2016).

Spreading of silent chromatin is another action-at-a-distance mechanism that silence genes (Lewin et al., 2011; Talbert and Henikoff, 2006). In particular, the *spread of DNA methylation* is believed to be a

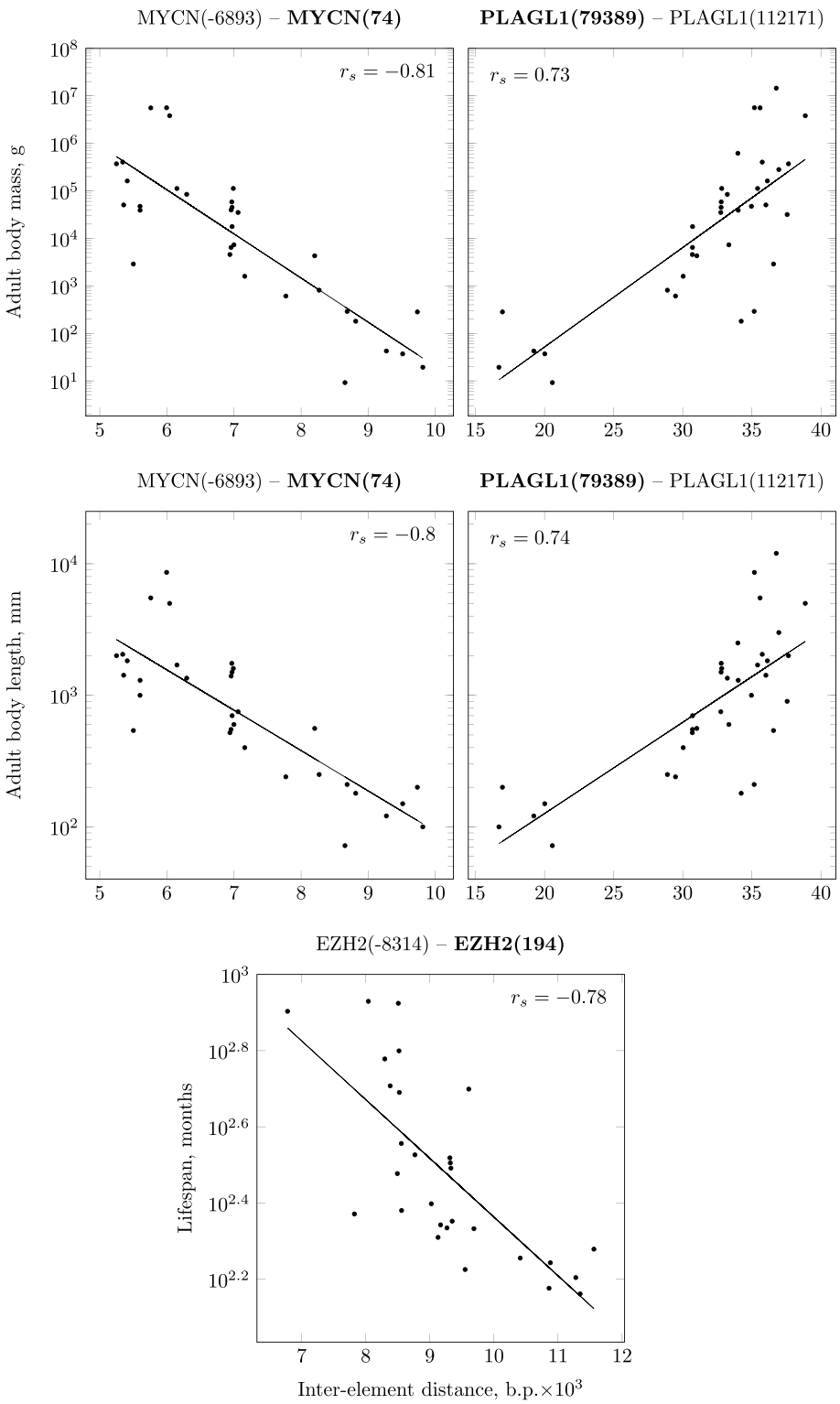

Fig. 1. Semilog plots of the inter-element distance and morpho-physiological trait for the pairs of the topmost correlated conserved elements in the neighborhoods of the *Mycn*, *Plagl1* and *Ezh2* genes. Elements, which overlap gene promoters, are highlighted in bold. Spearman's rank correlation coefficient is provided for every plot with the line of best fit.

potent silencing mechanism in mammals (Turker, 2002; Yates et al., 2003). It has been reported that DNA methylation spreading in higher eukaryotes appears to be analogous to Sir-dependent silencing in yeast (Sneppen and Dodd, 2015).

4.4. Correlation between the inter-conserved element distances and morpho-physiological traits of mammals

Genome distances between some conserved elements in neighborhoods of 17 genes studied were found to be significantly correlated with the adult body mass and length for some pairs of the conserved

Table 5
The DNA sequences of the topmost correlated conserved elements. Elements, which overlap gene promoters, are highlighted in bold.

Element	Sequence
MYCN(−6893)	TCCAGGACAGTCTCCAGCTGAACTTTGAGGAAAGAATTTGAATAAGCAGAAATGAAAGGAAAAGCCTTTCTAGGCCAGAAGTAGCAG-CTTCTGGAAAT
MYCN(74)	GAAAAGCAAGTGGCTTTTGGCGCGAAAGCCTTGGCGCCTCCCCTGATTTTTATGGAAATCAGGAGGGCGGGGTAAAGCCGCTTTCCT-CTCCTTTCTCCCTCCCCCTTGTCTGCGCCACAGCCCCCTTCTCTCCCCGCCCCCCGGGTGTGTCAGATTTTTCAGTTAATAATATCC-CCCGAGCTTCAAAGCGCAGGCTGTGACAGTCATCTGTCTGGACGCGCTGGGTGGATGCGGGGGGCTCCTGGGAACTGTGTTGGAGCC-GAGCAAGCGCTAGCCAGGCGCAAGCGCGCACAGACTGTAGCCATCC
PLAGL1(79389)	TTTAAAATTGCATATGAACCTACAGTGATCTCAAAATTATAATTAACAAATGCATGTGCAAATATTATACAAAACTAGTAATTTTTA-TTTTTGTTTGCAGATTGTTCCTGCCTGATTCATAAACATCGTG
PLAGL1(112171)	TCTTCTTCCCGTCTGATCCTCTCTCATCCCCAAAATAATCAACATGAGCAGAACGAAGCAGATGGCACAAGCCTTCTTTGGCATACT-GGCAGATCTGTATCACTCAGTATTTTTAGAATCAAGGGGACTGAAAATTATGTCCTGAGATGTTTTGGGGTAAGACTTACTATCTTT-CTAGGGTGACAGGTGTTGAAATAGCACCGTATTTAAGACTCAAATGCAATTTCTATGTCAAGTCACAGAACAGAATGACTCGCATCA-GGACCATTGCTCTGTAAGTTCTGTTAGGATCGTTGG
EZH2(−8314)	GTCTTCTTTCACTTAACACTATATTTATCAGATTCATTCACATTGTTGTATGTAGCTATAGATTCTCTGCATCATTGTAGGATATTC-CATTATGTAAATTAACCACAATTTATTCATTCTACCTTTGATGGGCATTTGGATAGTTTCCAGTTTGCATTATGA
EZH2(194)	CGCGGGGAAACGAGCGCGGCGGTTAAAACCGTTACCACCCCCGAGTTTTGAACTGGTTCAAACTTGGCTTCCAGCACCCGCCCCGCC-CCTCCCCCGCCCGGGAACTCTGCGGCGCCGGTTCCCGCCAAGAGCCGCCGGCGCTTCGTCCCGCCCTTCGGCCGGTTCCCGCCACCT-ATCCTCCCCGCCTCCCGTCCGCGGCGGGCTCCGGGCCCCCGCGATGTCTCCCGGTCCCCGCGTGCCTGCACACCGCCTTCCTGAGAG-GCGCCGTGTGTTCAGCGAAAGAACAAAGAGACGGCGGCGGCGCTTCCACACGGCCAGTGGCGTCCCTTACAGCGAACCCCGCCGCCG-CCCGCGCGCGCACGCGCTGCCAGTGCCCGCCCGCCCACGAGCCCTGAGCGCACTCTGCGTGGGGCTGGCTCGGCGCCTCCGAGCCCG-GCGGGCCCTGTGATTGGACGGGCGCCCGCCTCGCGTCCCGCCAATCGGGGCGGCGCTTGATTGGGCTGGGGGGGCCAAATAAAAGCG-ATGGCGATTGGGCTGCCGCGTTTGGCGCTCGGTCCGGTCGCGTCCGACACCCGGTGGGACTCAGAAGGCAGTGGAGCCCCGGCGGCG-GCGGCGGCGGCGCGCGGGGGCGACGCGCGGGAACAACGCGAGTCGGCGCGCGGGACGAAGGTAACGCGCCGCTGCGGGCGGCCCGGC-CGGCGGGGCTCCGGGAGTGCGAACCGGGCGGCG

elements in the neighborhoods of *Mycn* and *Plagl1* genes and with the lifespan for some pairs of the conserved elements in the neighborhood of *Ezh2* gene.

It should be noted that for each of these genes the topmost correlated pair of the conserved elements contained the element, which overlapped the promoter or the start site of the same gene, the relationship between inter-element distance and morpho-physiological trait appearing to be exponential. It suggests that interaction of the promoter and distal regulatory site can be involved in regulation of the *Mycn*, *Plagl1*, and *Ezh2* genes, the influence depending on both site epigenetic state and genome distance between the site and gene promoter.

The variance of inter-element distances between species appears to be resulting from indels of the SINE elements (like the Alu elements in primates or another species-specific interspersed elements) and short tandem repeat (STR) length variations in non-coding genome. Indeed, longer distance between the MYCN(-6893) and MYCN(74) elements in *Homo sapiens* vs *Loxodonta africana* is stipulated by the insertion of some Alu elements. At the same time, longer distance between these elements in *Mus musculus* vs *Homo sapiens* seems to be connected with the expansion of STRs and insertions of mouse-specific SINEs. Remarkably, many conserved regions are enriched with STRs that co-localize with regulatory *cis*-elements. Moreover, the STR length variation can modulate certain histone modifications, eventually affecting the expression of the gene (Gymrek et al., 2016).

Both positive and negative correlation between inter-conserved element distances and morpho-physiological traits can indicate that two distinct modulation mechanisms can be involved.

4.4.1. Negative correlation

Negative correlation was revealed between the adult body mass and length and some inter-conserved element distances in the neighborhood of the *Mycn* gene and between the lifespan and some inter-conserved element distances in the neighborhood of the *Ezh2* gene.

The proximity of the regulatory site and the gene promoter in massive mammals and their remoteness in small mammals suggests *activation* influence. It was observed that adult body mass drops almost exponentially with increasing of inter-conserved element distance. Thus, distance dependent activation mechanism of interaction between the distal regulatory site and the promoter, where influence decreases with inter-element distance, may lie at the basis of this regulation.

The enhancer-promoter interaction could explain these observations. Indeed, it has recently been demonstrated that enhancer's influence can fall off *exponentially* with distance from the promoter (Quintero-Cadena and Sternberg, 2016). It is interesting to note that one element — the MYCN(−6893) — from the topmost negatively correlated pair of conserved elements in the neighborhood of the *Mycn* gene — MYCN(−6893) and MYCN(74) — overlaps the GATA2 binding site, the GATA2 having been evidenced to be a component of transcription enhancer complexes (He et al., 2014).

On the other hand, it has been shown that enhancer methylation may serve as main determinant of gene expression, gene expression appearing to be inversely proportional to degree of enhancer methylation (Aran et al., 2013). Another study has also demonstrated that enhancer's influence decreases with the increase of degree of its methylation (Hwang et al., 2015). It is worth noting that enhancer methylation may be more closely related to changes in gene expression than promoter methylation (Aran and Hellman, 2013).

Thus, the next regulatory model is proposed:

- Distal regulatory site is an enhancer.
- Enhancer influence decreases *exponentially* with enhancer-promoter distance.
- Enhancer influence decreases *linearly* with degree of enhancer's methylation.
- Degree of enhancer's methylation increases logarithmically until adulthood and then slows to linear dependence later in life.
- In the initial moment regulatory site is not methylated.
- Growth velocity is proportional to the gene expression.

As a consequence, the closer regulatory element is located to promoter, the longer gene is expressed, gradually downregulating with age, inter-element distance *modulating* gene expression and eventually determining phenotypic variations between species.

4.4.2. Positive correlation

Positive correlation was observed between the adult body mass and length and the inter-conserved element distance for only one pair of the conserved elements in the neighborhood of the *Plagl1* gene.

The proximity of the regulatory site and the gene promoter in small

Table 6
Known genomic elements from various databases, which overlap conserved elements in the neighborhoods of growth-regulating *Mycn*, *Plagl1* and *Ezh2* genes. Elements, which overlap gene promoters, are highlighted in bold. Some conserved elements overlapped many known genomic elements, the total number of overlapped elements being only provided and highlighted in italic.

Element	EPD	ENCODE	JASPAR	ORegAnno	Clin. sig. dbSNP	Clin. sig. dbVar	DENdb DHS
MYCN(−25368)		GATA2					
MYCN(−6893)		GATA2		STAT1			
MYCN(−4493)							
MYCN(−1587)		*6 TFs*	ESRRB	PRDM14, TBX3			
MYCN(74)	MYCN_2, MYCN_1, Mycn_1	*12 TFs*		TFAP2C, SMARCA4		nsv997222	*
MYCN(1273)		SIN3A, RAD21, EZH2, TAF1				nsv997222	*
MYCN(3711)		*12 TFs*	FOXA1, ESRRB, Klf4	FOXA1, CTCF, ETS1, ATOH1, RUNX1		nsv997222	*
PLAGL1(79389)				Cdx1		nsv997068, nsv510056	
PLAGL1(112171)			NFIC	Cdx1		nsv997068	
EZH2(−8314)				E2F1			
EZH2(194)	Ezh2_1, EZH2_1, Ezh2_2	*62 TFs*	EGR1, TCF7L2, NFYB, STAT3, E2F1, E2F3, E2F4, NRF1, E2F6	RB1, RBL2, EGR1, TFAP2C, STAT1, E2F4, FOS, ETS1, SMARCA4	rs550913824, rs886062082, rs545716282, rs886062081, rs886062083		*
EZH2(1627)	EZH2_2	*12 TFs*		STAT1, ETS1, SMARCA4			*
EZH2(2278)		*8 TFs*	STAT1, STAT3, JUND	STAT1, SMARCA4			*

Table 7
Clinically significant dbVar elements, which *overlap* the genomic interval between pairs of the topmost correlated conserved elements. Elements, which overlap gene promoters, are highlighted in bold.

Element 1	Element 2	Clin. sig. dbVar
MYCN(−25368)	MYCN(−4493)	
MYCN(−25368)	**MYCN(74)**	
MYCN(−6893)	MYCN(−1587)	
MYCN(−6893)	**MYCN(74)**	
MYCN(−6893)	**MYCN(1273)**	nsv997222
MYCN(−6893)	MYCN(3711)	nsv997222
PLAGL1(79389)	PLAGL1(112171)	nsv997068, nsv510056
EZH2(−8314)	**EZH2(194)**	
EZH2(−8314)	**EZH2(1627)**	
EZH2(−8314)	EZH2(2278)	

mammals and their remoteness in massive mammals suggests *repressive* influence. It was found out that the adult body mass slowly increases with distance and then drastically rises on reaching a threshold distance. Thus, distance dependent repressive mechanism of interaction between the distal regulatory site and the promoter, where repressive influence decreases with the inter-element distance, may lie at the basis of this regulation. Silencer-promoter interaction, similarly to the aforementioned enhancer-promoter interaction, does not appear to underlie this regulation, since, in such case, the repression would drop exponentially with the distance and the adult body mass in small mammals would considerably rise with small changes in the distance between regulatory sites.

The only repressive mechanism of long-distance modulation of gene expression, retaining throughout the cell division, is the spread of silent chromatin. In particular, silencing can be caused by DNA methylation spreading.

It has been demonstrated that gene expression can depend on both degree of the regulatory site methylation and distance from regulatory site — methylation center — to the gene promoter, the DNA methylation spreading having been assumed as the basis of this phenomenon (Turker, 2002; Yates et al., 2003). Another study has reported that DNA methylation can inhibit a flanking promoter but a threshold of the modified CpGs was required to organize a stable, diffusible chromatin structure (Curradi et al., 2002). It should be emphasized that repressive effect may not require promoter methylation.

It is widely recognized that degree of epigenetic modification of the promoter inversely correlates with gene expression (Uno et al., 2011; Wagner et al., 2014; Walker et al., 2013). Thus, the ability of methylated DNA to influence an adjacent promoter may be a function of both the number of modified CpGs and the distance between the promoter and methylated DNA.

It is noteworthy that the *Plagl1* gene is known to be imprinted and DNA methylation spreading is widely known to affect imprinted genes (Marcho et al., 2015). It has been demonstrated that promoters of the imprinted genes can be sensitive to methylation of adjacent sequences (Stelzer and Jaenisch, 2015; Stelzer et al., 2015). Another study has revealed that insertion of the *imprinting control region* (ICR) in the neighborhood of a gene can silence even non-imprinted genes, proximity of the ICR to the gene promoter modulating repressive influence (Gebert et al., 2016). The ICR mediated silencing has not been shown to occur through DNA methylation of the promoter, however the spread of DNA methylation has been assumed to cause silencing.

Notably, it has been previously hypothesized that the coordinate decline in expression of the imprinted *Plagl1* gene is caused by altered methylation and consequent silencing of the expressed allele, but the methylation status of the *Plagl1* promoter region has been demonstrated not to change with age (Lui et al., 2008). Nevertheless, considering the above-mentioned, the spread of DNA methylation still might be involved in silencing the *Plagl1* gene, promoter methylation being not required.

Overall, the next regulatory model is proposed:

- Distal regulatory site is a methylation center.
- Methylation level of adjacent regions slowly falls with distance and drastically drops on reaching a threshold distance.
- Gene expression decreases linearly with the increase of degree of methylation of gene promoter.
- Growth velocity is proportional to the gene expression.
- Degree of methylation of regulatory site increases logarithmically until adulthood and then slows to linear dependence later in life.
- In the initial moment regulatory site is not methylated.

As a consequence, the farther the regulatory element is located to promoter, the longer the gene is expressed, gradually downregulating with age, inter-element distance *modulating* gene expression and eventually determining phenotypic variations between species.

5. Conclusion

The genome neighborhoods of 17 growth-promoting genes were investigated. We found that genome distances between some conserved elements in the neighborhoods of the *Mycn*, *Plagl1* and *Ezh2* genes were significantly correlated with some of the morpho-physiological traits in mammals.

Strikingly, for each of these genes the pair of the topmost correlated elements overlapped promoter of the corresponding gene. Moreover, in all cases the dependence between the inter-element distance and morpho-physiological trait — whether negatively or positively correlated — appeared to be exponential.

We addressed the difference of morpho-physiological traits between species with different modulation of growth-promoting gene expression. For cases of both negatively and positively correlated conserved elements, the corresponding model of gene regulation was proposed.

We first hypothesize that the genome distance between gene promoter and a regulatory site may be the main factor of evolutionary modulation of growth-promoting gene expression and eventually phenotype.

We believe in the technique we applied in our research to be used to seek for new genotype-phenotype relationships. We foresee that if modulation of the gene expression leads to conformable changing in the phenotype and difference in phenotype may point out on the type of modulation of the gene expression. We consider the presence of significant correlation between phenotype and inter-conserved element distance as a new line of evidences that may confirm predicted regulatory sites.

We bring to light the most probable genomic elements that we consider to be the targets for genome manipulations. To our mind, the most intriguing prospect of our research is the possibility to influence morpho-physiological trait of mammals simply adjusting some inter-conserved element distances in neighborhoods of some growth-promoting genes.

Data availability

Processed genome reference sequences with dot-plot annotated conserved elements used to support the findings of this study are available from the corresponding author upon request.

Declaration of competing interest

The authors have declared no competing interests.

Acknowledgements

The study was performed on the equipment of the Center of collective use “High technologies” (Southern Federal University, Rostov-on-Don, Russia) in the framework of the basic part of the Ministry of Education and Science of the Russian Federation state task on the theme: “The study of the genetic polymorphisms and microRNAs functional role in the genomes of humans and animals” [project 6.6762.2017 BP].

References

Aran, D., Hellman, A., 2013. DNA methylation of transcriptional enhancers and cancer predisposition. Cell 154 (1), 11–13.
Aran, D., Sabato, S., Hellman, A., 2013. DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. Genome Biol. 14 (3), R21.
Bogin, B., 1999. Evolutionary perspective on human growth. Annu. Rev. Anthropol. 28, 109–153.
Christensen, B.C., Houseman, E.A., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R., Sugarbaker, D.J., Yeh, R.F., Wiencke, J.K., Kelsey, K.T., 2009. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet. 5 (8), e1000602.
Curradi, M., Izzo, A., Badaracco, G., Landsberger, N., 2002. Molecular mechanisms of gene silencing mediated by DNA methylation. Mol. Cell. Biol. 22 (9), 3157–3173.
Delaney, A., Padmanabhan, V., Rezvani, G., Chen, W., Forcinito, P., Cheung, C.S., Baron, J., Lui, J.C., 2014. Evolutionary conservation and modulation of a juvenile growth-regulating genetic program. J. Mol. Endocrinol. 52 (3), 269–277.
Du, M., Zhang, Q., Bai, L., 2017. Three distinct mechanisms of long-distance modulation of gene expression in yeast. PLoS Genet. 13 (4), e1006736.
Finkielstain, G.P., Forcinito, P., Lui, J.C., Barnes, K.M., Marino, R., Makaroun, S., Nguyen, V., Lazarus, J.E., Nilsson, O., Baron, J., 2009. An extensive genetic program occurring during postnatal growth in multiple tissues. Endocrinology 150 (4), 1791–1800.
Gebert, C., Rong, Q., Jeong, S., Iben, J., Pfeifer, K., 2016. H19ICR mediated transcriptional silencing does not require target promoter methylation. Biochem. Biophys. Res. Commun. 476 (3), 121–126.
Gymrek, M., Willems, T., Guilmatre, A., Zeng, H., Markus, B., Georgiev, S., Daly, M.J., Price, A.L., Pritchard, J.K., Sharp, A.J., Erlich, Y., 2016. Abundant contribution of short tandem repeats to gene expression variation in humans. Nat. Genet. 48 (1), 22–29.
Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.B., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S., Ideker, T., Zhang, K., 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell 49 (2), 359–367.
Harmston, N., Lenhard, B., 2013. Chromatin and epigenetic features of long-range gene regulation. Nucleic Acids Res. 41 (15), 7185–7199.
He, B., Lanz, R.B., Fiskus, W., Geng, C., Yi, P., Hartig, S.M., Rajapakshe, K., Shou, J., Wei, L., Shah, S.S., Foley, C., Chew, S.A., Eedunuri, V.K., Bedoya, D.J., Feng, Q., Minami, T., Mitsiades, C.S., Frolov, A., Weigel, N.L., Hilsenbeck, S.G., Rosen, D.G., Palzkill, T., Ittmann, M.M., Song, Y., Coarfa, C., O'Malley, B.W., Mitsiades, N., 2014. GATA2 facilitates steroid receptor coactivator recruitment to the androgen receptor complex. Proc. Natl. Acad. Sci. U. S. A. 111 (51), 18261–18266.
Horvath, S., 2013. DNA methylation age of human tissues and cell types. Genome Biol. 14 (10), R115.
Hsu, C.L., Chang, H.Y., Chang, J.Y., Hsu, W.M., Huang, H.C., Juan, H.F., 2016. Unveiling MYCN regulatory networks in neuroblastoma via integrative analysis of heterogeneous genomics data. Oncotarget 7 (24), 36293–36310.
Hwang, W., Oliver, V.F., Merbs, S.L., Zhu, H., Qian, J., 2015. Prediction of promoters and enhancers using multiple dna methylation-associated features. BMC Genomics 16 (7), S11.
Issa, J.P., 2014, Jan, Jan. Aging and epigenetic drift: a vicious cycle. J. Clin. Invest. 124 (1), 24–29.
Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L., Safi, K., Sechrest, W., Boakes, E.H., Carbone, C., Connolly, C., Cutts, M.J., Foster, J.K., Grenyer, R., Habib, M., Plaster, C.A., Price, S.A., Rigby, E.A., Rist, J., Teacher, A., Bininda-Emonds, O.R.P., Gittleman, J.L., Mace, G.M., Purvis, A., 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology 90, 2648.
Kans, J., 2013. Entrez Direct: E-utilities on the UNIX Command Line. (Apr).
Krivega, I., Dean, A., 2012. Enhancer and promoter interactions-long distance calls. Curr. Opin. Genet. Dev. 22 (2), 79–85.
Lee, S.R., Roh, Y.G., Kim, S.K., Lee, J.S., Seol, S.Y., Lee, H.H., Kim, W.T., Kim, W.J., Heo, J., Cha, H.J., Kang, T.H., Chung, J.W., Chu, I.S., Leem, S.H., 2015. Activation of EZH2 and SUZ12 regulated by E2F1 predicts the disease progression and aggressive characteristics of bladder cancer. Clin. Cancer Res. 21 (23), 5391–5403.
Lewin, B., Krebs, J., Kilpatrick, S., Goldstein, E., 2011. Lewin's GENES X. In: Jones and Bartlett books in computer science Jones & Bartlett Learning.
Liu, G.E., Weirauch, M.T., Van Tassell, C.P., Li, R.W., Sonstegard, T.S., Matukumalli, L.K., Connor, E.E., Hanson, R.W., Yang, J., 2008. Identification of conserved regulatory elements in mammalian promoter regions: a case study using the PCK1 promoter. Genomics Proteomics Bioinforma. 6 (3-4), 129–143.
Lui, J.C., Baron, J., 2011. Mechanisms limiting body growth in mammals. Endocr. Rev. 32 (3), 422–440.
Lui, J.C., Baron, J., 2013. Evidence that Igf2 down-regulation in postnatal tissues and up-regulation in malignancies is driven by transcription factor E2f3. Proc. Natl. Acad.

Sci. U. S. A. 110 (15), 6181–6186.
Lui, J.C., Chen, W., Barnes, K.M., Baron, J., 2010a. Changes in gene expression associated with aging commonly originate during juvenile growth. Mech. Ageing Dev. 131 (10), 641–649.
Lui, J.C., Chen, W., Cheung, C.S., Baron, J., 2014. Broad shifts in gene expression during early postnatal life are associated with shifts in histone methylation patterns. PLoS ONE 9 (1), e86957.
Lui, J.C., Finkielstain, G.P., Barnes, K.M., Baron, J., 2008. An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. Am. J. Physiol. Regul. Integr. Comp. Physiol. 295 (1), R189–196.
Lui, J.C., Forcinito, P., Chang, M., Chen, W., Barnes, K.M., Baron, J., 2010b. Coordinated postnatal down-regulation of multiple growth-promoting genes: evidence for a genetic program limiting organ growth. FASEB J. 24 (8), 3083–3092.
Lui, J.C., Garrison, P., Baron, J., 2015. Regulation of body growth. Curr. Opin. Pediatr. 27 (4), 502–510.
MacIsaac, K.D., Lo, K.A., Gordon, W., Motola, S., Mazor, T., Fraenkel, E., 2010. A quantitative model of transcriptional regulation reveals the influence of binding location on expression. PLoS Comput. Biol. 6 (4), e1000773.
Marcho, C., Bevilacqua, A., Tremblay, K.D., Mager, J., 2015. Tissue-specific regulation of Igf2r/Airn imprinting during gastrulation. Epigenetics Chromatin 8, 10.
Nandakumar, S.K., Johnson, K., Throm, S.L., Pestina, T.I., Neale, G., Persons, D.A., 2015. Low-level GATA2 overexpression promotes myeloid progenitor self-renewal and blocks lymphoid differentiation in mice. Exp. Hematol. 43 (7), 565–577.
Quinlan, A.R., 2014. BEDTools: The Swiss-Army tool for genome feature analysis. Curr. Protoc. Bioinforma. 47, 1–34.
Quintero-Cadena, P., Sternberg, P.W., 2016. Enhancer Sharing Promotes Neighborhoods of Transcriptional Regulation Across Eukaryotes. G3 (Bethesda) 6 (12), 4167–4174.
Romanov, D.E., Butenko, E.V., Shkurat, T.P., 2019. Genome distance between growth-regulating genes and telomeres is correlated with morpho-physiological traits in mammals. Gene Rep. 14, 124–128.
Shkurat, T., Romanov, D., Pshenichnyy, E., Ponomareva, N., Alexsandrova, A., Bakhtadze, G., Lomteva, S., 2015. Prevalence of mirnas in introns and cis-regulatory regions of genes of the somatotropic axis in mammals. Am. J. Appl. Sci. 12, 1–7.
Sneppen, K., Dodd, I.B., 2015. Cooperative stabilization of the SIR complex provides robust epigenetic memory in a model of SIR silencing in *Saccharomyces cerevisiae*. Epigenetics 10 (4), 293–302.
Stelzer, Y., Jaenisch, R., 2015. Monitoring dynamics of DNA methylation at single-cell resolution during development and dsease. Cold Spring Harb. Symp. Quant. Biol. 80, 199–206.
Stelzer, Y., Shivalila, C.S., Soldner, F., Markoulaki, S., Jaenisch, R., 2015. Tracing dynamic changes of DNA methylation at single-cell resolution. Cell 163 (1), 218–229.
Talbert, P.B., Henikoff, S., 2006. Spreading of silent chromatin: inaction at a distance. Nat. Rev. Genet. 7 (10), 793–803.
The Apache Software Foundation, 2006. Commons math: The apache commons mathematics library. (Aug).
Turker, M.S., 2002. Gene silencing in mammalian cells and the spread of DNA methylation. Oncogene 21 (35), 5388–5393.
Uno, M., Oba-Shinjo, S.M., Camargo, A.A., Moura, R.P., Aguiar, P.H., Cabrera, H.N., Begnami, M., Rosemberg, S., Teixeira, M.J., Marie, S.K., 2011. Correlation of MGMT promoter methylation status with gene and protein expression levels in glioblastoma. Clin. (Sao Paulo) 66 (10), 1747–1755.
Wagner, J.R., Busche, S., Ge, B., Kwan, T., Pastinen, T., Blanchette, M., 2014. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. Genome Biol. 15 (2), R37.
Walker, C.G., Littlejohn, M.D., Meier, S., Roche, J.R., Mitchell, M.D., 2013. DNA methylation is correlated with gene expression during early pregnancy in *Bos taurus*. Physiol. Genomics 45 (7), 276–286.
Yates, P.A., Burman, R., Simpson, J., Ponomoreva, O.N., Thayer, M.J., Turker, M.S., 2003. Silencing of mouse Aprt is a gradual process in differentiated cells. Mol. Cell. Biol. 23 (13), 4461–4470.