nature ecology & evolution

SUPPLEMENTARY INFORMATION

VOLUME: 1 | ARTICLE NUMBER: 0055

In the format provided by the authors and unedited.

# Antagonistic pleiotropy and mutation accumulation influence human senescence and disease

**Authors:** Juan Antonio Rodríguez, Urko M. Marigorta, David A. Hughes, Nino Spataro, Elena Bosch and Arcadi Navarro

Correspondence to: arcadi.navarro@upf.edu, elena.bosch@upf.edu

**Supplementary Sections:**

## Section 1: Basic background on ageing research

The evolutionary and immediate causes of ageing are a fundamental biological problem that has been the subject of extensive literature. The objective of this note is to provide some basic background on the evolutionary theories and empirical data on the problem of ageing, so that our results can be better appreciated by readers.

Many hypotheses have been proposed to explain senescence (for recent reviews see references 1–5), but the sheer diversity of ageing patterns across the tree of life indicates that there is no single theory that can be reconciled with all the data[6]. This huge heterogeneity is at the core of evolutionary research on ageing, which tries to understand not only why some species and not others decay with age, but also what explains the differences in their rates of senescence[4,7]. Some hypotheses offer explanations to account for these differences that are more or less adequate depending on the particular organism they consider, as well as on its corresponding environmental and/or life-history conditions. For instance, the ageing mechanisms that might give the best explanation for the lack of senescence in Hydras are likely to be different from the ageing factors in salmons or humans.

A brief overview on the main theories of ageing can be seen in Supplementary Table 1. It can be appreciated that, over time, many hypotheses have been proposed. However, when it comes to explaining senescence patterns in our species, two classical, non-mutually exclusive, theories have become the most

popular among evolutionary biologists: **the Mutation Accumulation**[8] **(MA)** and the **Antagonistic Pleiotropy (AP)**[9] theories of ageing.

| Group | Name | Definition | Refs. |
|---|---|---|---|
| **Maladaptive ageing** | Mutation Accumulation | Since late-acting deleterious alleles are less effectively removed by natural selection than early acting deleterious alleles, harmful mutations will tend to accumulate | 8 |
| | Balance wear/tear - genetic repair | Patterns of senescence are a balance between somatic wear/tear *vs.* genetic repair controls. | 7 |
| | Somatic Damage | Senescence comes as an accumulation product of physical / mechanical wearing. | 10 |
| **Secondary ageing** | Antagonistic Pleiotropy | Natural selection will favor beneficial mutations in the fertile stages of life at the cost of deleterious effects later, contributing to senescence. | 9 |
| | Disposable Soma | Life history strategies favor either longevity (repair/maintenance of the organism, i.e. the Soma) or reproduction. | 11 |
| | Allocation Theory | Related to Disposable Soma. Species face compromises of allocating limited resources to growth vs. maintenance vs. reproduction vs. escaping predators and pathogens. Ageing patterns will depend on the set of choices made by the species. | 12–15 |
| **Assisted / Programmed ageing** | Assisted death | A controversial view according to which longevity is seen as a detrimental trait for the organism, while death is seen as a favorable trait, so certain genetic programs activate to inhibit or delay a longer lifespan. | 16 |

**Supplementary Table 1 |** Summary of some of the main theories of ageing. Theories by themselves can be grouped into a higher classification order (first column), depending on the senescence outcome expected from the theory. Modified from Trindade *et al.,* 2013 [4]

Building on Fisher's concept of reproductive value[17], Peter Medawar proposed the foundations of the MA theory in 1951[18]. Medawar proposed the general concept, but its correct mathematical development was provided 20 years later

by William D. Hamilton[19] (but see a comment by Brian Charlesworth[20]). There is a steady reduction in the numbers of individuals belonging to a given birth cohort as age increases, as a result of extrinsic sources of mortality. Mutations which increase survival early during the reproductive period will affect the numbers of their carriers more greatly than mutations acting later; mutations affecting fertility will be selected most strongly, the earlier they act. In other words: the action of natural selection will be more effective in improving survival or fertility early rather than late in life. As a result, late-acting deleterious alleles will be under weaker purifying selection and will tend to reach higher frequencies, compared with early-acting one.

Medawar further suggested that the same allele could have different effects at different periods in life, such that a genetic variant increasing fitness early in life could be favored despite a negative effect later in life, particularly if the organism is no longer able to reproduce. Such pleiotropic effects were at the basis of the AP theory of ageing, verbally developed by George C. Williams in 1957[9] and mathematically modeled in 1994 by Charlesworth[21]. Recent theoretical work[22] shows that the appearance by mutation of antagonistic pleiotropy alleles is theoretically expected, but that it would be curtailed by senescence itself. In other words, the phenomenon of AP may happen, but only up to a certain point since senescence *per se* reduces the probability that pleiotropic mutations appear.

Even if theory shows that, under certain conditions AP is to be expected, previous work dealing with that theory resulted in a variety of somewhat

contradictory (and sometimes unexpected) experimental results. The existence of the AP phenomenon is supported by empirical work in *Drosophila*[23–25]. In 2000 Stearns *et al.*[25] showed that, as predicted by both MA and the AP hypothesis of ageing[9], a higher extrinsic mortality rate (such as diseases and predation) favors the evolution of a higher intrinsic mortality rate or actuarial senescence (deterioration of the organism, independently of the environment). Classical theories of ageing assume this separation between intrinsic and extrinsic mortality rate. However, Koopman *et al.* [26] argue that senescence and death are the result of interactions between genes and the environment and that, thus, both types of mortality rates should not be partitioned. Further evidence for AP was found by Bryant *et al.,* (2004)[27], when patterns of actuarial senescence were quantified from field data for two distinct guppy (*Poecilia reticulata*) populations. Both populations were fetched from the wild, but one of them came from a high predation habitat. The observation that females from this latter population suffered an earlier onset and higher rate of senescence fulfilled the expectations of the classical theories. This result was further corroborated in wild swan populations followed through 36 years[28]. However, other authors have found the opposite patterns: that extrinsic mortality does not necessarily trigger a faster onset of senescence[29–31], thus contradicting some of the original predictions by Williams[9]. In particular, Reznick *et al*. (2004)[32] observed that guppies from a high predation habitat did not show earlier onset of senescence, as measured by either mortality or reproduction, relative to low predation guppies. These last findings do not satisfy neither the MA nor the AP hypotheses and highlight that different mechanisms of ageing may have had distinct roles in the evolutionary histories of different species. However, in the

same paper, the authors showed that guppies coming from high predation habitats presented an earlier onset of senescence in terms of physiological responses, measured as swimming performance. That is, organismal decay does seem to occur after the expected age, at least when measured with particular parameters.

The results above highlight another important question: how should senescence be measured? One of the most debated problems on senescence is its quantification[33]. Given that our analysis focuses on modern humans and disease risk, we did not use other relevant parameters such as patterns of reproduction, fertility, or intrinsic senescence. However, GWAS' data on age-specific diseases and their associated genetic variants do fulfill some of the most obvious predictions of MA and AP. Also, our results indicate that the onset of senescence occurs at an age threshold of around 40-50 years old, which fits with many sources of evidence[34–37]. Even if in modern societies mortality does not generally rise from that age[38], it is clear that a physiological decline does take place.

At any rate, our results cannot possible indicate that AP or MA are the only senescence hypotheses that have played a role in our evolution. Other theories could also be important, but available data precludes testing them. One particularly well-known alternative or complementary theory is the "**Disposable soma theory" (DS)**, formulated by Thomas Kirkwood in 1977[11] and further developed by Kirkwood and Holliday (1979)[39]. This theory suggests that organisms have to choose between dedicating energy to reproduction or invest

it in soma preservation and growth, with the trade-off of an increased lifespan and delayed senescence. Depending on an organism's particular life-history traits and on the outcome of both processes certain resource-allocation choices will produce higher and more optimal fitness than others. In brief, the "*soma*" is just a large container for reproductive material that, under certain circumstances, can be disposed of when prioritizing reproduction. This idea can be applied to all organisms presenting a distinction between somatic and reproductive structures. Throughout life, germinal cell line does not accumulate damage, but somatic cells do, manifesting senescence[40,41]. However, there is an exception to this: In the genus *Hydra*, a water polyp (phylum Cnidaria) there is not distinction between the somatic and the reproductive structures. These organisms do not show symptoms of senescence[42] and if maintained under benign conditions and asexual reproduction they could live for an indefinite time[6,41].

In any case, both the AP and DS theories can easily be reconciled. Many authors, including T.B.L Kirkwood, agree that DS is a particular instance of AP[39,43]. This could be expressed through the idea that there may be mutations that foster reproduction at the cost of disabling or impairing growth and maintenance of the somatic parts of the organism. In the end, both theories support the idea that ageing is the result of the action of genes and of natural selection[44]. In addition, both theories can be framed within a subgroup of "*secondary aging*" theories (Supplementary Table 1) which, in general terms, suggests that senescence is simply a by-product of an investment early in life: AP specifies that particular genetic variants from particular genes will "cause"

ageing or physiological decay in exchange of their positive effects early in life; while DS poses that senescence occurs because of the genetic variants that have been favored when fostering reproduction at the expense of somatic maintenance and repair functions. This lifelong, but limited, system of repair and maintenance will eventually lead to the accumulation of molecular and cellular damage, deriving stochastically into a senescence process. Again, and given the nature of our data, our observations cannot possible exclude that the DS theory is playing a fundamental role in the senescence patterns of our species.

Neither of these theories has been free of criticism. One of the main issues, raised by Vaupel *et al.* 2004[45], would be that both theories, AP and MA consider the Hamiltonian view of senescence[19], which, according to these authors, poses the universality of senescence and the decreasing effectiveness action of natural selection with age. Hamilton stated that "*senescence is universal*" and indeed, as Medawar first pointed out, a life-history that has an age-independent mortality rate generates the pressure to evolve senescence[8]. In spite of these clear theoretical expectations, cornerstones for skepticism were set by Curtsinger *et al*. (1992)[46] and Carey *et al.* (1992)[47] who failed to find direct evidence of the expected patterns of senescence after maturity in *Drosophila* and medflies. For both organisms, there is an increase in mortality after reproduction but it levels off late in life, which some authors have explained in terms of AP and MA[48–50]. Still, theoretical expectations are not always obviously met. For instance, an exhaustive literature exploration of death rates in six animal species by Vaupel *et al.,* (1998)[38] showed that senescence patterns are not universally increasing after sexual maturity. Recent work by Jones *et al*.

(2014)[12], has stressed this fact, revealing a huge variety of senescence patterns in nature that range from positive senescence, as depicted by Hamilton[19] to negative senescence and including negligible senescence, already defined by C. E. Finch (1990)[36]. The concept of negative senescence refers to death rates falling with age. It was introduced by Vaupel, Baudisch and colleagues in three seminal works[14,45,51] and it is particularly important since it illustrates that an all-encompassing theory of senescence has not yet been conceived. Within a sophisticated theoretical framework, Vaupel *et al.,* (2004)[45] showed that it was conceptually possible to develop negative senescence for particular species under particular conditions and proved that Hamilton's predictions were, at least, not always correct for all species. The comprehensive work of Jones *et al*. (2014)[6] assesses ageing patterns across the tree of life. This high diversity in ageing patterns demonstrates by itself that classical ageing theories cannot explain the enormous extant amount of variation in life-history trajectories. For example, the already discussed *Hydra magnipapillata*[41], which may show negligible senescence during its whole life; the desert tortoise (*Gopherus agassizii*), or plants such as the white mangrove (*Avicernia marina*) which may present, according to some authors, negative senescence[6].

These results question the Hamiltonian view of senescence, but other kinds of evidence support it. More technically, Michael Rose[52] claimed that even if the models by Vaupel and Baudisch[45,51] are interesting for life-history models, they are focused on clonally reproducing organisms for which the Hamiltonian models can not apply. According to Rose et al[52], Hamilton's equations have been misinterpreted. The results of the late-life plateau in mortality, the slowing

down of mortality rates previously reported by Carey, Curtsinger and others[46,47], can be reconciled with the Hamiltonian theory, since that the force of natural selection will eventually fall to zero when survival and reproduction stop in a population[53,54]. These plateaus will extend for an indefinite late age period. But the fact that they do not decrease implies that slowing of mortality rates seen in particular organisms under laboratory conditions[38] can be explained by the incapability of natural selection to distinguish between fitness differences in survival at different ages after Hamilton's natural selection functions plateau at zero[52]. Wrapping it all up, the cases of late-life plateaus for mortality and fecundity may be considered corollaries of Hamilton's theory, even if this is not apparent at first sight and one needs to add on postulates concerning effects of late-acting mutations on mortality during the reproductive period[49]. Another plausible option to explain these plateaus is genetic heterogeneity, as suggested by several authors[55]. This possibility was rejected for *Drosophila*, since plateaus were seen for both isogenic and heterogenic lines. Contrasting and ambiguous results were obtained for *C.elegans* [56,57].

In short, our findings should be framed within the large body of literature whose surface we have only scratched above. The disease-mediated effects we detect could also exert an impact in unknown endophenotypes and contribute, trough disease protection, to a "healthier ageing" after 80's. In fact, genetic variants acting like this have been previously found[58,59]. Our results fit within the general observations in our species, but, of course, that does not rule out other complementary explanations for our senescence patterns.

## Section 2: Huntington's disease and Haemochromatosis as potential examples of mutation accumulation and antagonistic pleiotropy

*Huntington's disease*

Huntington's disease (HD) is a rare autosomal dominant neurodegenerative disease, with an onset generally in post-reproductive stages. Antagonistic pleiotropy has been suggested to play a role in this condition elsewhere[60–62] as mechanism to explain its prevalence and manifestation in late ages [60]. Increased fitness (as measured by number of siblings) in HD affected individuals was already reported during the sixties and seventies[61], even before the causative molecular mechanism of the intergenerationally increasing number of CAG trinucleotide repeats in HD was known[63]. In particular, the first solid observation of increased offspring in HD patients came in 1975, when it was described that HD patients' offspring was up to 39% higher than of the unaffected controls[64]. At the same time, HD patients have been shown to present reduced rates of some cancers when compared to controls[62,65]. Such reduction has been associated to an increase in *p53* tumor suppressor activity[66], which induces a higher apoptosis rate. While this higher apoptosis rate may be the cause for the neurodegeneration episodes in HD, it has been suggested that it may also be protective for progression of malignant cancer cells. Thus, a yet unknown mechanism increasing fertility and a higher cell apoptosis rate in HD patients represent two potential cases of antagonistically pleiotropic behavior in carriers of this condition. However, an explanation for the

increasing fitness benefit has not been fully proved for this trait, and mutation accumulation could provide a more suitable explanation for it[67]

*Haemochromatosis*

Hereditary haemochromatosis (HH) is an autosomal recessive disorder, which increases the amount of iron in blood with age[68]. It is more frequent in populations of European ancestry (~10%)[69], where 90% of affected patients present the C282Y missense mutation (rs1800568) in the *HFE* gene[70]. Affected individuals benefit in early reproductive ages by the associated enhanced dietary iron absorption. This additional iron load results in a more efficient immune system, particularly against *Salmonella typhi* and *Mycobacterium tuberculosis*. Notably, the diseases caused by these pathogens (typhoid fever and tuberculosis, respectively), although prevalent throughout human evolution, expanded with the crowded conditions of large cities[70]. However, the deleterious effects of the C282Y mutation appear after the fertile period, when increased iron accumulation can represent up to a 33% morbidity increase for C282Y homozygotes[71]. In particular, conditions such as certain infections, neoplasias, cardiomyopathy and other neurodegenerative, endocrine and orthopedic diseases have been suggested to be fostered and worsened by raised iron loads[72].

## Section 3: Genomic mapping of pleiotropic SNPs

Ascribing genes to SNPs detected by GWAS is notoriously difficult in many cases and, correspondingly, several genes can be annotated to a single GWAS hit. The "*Reported genes*" field in the GWAS Catalog informs about any gene reported in the original study; while the "*Mapped genes*" field provides information about the gene/s closest to each associated marker. Moreover, the "*context"* column informs about the location of the SNP (exon, intron, 5'UTR). We classified each of the unique 219 SNPs linked to pleiotropies by considering both the "*Mapped genes"* and the "*Context*" columns by calling a single genomic locus for every unique combination of both entries. Doing so, we identified a total of unique 166 genomic loci for the whole set of 219 SNPs, out of which 55 were located in intergenic regions, 71 in introns, 7 in coding sequences, and 15 were missense mutations. The remaining loci were located either in ncRNAs, UTRs, or near-gene regions (Supplementary Table 2).

| | **Genome Wide** | **Ageing Set I** (Sousa-Victor *et al.*, 2014) | **Ageing Set II** (Magalhães *et al.*, 2009) | **Expression Set** (Harries *et al.*, 2011) |
|---|---|---|---|---|
| **Disease-associated SNPs** | 2,559 | 87 | 147 | 72 |
| **Pleiotropies** | 266 | 39 | 58 | 21 |
| **SNPs in pleiotropies** | 219 | 24 | 35 | 19 |
| **Genes/Loci in pleiotropies** | 166 loci* | 135 genes | 298 genes | 269 genes |
| **intergenic regions* | *55* | - | - | - |
| **introns* | *71* | - | - | - |
| **coding sequence* | *7* | - | - | - |
| **missense mutations* | *15* | - | - | - |
| **ncRNAs, UTRs and near-gene regions* | *18* | - | - | - |

**Supplementary Table 2 |** Summary table for number of genomic features inferred from each of the analyzed data sets.

## Section 4: Testing the proportions of agonistic *vs.* antagonistic pleiotropies using (a) three time periods; and (b) different LD thresholds

### (a) Using different time periods

In previous analyses we compared pleiotropies occurring in same *vs*. different periods of life, but it is also worth considering separately the pleiotropies that involve diseases with ages of onset within the same period of life. In that way, these pleiotropies get further classified into "*early-early*" and "*late-late*" categories, generating a 2x3 table. Under these conditions, the excess of early-late antagonistic pleiotropies around 42-50 years is still found, providing further support to our main result. Manual inspection of the corresponding tables considering the three time periods also confirms that this excess is due to the antagonistic early-late category. In particular, the ratio of antagonistic to agonistic pleiotropies, at 46 years, is higher for the antagonistic early-late group when compared to the other groups (Supplementary Table 3). Data for 2x2 analyses can be seen in Supplementary Data 4.

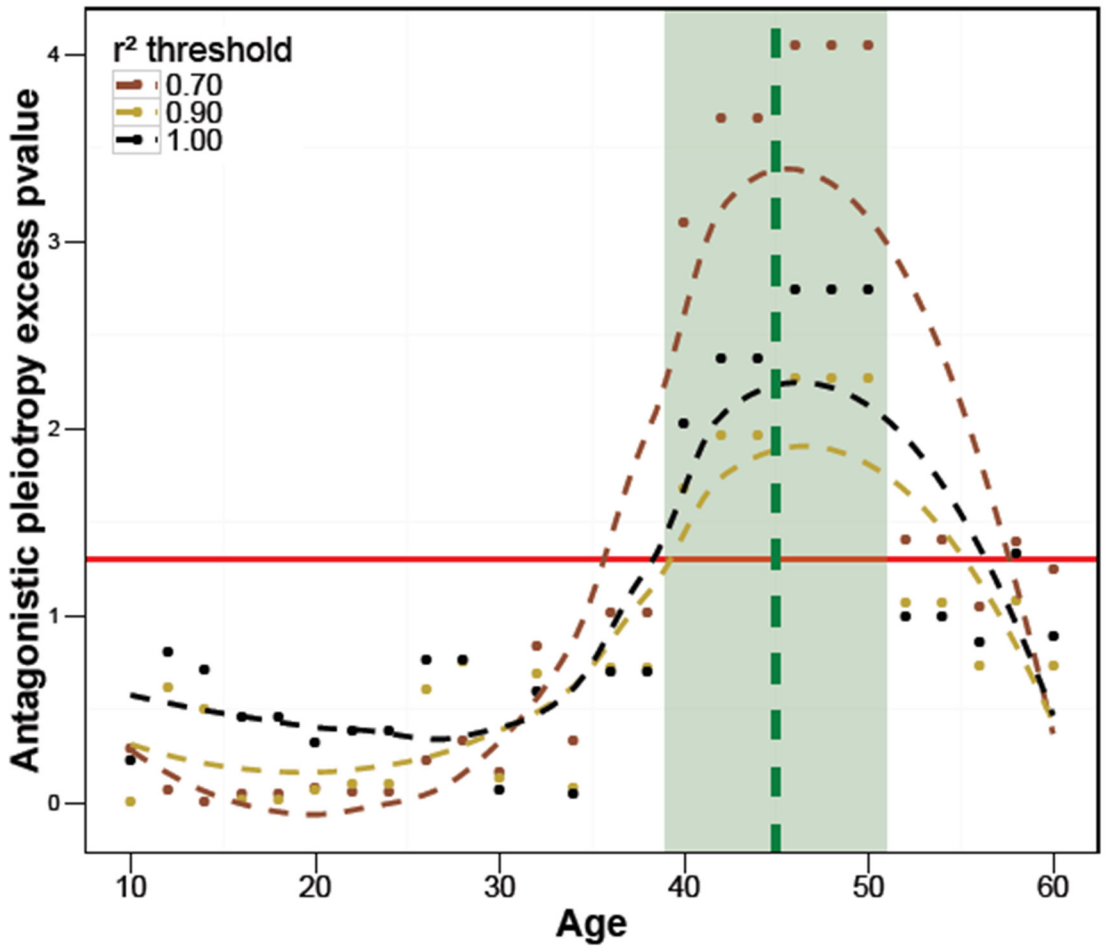
| | Antagonistic | Agonistic | Ratio |
|---|---|---|---|
| **Early-Early** | 43 | 147 | 0.29:1 |
| **Early-Late** | 26 | 27 | 0.96:1 |
| **Late-Late** | 3 | 20 | 0.20:1 |

**Supplementary Table 3 |** Distribution of pleiotropies in a 2x3 table considering an early-late transition threshold of 46 years. Note, once more, the excess proportions of antagonistic early-late pleiotropies.

### (a) Using different LD Thresholds

In Figure 1C we used as putative pleiotropies not only SNPs linked to two different diseases, but also any pair of SNPs that, even if associated to different

diseases, present LD>0.8 (see Methods). The pattern maintains when using other LD thresholds.




**Supplementary Figure 1 |** The significant excess of antagonistic early-late pleiotropies between 40 and 50 years old is consistent when using different $r^2$ thresholds. For different $r^2$ values, the Y-axis indicates -log10 (p-value) of the Chi-Square tests performed for pleiotropies at each age threshold.

**Section 5: Assessment of the overlap between pleiotropies and gene sets/genomic regions known to be related to senescence**

For some analyses, we explored pleiotropies not only for individual genes but also for particular sets of genes known to be involved either in senescence[73,74] or in expression change with ageing[75]. To do so, we compared the number of pleiotropies mapping in each particular gene set against those that would be expected from genome-wide proportions. We proceeded as follows: genes in such datasets were crossed against the downloaded GWAS database and for each gene in each list we computed how many GWAS hits had been assigned to that gene, considering as assigned if the gene was present either in the *"Reported genes"* or "*Mapped genes*" column in the GWAS Catalog. The resulting SNP set was used to find pleiotropies associated with a particular gene set. A standard hypergeometric test was then performed to evaluate whether each gene set contains an excess of pleiotropies relative to genome-wide expectations (Supplementary Data 7). To discard the possibility of an excess of pleiotropies resulting from a small set of SNPs (as one SNP can be involved in multiple pleiotropies), we compared not the number of pleiotropies but the number of SNPs found in pleiotropies. Again a hypergeometric test was used to test this SNP count table for each gene set against the background, namely all genes in the GWAS Catalog (Supplementary Data 7).

The first data set we used consists of 135 genes (Supplementary Table 2) whose expression is altered between presenescent and senescent states in a mouse model[73] (Supplementary Data 8). The human orthologs of these genes

harbor 87 unique SNPs associated with disease in the GWAS Catalog (Supplementary Data 9). These SNPs participate in a total of 39 pleiotropies (Supplementary Data 10), which is significantly higher than the genome-wide proportions of pleiotropies (266 pleiotropies out of a total of 2,559 SNPs, hypergeometric test *p*-value < $2.2x10^{-16}$) (Supplementary Data 7). The second gene set (Supplementary Table 2) consisted of 298 genes (Supplementary Data 11) that have been linked to ageing phenotypes in humans and a variety of experimental ageing models. These genes have been assembled in the online database GenAge[76]. A total of 147 disease-associated SNPs were found in that set (Supplementary Data 12), participating in a total of 58 pleiotropies (Supplementary Data 13). Again, a hypergeometric test shows a highly significant excess of pleiotropies (*p*-value < $2.2x10^{-16}$) (Supplementary Data 7).

If we limit our analysis to only early-late antagonistic pleiotropies, rather than all pleiotropies statistical power is reduced as there are only 26 early-late pleiotropies mapping to 19 loci (age threshold of 46 years) (Supplementary Data 5). Nevertheless, when merging the two ageing datasets[73,76] the excess of early-late pleiotropies in senescence-related genes remains significant (*p*-value=0.0029) (Supplementary Data 7).

The relation between pleiotropies and senescence-related genes may be due to these genes having different functions during different periods of life. This functional shift could be linked to age-related changes in gene expression and/or methylation. To test these possibilities we examined a set of 269 genes (Supplementary Table 2 and Data 14), whose level of expression changes with

age in humans[75]. These genes harbor 72 disease-associated SNPs involved in 21 pleiotropies (Supplementary Data 15 and 16), which again exceeds random expectations ($p$-value = $1.48 \times 10^{-6}$) (Supplementary Data 7).

We also studied a series of genomic regions whose methylation patterns change between centenarians and newborns[77]. We reasoned that if pleiotropy is related to ageing, pleiotropic regions could significantly overlap differentially methylated regions (DMRs). We started by creating a set of genomic regions that are comparable between pleiotropies and DMRs. This step is required because when designing GWAS chip-arrays, segmental duplications, genome alignment gaps or tandem repeats are not considered as candidate regions where to set a tag-SNP, but when searching for DMRs these regions are explored by some technologies. Thus, after downloading the corresponding tracks from the UCSC Genome browser, these types of regions were removed to generate a common comparable universe with the list of DMRs[77]. Sex chromosomes and mitochondrial DNA were also excluded.

To create “pleiotropic regions” for the 266 pleiotropies, we defined a region around the SNP(s) involved in each pleiotropy as the physical space between both SNPs plus 60 kb windows up and downstream. Since different pleiotropies might overlap we were conservative and considered a single region for those pleiotropies that overlapped or mapped at less than 200 Kb from each other, resulting in a final list of 110 unique pleiotropic regions. Next, we performed three comparisons to assess the amount of overlap between the genomic regions involved in pleiotropies and DMRs, namely:

a) all pleiotropic regions *vs.* random genome resamplings (*see main text*);

b) antagonistic early-late pleiotropic regions *vs.* random genome resamplings;

and,

c) antagonistic early-late pleiotropic regions *vs.* the rest of pleiotropic regions.

For the first two comparisons (a, b) we performed 1,000 random resamplings of as many regions from the common universe as the number of regions in each tested category. Length of the random drawn regions was matched one-to-one with that of the regions for each tested category. Each of these 1,000 sets of *n* random genomic regions was intersected with the list of DMRs, and we created a null distribution with the number of the regions overlapping with at least one DMR. We finally placed the number of pleiotropic regions, either in the whole set overlapping a DMR (a: n=72) or in the antagonistic early-late pleiotropic regions (b: n=12) (Supplementary Table 4) over the null distribution.

| | **Agonistic** | **Antagonistic** |
|---|---|---|
| **Early-Early** | 21 (*28*) | 44 (*68*) |
| **Early-Late** | 12 (*19*) | 9 (*11*) |
| **Late-Late** | 2 (*2*) | 8 (*13*) |

**Supplementary Table 4 |** Number of pleiotropies overlapping with at least one DMR in each category. In parentheses, the total number of unique genomic regions corresponding to that pleiotropic category is shown. It does not add up to 110 because overlaps are calculated independently for each category, and a given pleiotropic region may be involved in more than one type of pleiotropy. To calculate the overlap with DMRs, we just considered each region once (n=72).

We found that pleiotropic regions as a whole are enriched for DMRs (n=72/110 overlapping with at least a DMR; *p*-value=0.01; 99$^{th}$ percentile in the null distribution for 1,000 resamplings, see Main text). Additionally, when considering separately the early-early and the late-late pleiotropies (Supplementary Information Section 3), the enrichment of DMRs was still found for both categories (same period pleiotropies: *p*-value=0.01; different period pleiotropies: *p*-value=0.02).

Given that pleiotropic regions/genes/variants tend to develop several functions, it does make sense that different methylation patterns or expression levels are presented at different times. However, antagonistic early-late pleiotropic regions were not enriched in DMRs neither when compared to the rest of the genome (*p*-value=0.16) (b), nor when compared to the rest of pleiotropies (c) (Supplementary Table 5).

| | **Overlaps with DMRs** | **No overlap with DMRs** | **TOTAL** |
|---|---|---|---|
| **Antagonistic Early-Late** | 12 | 7 | 19 |
| **All other** | 65 | 33 | 98 |
| **TOTAL** | 77 | 40 | 117 |
| | | Fisher's exact test p-value = 0.79 | |

**Supplementary Table 5:** Contingency table showing no difference in DMR overlaps between antagonistic early-late pleiotropies and the other classes of pleiotropy.

## Section 6: Pleiotropy and natural selection

Natural selection may have favoured pleiotropic alleles that benefit the organism at early ages in spite of their deleterious late-age effects. Using the same approach than for DMRs, we evaluated whether pleiotropies overlap genes or genomic regions that have been recognized as targets of positive selection. As discussed in the main text, it is quite unlikely that there is a significant overlap between pleiotropies and selection, since that would imply that pleiotropy and/or senescence would have been main drivers of human adaptation and, therefore, such signals would have been detected by the many selection scans that have been run over the last decade. Indeed, the most recent and comprehensive selection scans show that there is not such signature[78–80]. Still, and for the sake of completeness we tested for a significant overlap of pleiotropies with genes that have been shown to be either under recent or ancestral positive selection.

### Recent selection

We compiled two datasets of genes under recent selection in worldwide human populations. The first one is a set of > 700 validated genomic regions from multiple genome-wide scans of positive selection in humans mostly based on SNP data[81]. Performing the same procedure of random resampling from a comparable genomic universe than in Supplementary Information Section 5, we found no enrichment in signals of positive selection for the 110 pleiotropic regions in any of the three comparisons performed:

(a) All pleiotropic regions *vs.* random genome resamplings (Fisher's exact test *p*-value=0.54);

(b) antagonistic early-late pleiotropic regions *vs.* random genome resamplings (Fisher's exact test *p*-value=0.09);

and,

(c) antagonistic early-late pleiotropic regions *vs.* the rest of pleiotropies (*p*-value=0.24, Fisher's exact test, Supplementary Table 6).

| | Overlaps with selection regions | No overlap with selection regions | TOTAL |
|---|---|---|---|
| **Antagonistic Early-Late** | 4 | 15 | 19 |
| **All other** | 10 | 88 | 98 |
| **TOTAL** | 14 | 103 | 117 |
| | | Fisher's exact test p-value = 0.24 | |

**Supplementary Table 6 |** Contingency table showing no differences in number of antagonistic early-late against all other classes of pleiotropies overlapping with regions under positive selection (Akey et al., 2009). Numbers do not add up to 110 because when merging "All other" categories we considered only unique regions.

A second dataset comes from a recent paper by Colonna *et al.*[82], who identified signals of recent adaptation in data from the 1000 Genomes Project. They used derived allele frequency differences between three main worldwide human populations to identify regions under positive selection and reported 450 candidate genes. In this set of genes we can find 117 of the disease-related SNPs used in our study, participating in a total 13 pleiotropies (using an LD threshold of $r^2 \geq 0.8$). Again, no enrichment in pleiotropies was found in the candidate genes for recent positive selection in humans (binomial test *p*-value for $x \geq 13 = 0.44$; Supplementary Table 7).

| | Genome Wide | Colonna *et al.*, (2014) |
|---|---|---|
| **SNPs** | 2,559 | 117 |
| **Pleiotropies** | 266 | 13 |
| | | Binomial test p-value = 0.44 |

**Supplementary Table 7 |** Binomial table showing no excess of pleiotropies in a set of genes under recent positive selection (Colonna et al., 2014).

## Ancient selection

Next, we investigated whether genes harboring pleiotropies presented any particular pattern of ancestral evolution. Since most neutrality tests that are used to detect ancient episodes of selection are gene-based, we associated the 266 pleiotropies detected in our study to each gene or genes where the causal variant might lie. In particular, for each SNP within a pleiotropy we selected all genes in the GWAS Catalog described as "*Reported Genes*" or "*Mapped Genes*". In those cases where the "*Reported genes*" and the "*Mapped Genes*" were different or more than a unique gene was reported in both SNPs, we conservatively considered all of them. This added up to a set of 361 genes. To study rates of protein evolution, dN/dS values for these genes were extracted from the Ensembl Biomart Genes 78 database[83] considering three pairwise species comparisons: human-chimpanzee, human-macaque and human-mouse. Only one-to-one orthologous gene pairs were considered and no filters for sequence identity were applied. We performed the following comparisons:

(a) rates of protein evolution of the whole set of pleiotropic genes to the rest of the genome;

(b) rates of protein evolution values of the genes containing antagonistic early-late pleiotropies to the rest of the genome;

and,

(c) rates of protein evolution values of the genes containing antagonistic early-late pleiotropies to the rest of pleiotropic genes.

To do so we sampled 1,000 times, from the list of genes in the genome, sets of genes either as large as the number of genes in the complete set of all pleiotropies (a) or as large as the set of antagonistic early-late pleiotropies (b, c). For each resampling, the average dN/dS was calculated. The 1,000 resampled averages representing a null distribution were then compared to the average dN/dS of the tested pleiotropic genes. After this procedure, none of the three comparisons yielded clearly significant results.

Besides evaluating whether pleiotropies significantly overlap with genes or genomic regions under recent and ancient positive selection, we also tested whether any of the two sets of ageing genes[73,74] used above showed enrichment for positive adaptation. Again, none of the analysis displayed significant results.

## Section 7: Evolutionary analysis of antagonistic pleiotropy genes: other particular cases.

The fact that the genes or regions reported by selection scans do not present excess overlap with pleiotropies does not exclude the possibility of detecting and studying particular instances of natural selection. To ensure that any given case of age-related antagonistic pleiotropy is relevant to human adaption in relation to ageing and disease, it should fulfil certain conditions: (1) the early-onset condition should have had a significant impact on fitness; (2) display some signature of positive selection; and, (3) its pleiotropic effects should ideally be experimentally validated through animal models, expression levels or similar. After exploring signatures of adaptation on the 1000 Genomes Selection Browser[84,85] for all the antagonistic early-late pleiotropies identified in this study, we found some particular cases of positive selection partially fulfilling these conditions (Supplementary Table 8).

| EARLY ONSET | | | | LATE ONSET | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SNP | Disease | Risk Allele | Anc. All*. | SNP | Disease | Risk Allele | Anc. All.* | Derived allele = early protect | $r^2$ | Chr | Gene |
| rs1295686 | Asthma | T | T | rs20541 | Psoriasis | G | G | YES | 1 | 5 | *IL13* |
| rs1295686 | Atopic dermatitis | T | T | rs20541 | Psoriasis | G | G | YES | 1 | 5 | *IL13* |
| rs20541 | Hodgkin's lymphoma | A | G | rs20541 | Psoriasis | G | G | NO | 1 | 5 | *IL13* |
| rs2157719 | Glioma | C | C | rs523096 | Glaucoma | A | A | YES | 0.84 | 9 | *CDKN2A* |
| rs2157719 | Glioma | C | C | rs564398 | Type 2 diabetes | T | T | YES | 0.95 | 9 | *CDKN2A* |
| rs2157719 | Glioma | C | C | rs7865618 | Coronary heart disease | A | A | YES | 1 | 9 | *CDKN2A* |
| rs2157719 | Glioma | C | C | rs1412829 | Nasopharynx carcinoma | A | A | YES | 0.95 | 9 | *CDKN2A* |
| rs11755724 | Multiple sclerosis | A | A | rs11755724 | Age-related macular degen. | G | A | YES | 1 | 6 | *RREB1* |

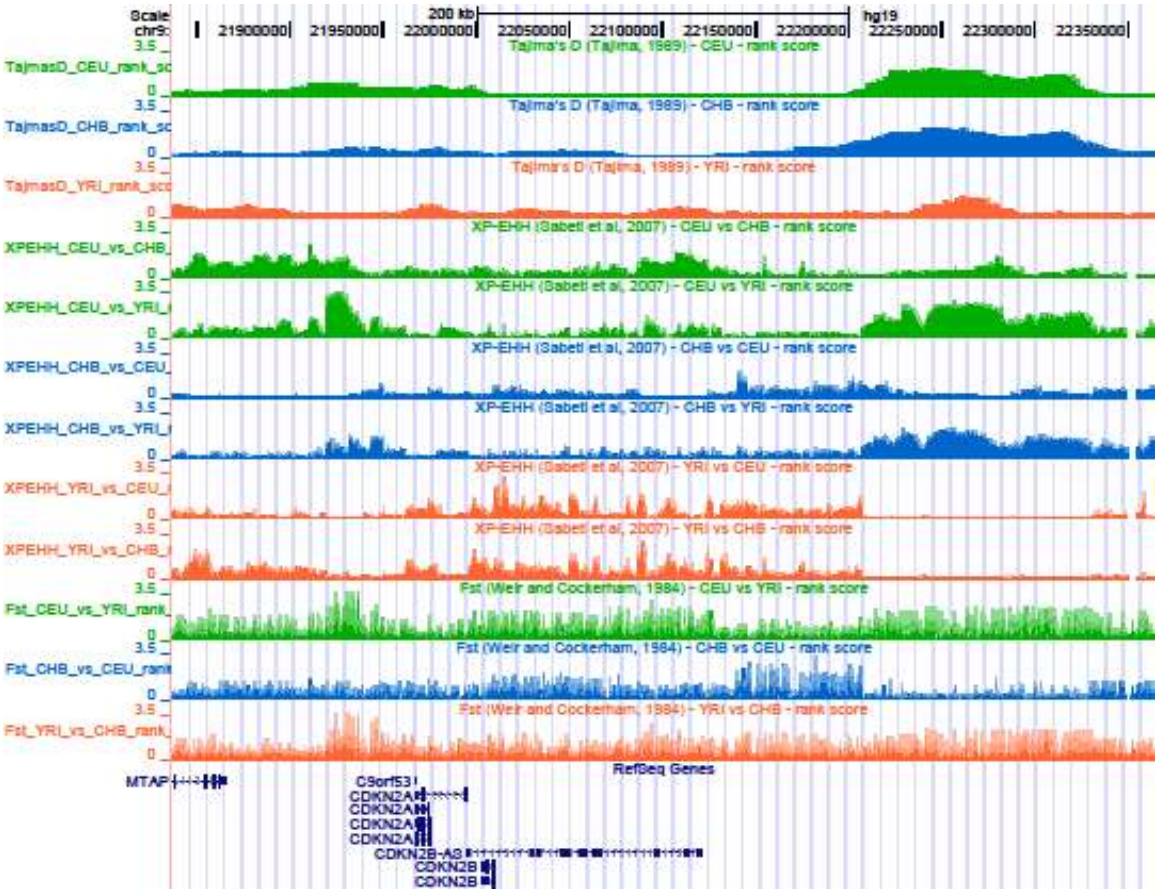
**Supplementary Table 8:** Linked markers identified as antagonistic early-late pleiotropies involving genes with signatures of positive selection on the 1000 Genomes Selection Browser (Pybus et al., 2014). **Anc. All=Ancestral Allele*.

### *CDKN2A*

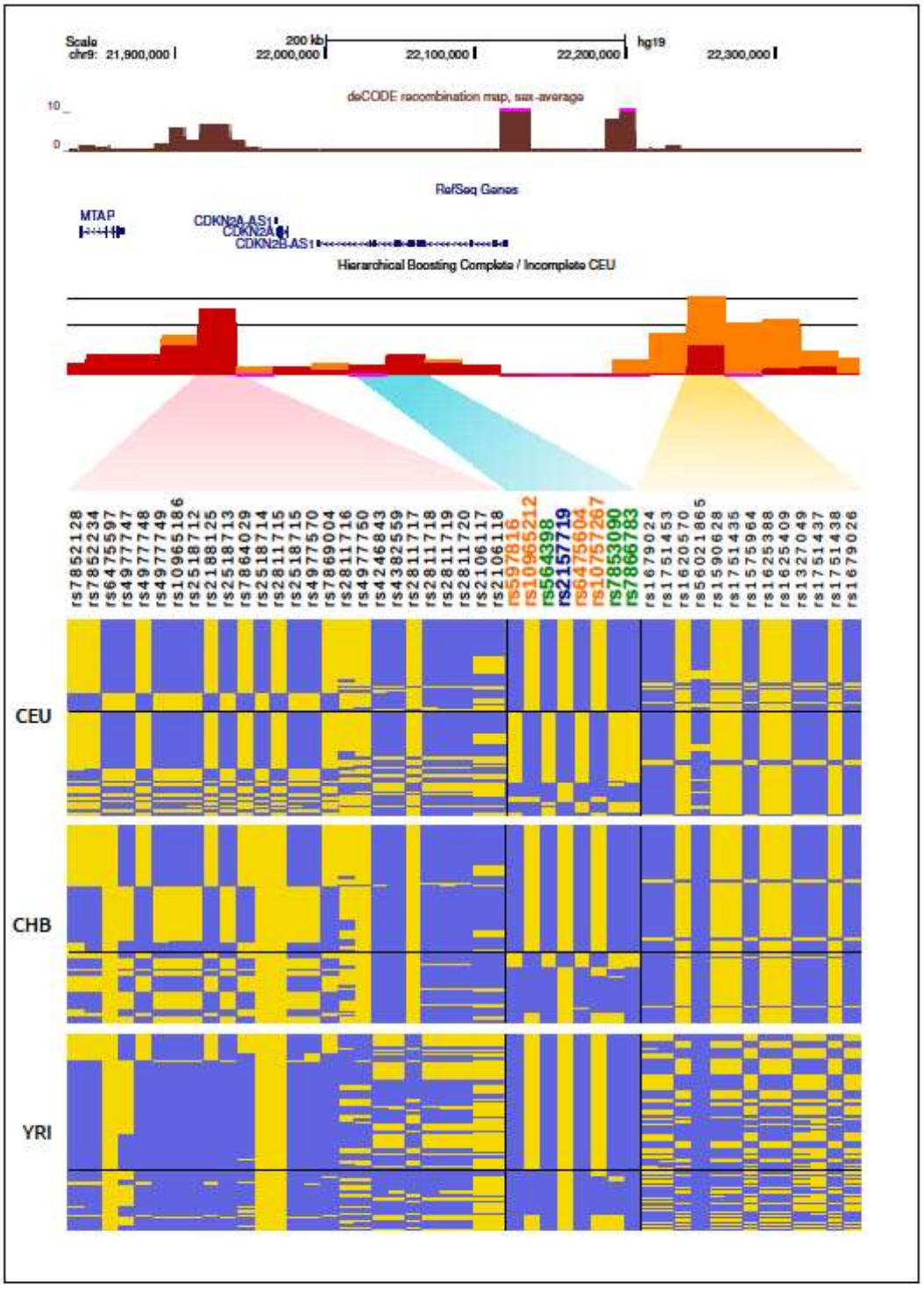
The case of *CDKN2A* (also called *p16*$^{INK4a}$) is particularly interesting. This gene, that has a clear role in ageing[73,86–88], encodes for a Cyclin-Dependent Kinase Inhibitor acting as tumor suppressor inducing cell cycle arrest. When muscular tissue is damaged, muscle stem cells exit quiescence and induce regeneration. However, this capability declines with age in parallel with an increase of the expression level of *CDKN2A* that induces sarcopenia - the decline in muscular regenerative capacities that develops even with “healthy” ageing[87]. Recent work has shown that wild-type ageing mice are able to revert to the original muscle regenerative functions after silencing *CDKN2A*[73]. This gene’s further phenotypic effects implicate that this locus has pleiotropic senescence significance. For instance, elevated expression levels in adipose tissue, skeletal muscle and in the eyes correlate with loss of subcutaneous adipose tissue, onset of spine deformities, skeletal muscle atrophy, arterial stiffness and cataracts. After the onset of these symptoms in a progeric mouse model, silencing of *CDKN2A* expression increased lifespan, and ageing conditions were reverted[87]. In contrast, knockout newborn mice for *CDKN2A* die at early age from tumors[87]. Consistent with these linked pleiotropic phenotypes, we found *CDKN2A* to be implicated in four antagonistic pleiotropies involving five SNPs and five diseases (Supplementary Table 8). The T allele (CEU freq = 54%) of the intronic SNP rs2157719 is protective for glioma, while increasing the risk of four different age-related conditions: type 2 diabetes, glaucoma, coronary heart disease and nasopharyngeal cancer. Interestingly, symptoms of these diseases are mimicked by the aforementioned elevated expression levels of this gene in adipose tissue, skeletal muscle and eye and by the tumors in young knockout

mice. Moreover, in glioma cases, G/G homozygotes have lower *CDKN2A* expression[89]. We hypothesize that protection from glioma, a relatively frequent, early onset, and often fatal cancer (<15% survival in 10 years[90]) is favored at the price of increased risk for four late-onset conditions.

To evaluate the selective footprint linked to this locus we began by exploring the data provided by the 1000 Genomes Human Selection Browser[84]. We observed high population differentiation between Europeans and Yorubans along the flanking genomic regions of *CDKN2A* (Supplementary Figure 2). In agreement with this observation, a hierarchical boosting method[85] revealed that both flanking regions present significant scores for selective sweeps, unique to Europeans. The downstream sweep was predicted to be complete, while the upstream sweep appears as incomplete (Supplementary Figure 3).




**Supplementary Figure 2 | Signatures of positive selection around the *CDKN2A* and *CDKN2B-2AS1* (or *ANRIL*) *genes*.** Tracks for Tajima's D, $F_{ST}$ and XP-EHH as extracted from the 1000 Genomes Selection Browser v1.0 (http://hsb.upf.edu/)[84] for the CEU, YRI and CHB populations. Most peaks observed for the three statistics in the CEU population overlap with the complete and incomplete sweep signals shown in Supplementary Fig. 3. When comparing CEU *vs.* YRI, XP-EHH reaches significant values between ~21,925,000 and ~21,937,000 bp, although is not appreciated due to zoom level. $F_{st}$ reached also significant values for SNPs within this same region. In line with this, we detected a non-significant but relatively high iHS value for the ancestral C allele at rs2157719 (iHS=1.184), consistent with the more recent and incomplete sweep favoring the ancestral C allele (data not shown).



**Supplementary Figure 3 | Hierarchical boosting scores for selective sweeps and haplotype patterns around the *CDKN2A* locus.** UCSC genome tracks for "*Decode recombination map, sex-average*", "*RefSeq genes"* and *"Hierarchical Boosting"* were obtained from the 1000 Genomes Selection browser v.1.0 (http://hsb.upf.edu/)[85]. In the "*Hierarchical Boosting*" track, the first black line from the bottom corresponds to the significance threshold for complete sweep (shown in red), while the second black line indicates the significance threshold for incomplete sweep (in orange). For those SNPs mapping in genomic regions with significant signals for selective sweeps as well as for the eight putative functional SNPs identified along the region (zoomed regions in the figure), haplotypes were extracted with Haploview[91] for the three main populations (CEU, CHB, YRI) of the 1000 Genomes Project. For each SNP, ancestral (blue) or derived (yellow) state was assessed by comparison with the chimpanzee, using data from 1000 Genomes Project, phase 1[92]. In each population, the 8-SNP central haplotype protective from glioma is presented above a black line. Colours in the SNP identifiers indicate different functional information when available: GWAS tag-SNP (blue), probably functional SNPs according to CADD[93] (orange) and exonic variant at *CDKN2B-2AS1* (or *ANRIL*) (green).

*In silico* functional analysis along the whole ~350 Kb region comprising the signatures of positive selection in Europeans led to the identification of a set of seven putative functional SNPs along a 35 Kb region in almost complete LD to rs2157719 (Supplementary Table 9). Four of these SNPs are among the top 10%, or even the top 1%, of the most harmful substitutions in the human genome according to CADD[93], while the remaining three SNPs map in exonic regions of a long non-coding RNA, known to modulate expression of *CDKN2A*, as detailed below. Together with rs2157719, these seven, presumed functional, SNPs define two main haplotypes in Europeans, which can be labeled as *risky* and *protective* from the perspective of the early-onset condition associated to the allelic variation at rs2157719 (Supplementary Figure 3). In contrast, in Africans and Asians, the glioma protective T-allele at rs2157719 is fixed or virtually fixed and a *protective* 8-SNP haplotype is present at a very high frequency (~69%).

| Variant ID | Position (hg19) | MAF CEU | MAF YRI | Feature | Predicted Functional | CADD PHRED Score | Fst | percentil Fst |
|---|---|---|---|---|---|---|---|---|
| rs597816 | 22021172 | 0,417647 | 0 | Intronic | YES | **13,14** | 0,41 | 0,985 |
| rs10965212 | 22023795 | 0,494118 | 0,21 | Intronic | YES | **10,21** | 0,16 | 0,896 |
| rs564398 | 22029547 | 0,435294 | 0 | Exonic | NO | <10 | 0,43 | 0,987 |
| rs2157719 | 22033366 | 0,447059 | 0 | Intronic (TAG) | NO | <10 | 0,44 | 0,988 |
| rs6475604 | 22052734 | 0,435294 | 0 | Intronic | YES | **20,3** | 0,43 | 0,987 |
| rs10757267 | 22052810 | 0,494118 | 0,21 | Intronic | YES | **16,74** | 0,15 | 0,887 |
| rs7853090 | 22056295 | 0,441176 | 0,10 | Exonic | NO | <10 | 0,25 | 0,948 |
| rs7866783 | 22056359 | 0,435294 | 0 | Exonic | NO | <10 | 0,43 | 0,987 |

**Supplementary Table 9:** Genomic features for the 8 SNPs defining a core haplotype inside CDKN2A-ANRIL locus.

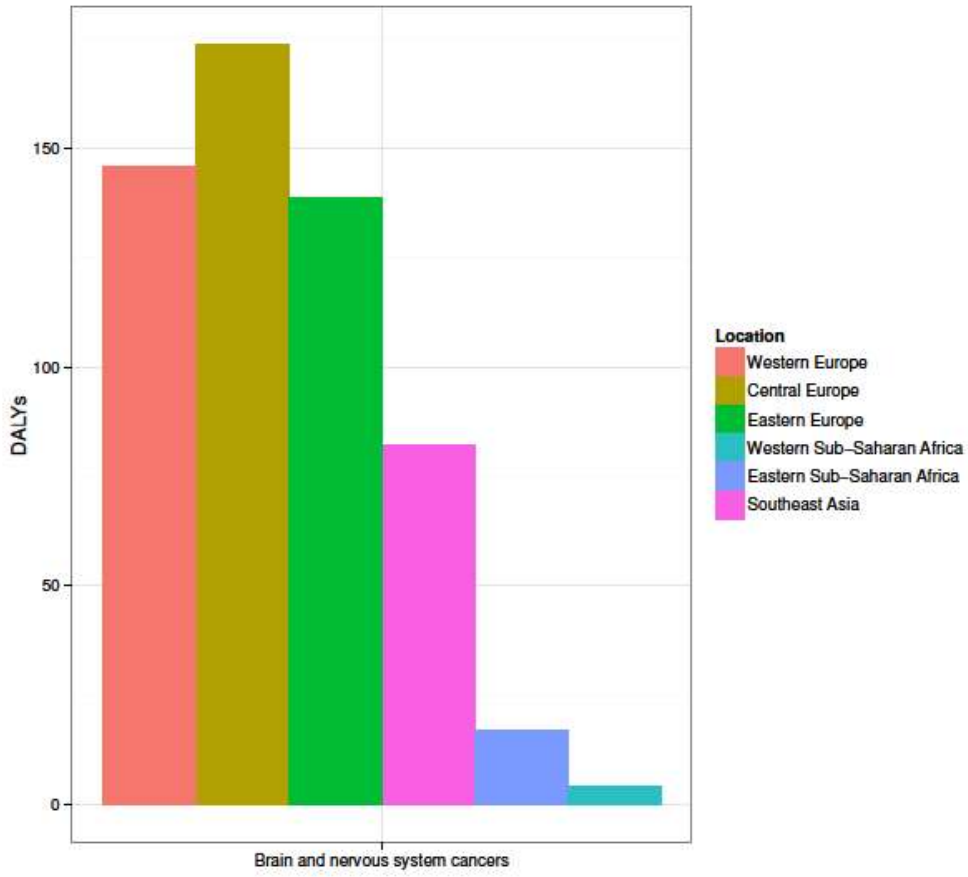
While the complex pattern of recombination does not allow detecting clear departures from neutrality along the 35 kb (Supplementary Figure 3), five of the eight SNPs (including rs2157719) fell over the 98$^{th}$ percentile (*p*-value = 0.02) of an empirical distribution of $F_{ST}$ values between Europeans and Yorubans (Supplementary Table 9). Moreover, the remaining three were found around 88$^{th}$ percentile (*p*-value = 0.12). As mentioned above, these 8 SNPs map within a long non-coding RNA (lncRNA), called *CDKN2B-AS1*, also known as *ANRIL* (for Antisense Non-coding RNA in the *INK4* Locus), which has interesting features regarding the pleiotropies and related diseases linked to this region.

This lncRNA was discovered in 2007[94] and maps ~300 bp upstream of the transcription start site of *CDKN2A*. It is also a hotspot for GWAS hits with a role in cellular ageing[95,96] and known to regulate the expression of three tumour suppressor genes *CDKN2A*, *CDKN2B* and *ARF*[95,97]. All three of these genes are involved in cellular senescence processes by inhibiting the cell cycle progression from phase $G_1$ to S, under cellular stress conditions[98]. Moreover, these genes are also involved in a wide spectrum of complex diseases,

including several forms of cancer, type 2 diabetes, periodontitis, stroke, coronary heart disease, and glaucoma among others[96,99]. The first implication of this lncRNA molecule with disease came from the observation that patients suffering neural system tumors have an increased risk of developing malignant melanoma[94]. A genetic examination of these patients showed that most of them were carrying deletions in this region, leading to a possible misregulation of the *CDKN2A/B* tumour suppressors. Association studies also demonstrated the relevance of *ANRIL* in many human age-related pathologies[99] and different biological processes related to cellular senescence[98]. For instance, the misregulation of the expression mediated by variation on *ANRIL* gene could alter the telomere shortening operated by the tumor suppressor genes within the *INK4/ARF* locus and provoke the degeneration of the replicative system of several cell classes with ageing[98]. Moreover, deterioration of β-cells in pancreas through this mechanism has been demonstrated to contribute to type 2 diabetes onset[100]. The *INK4/ARF* locus regulated through *ANRIL* expression has been investigated to explain the link with the deleterious phenotypes in which it is involved[95]. Most genetic variants uncovered by GWAS in the *INK4/ARF* gene cluster are located in non-coding regions, suggesting a role in gene expression[89,97,99]. Certain expression studies[95] have confirmed the co-regulation of genes in the 9p21.3 region: *p15/CDKN2B*, *p16/CDKN2A*, *p14/ARF* and *ANRIL*. Out of these four genes, *ANRIL* showed the strongest and most solid associations with the diseases associated to the region[95]. Cunnington *et al*.[95] demonstrated that several SNPs were affecting expression levels of *ANRIL*, suggesting that multiple variants, and not a single one, were involved in disease. In particular, rs564398 and rs1063192 were reported as direct

expression modulators associated with coronary heart disease and glioma, respectively. While rs564398 is one of the 8 functional SNPs considered here, rs1063192 is in almost perfect LD ($r^2 = 0.98$) with rs2157719 (GWAS tag SNP for early onset glioma), with both SNPs presenting their risk alleles in phase. Cunnington *et al.*,[95] also examined the expression levels of *ANRIL* depending on the presence of particular risk alleles for the diseases mapping in the 9p21.3 region. The risk allele T at rs1063192 for coronary heart disease and type 2 diabetes was associated with reduced *ANRIL* expression levels. However, in gliomas, the risk allele C at rs1063192 highly correlated with increased *ANRIL* expression levels. Further support to the antagonistic pleiotropies described here for this region is provided by the fact that expression levels for *ANRIL* are inversely correlated between glioma – diabetes and glioma – coronary heart disease, and that these diseases are involved in antagonistic pleiotropies concerning *ANRIL*.

The differential impacts of brain neoplasms between Sub-Saharan Africa, Europe and Asia agree with this pattern of population differentiation, since childhood brain tumors are far more prevalent in Europe (Supplementary Figure 4), where the protective variant only shows intermediate frequency.




**Supplementary Figure 4 | Disability Adjusted Life Years (DALYs) for brain and nervous system cancers in three world regions.** DALYs are units used by the World Health Organisation as a health statistic. One DALY can be thought of as one lost year of "healthy" life. DALYs are provided in rate/100,000, that is, out of 100,000 DALYs lost in the population how many of correspond to brain cancers in each region. Data was gathered from the Institute for Health Metrics and Evaluation, downloaded from http://www.healthdata.org/gbd/data

### *RREB1*

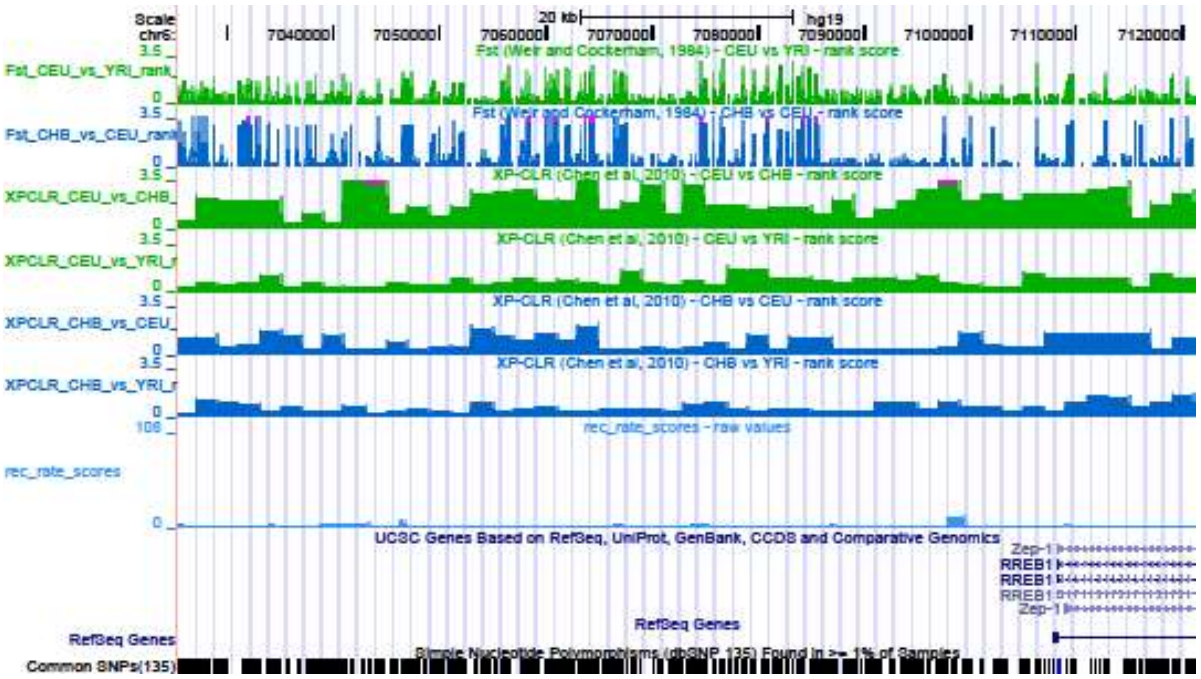
The intronic SNP rs11755724 in the *RREB1* gene was antagonistically associated with multiple sclerosis and age-related macular degeneration. In particular, the derived allele (G allele) is protective for multiple sclerosis but contributes risk for age-related macular degeneration (Supplementary Table 8). This allele is most frequent in the European population (63.3% HapMap CEU; 71% HapMap TSI), while being almost absent in Asians and 17% in

Africans, a pattern that matches the prevalence of multiple sclerosis, which in the two latter populations is around 40 times less prevalent than in Europeans[101–103] (Supplementary Table 10).

| Location | Sex | DALYs lost | Lower Bound | Upper Bound |
|---|---|---|---|---|
| **East Asia** | Male | 1427 | 1073 | 1865 |
| **East Asia** | Female | 2685 | 2045 | 3350 |
| **East Asia** | Both | 2034 | 1564 | 2530 |
| **Western Europe** | Male | 55610 | 38230 | 72570 |
| **Western Europe** | Female | 96710 | 75510 | 117510 |
| **Western Europe** | Both | 76570 | 60780 | 91960 |

**Supplementary Table 10 |** DALYs lost to multiple sclerosis in East Asia compared to Europe. DALYs lost to multiple sclerosis in East Asia vs. Western Europe. DALYs are expressed in DALYs/100,000, meaning that out of 100,000 DALYs lost in the population the value in the table corresponds to those from multiple sclerosis. Data collected from http://vizhub.healthdata.org/gbd-compare/

Notably, we observe a significant signal for positive selection in the upstream region of this gene, especially when comparing populations of European and Asian origin (XP-CLR in *CEU vs. CHB* and $F_{ST}$ in *CEU vs. CHB*) (Supplementary Figure 5).



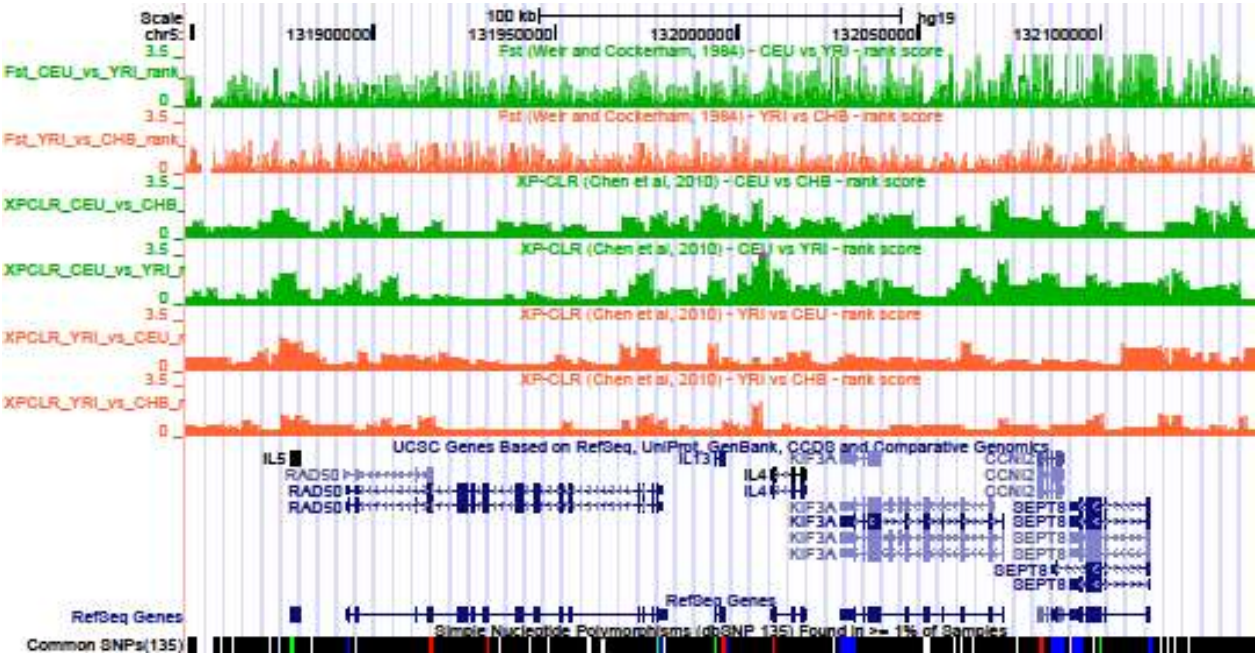

**Supplementary Figure 5 | Signatures of positive selection around the *RREB1* gene region.** Tracks for $F_{ST}$ and XP-CLR between pairs of populations as extracted from the 1000 Genomes Selection Browser v1.0 (http://hsb.upf.edu/)[84]. In the context of overall higher values, departures from neutrality are detected (in pink, inside the box) for both $F_{ST}$ and XP-CLR when comparing CEU *vs.* CHB populations.

It has been shown that multiple sclerosis patients have uric acid deficiency, but its increase has as well been associated with age-related macular degeneration[104,105]. We searched for further associations of this gene in the whole GWAS Catalog database, finding SNP rs675209 mapping also in *RREB1* and associated with modulation of uric acid levels in blood[106]. Both SNPs present the strongest LD in CEU populations ($r^2$ = 0.576; D'=0.94), where frequencies of rs11755724 are higher (recall that rs11755724-G is fixed in Asia and LD between both SNPs is $r^2 < 0.1$ in Africa, where its frequency is 0.17). Given the role that uric acid plays in fighting ageing and cancer by acting as a highly efficient scavenger of free radicals[107], it is tempting to hypothesize that this pleiotropy contributes to the risk for age-related macular degeneration as a

trade-off for protection against multiple sclerosis, mainly in European populations where this disease is highly prevalent.

***IL13***

Our results showed three antagonistic early-late pleiotropies linked to this gene (Supplementary Table 8). In particular, the protective alleles of three SNPs associated to early-onset conditions (namely, rs1295686 for Hodgkin's lymphoma, rs1295686 for atopic dermatitis and rs20541 for asthma) are included in the same haplotype with perfect LD, or even coincide, with the risk allele for psoriasis, a late onset condition, in SNP rs20541. Significant signatures of selection were detected around the *IL13* gene in the XP-CLR test when comparing CEU with CHB, and CEU with YRI (Supplementary Figure 6). Moreover, this gene has also been pinpointed as a candidate of positive selection by Akey *et al.*[81]. These results are consistent with a previous report of positive selection in this gene that already suggested a relation between selection and disease[108]. Our results suggest that increased risk of psoriasis may come as a trade-off for decreased risk of the aforementioned early-onset diseases.




**Supplementary Figure 6 | Signatures of positive selection around the IL13 gene region.** Tracks for FST and XP-CLR between pairs of populations as extracted from the 1000 Genomes Selection Browser v1.0 (http://hsb.upf.edu/)[84]. Departures from neutrality can be detected (in pink, inside the box) for XP-CLR when comparing CEU vs. YRI.

Further indirect support for this idea comes from a recent study by Mathieson *et al.*[109], who suggested that recent positive selection has acted in a gene associated with celiac disease, *SLC22A4*, mapping ~300 kb downstream from *IL13*. They hypothesize that this gene, which codes for an ergothioneine transporter, was selected to compensate for diet deficiencies, even at the cost of celiac disease. Ergothioneine is an amino-acid found also at high levels in the skin, where it acts as a powerful antioxidant, scavenging hydroxyl radicals and hypochlorous acid[110]. It is striking, even if possibly casual, that all the pleiotropies mapping in *IL13* seem to be associated with skin-related conditions: either psoriasis or atopic dermatitis, or both.

## References


1. Vaupel, J. W. Biodemography of Human Ageing. *Nature* **464,** 536–42 (2010).
2. Kirkwood, T. B. Understanding the odd science of aging. *Cell* **120,** 437–47 (2005).
3. Flatt, T. Ageing: Diet and longevity in the balance. *Nature* **462,** 989–990 (2009).
4. Trindade, L. S. *et al.* A novel classification system for evolutionary aging theories. *Front. Genet.* **4,** 1–8 (2013).
5. López-Otín, C., Blasco, M. a, Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153,** 1194–217 (2013).
6. Jones, O. R. *et al.* Diversity of ageing across the tree of life. *Nature* 5–10 (2014). doi:10.1038/nature12789
7. Ricklefs, R. E. Evolutionary theories of aging: confirmation of a fundamental prediction, with implications for the genetic basis and evolution of life span. *Am. Nat.* **152,** 24–44 (1998).
8. Medawar, P. B. An Unsolved Problem of Biology. *Evolution in Health and Disease* 24 (1951). doi:10.1016/S0140-6736(00)99799-X
9. Williams, G. C. Pleiotropy, Natural Selection, and the Evolution of Senescence. *Evolution (N. Y).* **11,** 398–411 (1957).
10. Weismann, A. Essays Upon Heredity and Kindred Biological Problems. *Oxford Clarendon Press* **1,** 1–5 (1891).
11. Kirkwood, T. B. Evolution of ageing. *Nature* **170,** 201–4 (1977).
12. Stearns, S. C. *The Evolution of Life Histories*. (Oxford University Press, 1992).
13. Baudisch, A. & Vaupel, J. W. Evolution. Getting to the root of aging. *Science* **338,** 618–9 (2012).
14. Baudisch, A. *Inevitable Aging? Contributions to Evolutionary-Demographic Theory*. (Springer-Verlag Berlin Heidelberg, 2008).
15. Baudisch, A. Birds Do It, Bees Do It, We Do It: Contributions of Theoretical Modelling to Understanding the Shape of Ageing across the Tree of Life. *Gerontology* **58,** 481–489 (2012).
16. Martins, A. C. R. Change and aging senescence as an adaptation. *PLoS One* **6,** (2011).
17. Fisher, R. a. The Genetical Theory of Natural Selection. *Genetics* **154,** 272 (1930).
18. Medawar, P. B. Un unsolved problem of biology. in *An Inaugural Lecture Delivered at University College, London* (University College London, 1951).
19. Hamilton, W. D. The moulding of senescence by natural selection. *J. Theor. Biol.* **12,** 12–45 (1966).
20. Charlesworth, B. Fisher, Medawar, Hamilton and the evolution of aging. *Genetics* **156,** 927–931 (2000).
21. Charlesworth, B. *Evolution in Age-Structured Populations*. (Cambridge University Press, 1994).
22. Moorad, J. A. & Hall, D. W. Age-dependent mutational effects curtail the evolution of senescence by antagonistic pleiotropy. *J. Evol. Biol.* **22,** 2409–2419 (2009).
23. Rose, M. R. Laboratory Evolution of Postponed Senescence in

Drosophila melanogaster. *Evolution (N. Y).* **38,** 1004–1010 (1984).
24. Partridge, L., Prowse, N. & Pignatelli, P. Another set of responses and correlated responses to selection on age at reproduction in Drosophila melanogaster. *Proc. R. Soc. B Biol. Sci.* **266,** 255–261 (1999).
25. Stearns, S. C., Ackermann, M., Doebli, M. & Kaiser, M. Experimental evolution of aging, growth, and reproduction in fruit flies. *Proc. Natl. Acad. Sci. USA* **97,** 3309–3313 (2000).
26. Koopman, J. J. E., Wensink, M. J., Rozing, M. P., van Bodegom, D. & Westendorp, R. G. J. Intrinsic and extrinsic mortality reunited. *Experimental Gerontology* **67,** 48–53 (2015).
27. Bryant, M. J. & Reznick, D. Comparative studies of senescence in natural populations of guppies. *Am. Nat.* **163,** 55–68 (2004).
28. Charmantier, A., Perrins, C., McCleery, R. H. & Sheldon, B. C. Quantitative genetics of age at reproduction in wild swans: support for antagonistic pleiotropy models of senescence. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 6587–92 (2006).
29. Williams, P. D. & Day, T. Antagonistic pleiotropy, mortality source interactions, and the evolutionary theory of senescence. *Evolution (N. Y).* **57,** 1478–88 (2003).
30. Morbey, Y. E., Brassil, C. E. & Hendry, A. P. Rapid senescence in pacific salmon. *Am. Nat.* **166,** 556–568 (2005).
31. Miller, R. a, Harper, J. M., Dysko, R. C., Durkee, S. J. & Austad, S. N. Longer life spans and delayed maturation in wild-derived mice. *Exp. Biol. Med. (Maywood).* **227,** 500–508 (2002).
32. Reznick, D. N., Bryant, M. J., Roff, D., Ghalambor, C. K. & Ghalambor, D. E. Effect of extrinsic mortality on the evolution of senescence in guppies. *Nature* **431,** 1095–1099 (2004).
33. Williams, P. D., Day, T., Fletcher, Q. & Rowe, L. The shaping of senescence in the wild. *Trends Ecol. Evol.* **21,** 458–463 (2006).
34. Parker, S. L., Tong, T., Bolden, S. & Wingo, P. a. Cancer statistics, 1997. *CA. Cancer J. Clin.* **47,** 5–27 (1997).
35. Heron, M. *et al. Deaths: final data for 2006. National vital statistics reports : from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* **57,** (2009).
36. Finch, C. E. *Longevity, Senescence, and the Genome*. (Academic Press, 1990).
37. Finch, C. E. *The Biology of Human Longevity: Inflammation, Nutrition, and Aging in the Evolution of Lifespans*. (2007).
38. Vaupel, J. W. *et al.* Biodemographic Trajectories of Longevity. *Science (80-. ).* **280,** 855–860 (1998).
39. Kirkwood, T. B. & Holliday, R. The evolution of ageing and longevity. *Proc. R. Soc. Lond. B. Biol. Sci.* **205,** 531–546 (1979).
40. Dańko, M. J., Kozłowski, J. & Schaible, R. Unraveling the non-senescence phenomenon in Hydra. *J. Theor. Biol.* **382,** 137–149 (2015).
41. Schaible, R. *et al.* Constant mortality and fertility over age in Hydra. *Proc Natl Acad Sci U S A* **112,** 15701–15706 (2015).
42. Martínez, D. E. Mortality patterns suggest lack of senescence in hydra. *Exp. Gerontol.* **33,** 217–225 (1998).
43. Gavrilov, L. a & Gavrilova, N. S. Evolutionary theories of aging and longevity. *ScientificWorldJournal.* **2,** 339–356 (2002).

44. Kirkwood, T. B. Understanding ageing from an evolutionary perspective. in *Journal of Internal Medicine* **263,** 117–127 (2008).
45. Vaupel, J. W., Baudisch, A., Dölling, M., Roach, D. A. & Gampe, J. The case for negative senescence. *Theor. Popul. Biol.* **65,** 339–351 (2004).
46. Curtsinger, J. W., Fukui, H. H., Townsend, D. R. & Vaupel, J. W. Demography of genotypes: failure of the limited life-span paradigm in Drosophila melanogaster. *Science (80-. ).* **258,** 461–3 (1992).
47. Carey, J. R., Liedo, P., Orozco, D. & Vaupel, J. W. Slowing of mortality rates at older ages in large medfly cohorts. *Science* **258,** 457–461 (1992).
48. Rose, M. R. & Charlesworth, B. A test of evolutionary theories of senescence. *Nature* **287,** 141–142 (1980).
49. Charlesworth, B. Patterns of Age-specific Means and Genetic Variances of Mortality Rates Predicted by the Mutation-Accumulation Theory of Ageing. *J. Theor. Biol.* **210,** 47–65 (2001).
50. Borash, D. J., Rose, M. R. & Mueller, L. D. Mutation accumulation affects male virility in Drosophila selected for later reproduction. *Physiol. Biochem. Zool.* **80,** 461–472 (2007).
51. Baudisch, A. Hamilton's indicators of the force of selection. *Proc. Natl. Acad. Sci.* **102,** (2005).
52. Rose, M. R., Rauser, C. L., Benford, G., Matos, M. & Mueller, L. D. Hamilton's forces of natural selection after forty years. *Evolution (N. Y).* **61,** 1265–1276 (2007).
53. Mueller, L. D. & Rose, M. R. Evolutionary theory predicts late-life mortality plateaus. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 15249–15253 (1996).
54. Rauser, C. L., Mueller, L. D. & Rose, M. R. The evolution of late life. *Ageing Res. Rev.* **5,** 14–32 (2006).
55. Finch, C. E. & Kirkwood, T. B. *Chance, Development, and Aging*. (Oxford University Press, 2000).
56. Brooks, A., Lithgow, G. J. & Johnson, T. E. Mortality rates in a genetically heterogeneous population of Caenorhabditis elegans. *Science (80-. ).* **263,** 668–671 (1994).
57. Wu, D., Rea, S. L., Yashin, A. I. & Johnson, T. E. Visualizing hidden heterogeneity in isogenic populations of C. elegans. *Exp. Gerontol.* **41,** 261–270 (2006).
58. Barzilai, N., Atzmon, G., Derby, C. A., Bauman, J. M. & Lipton, R. B. A genotype of exceptional longevity is associated with preservation of cognitive function. *Neurology* **67,** 2170–2175 (2006).
59. Erikson, G. A. *et al.* Whole-Genome Sequencing of a Healthy Aging Cohort. *Cell* **165,** 1002–1011 (2016).
60. Carter, A. J. R. & Nguyen, A. Q. Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. *BMC Med. Genet.* **12,** 160 (2011).
61. Albin, R. L. Antagonistic pleiotropy, mutation accumulation, and human genetic disease. *Genetica* **91,** 279–286 (1993).
62. Sørensen, S. a, Fenger, K. & Olsen, J. H. Significantly lower incidence of cancer among patients with Huntington disease: An apoptotic effect of an expanded polyglutamine tract? *Cancer* **86,** 1342–6 (1999).
63. Several. A Novel Gene Containing a Trinucleotide That Is Expanded and Unstable on Huntington ' s Disease Chromosomes. *Cell* **72,** 971–983 (1993).

64. Shokeir, M. H. Investigations on Huntington's disease in the Canadian Prairies. I. Prevalence. *Clin. Genet.* **7,** 345–348 (1975).
65. Nesse, R. M. On the difficulty of defining disease: a Darwinian perspective. *Med. Health Care. Philos.* **4,** 37–46 (2001).
66. Eskenazi, B. R., Wilson-Rich, N. S. & Starks, P. T. A Darwinian approach to Huntington's disease: Subtle health benefits of a neurological disorder. *Med. Hypotheses* **69,** 1183–1189 (2007).
67. Charlesworth, B. *Evolution in age structured populations*. *CAMBRIDGE STUDIESIN MATHEMATICALBIOLOGY131994ixiii1306illustr* (Cambridge Univ. Press. Cambridge, 1980). doi:10.1016/0531-5565(95)90013-6
68. Adams, P. C. & Barton, J. C. Haemochromatosis. *Lancet* **370,** 1855–1860 (2007).
69. Distante, S. *et al.* The origin and spread of the HFE-C282Y haemochromatosis mutation. *Hum. Genet.* **115,** 269–279 (2004).
70. Weinberg, E. D. Survival advantage of the hemochromatosis C282Y mutation. *Perspect. Biol. Med.* **51,** 98–102 (2008).
71. Whitlock, E. P., Garlitz, B. A., Harris, E. L., Beil, T. L. & Smith, P. R. Screening for hereditary hemochromatosis: a systematic review for the U.S. Preventive Services Task Force. *Ann. Intern. Med.* **145,** 209–23 (2006).
72. Weinberg, E. D. Iron loading in humans: a risk factor for enhanced morbidity and mortality. *J. Nutr. Environ. Med.* **16,** 43–51 (2007).
73. Sousa-Victor, P. *et al.* Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature* **506,** 316–21 (2014).
74. de Magalhães, J. P., Curado, J. & Church, G. M. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics* **25,** 875–881 (2009).
75. Harries, L. W. *et al.* Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. *Aging Cell* **10,** 868–878 (2011).
76. de Magalhães, J. P. *et al.* The Human Ageing Genomic Resources: Online databases and tools for biogerontologists. *Aging Cell* **8,** 65–72 (2009).
77. Heyn, H. *et al.* Distinct DNA methylomes of newborns and centenarians. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 10522–10527 (2012).
78. Raj, T. *et al.* Common Risk Alleles for Inflammatory Diseases Are Targets of Recent Positive Selection. *Am. J. Hum. Genet.* **92,** 517–529 (2013).
79. Dulai, K. S., Von Dornum, M., Mollon, J. D. & Hunt, D. M. The evolution of trichromatic color vision by opsin gene duplication in new world and old world primates. *Genome Res.* **9,** 629–638 (1999).
80. Hancock, A. M. *et al.* Adaptations to climate-mediated selective pressures in humans. *PLoS Genet.* **7,** e1001375 (2011).
81. Akey, J. M. Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Res.* **19,** 711–722 (2009).
82. Colonna, V. *et al.* Human genomic regions with exceptionally high levels of population differentiation identified from 911 whole-genome sequences. *Genome Biol.* **15,** R88 (2014).
83. Cunningham, F. *et al.* Ensembl 2015. *Nucleic Acids Res.* **43,** D662-9 (2014).
84. Pybus, M. *et al.* 1000 Genomes Selection Browser 1.0: A genome

browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Res.* **42,** (2014).
85. Pybus, M. *et al.* Hierarchical boosting: A machine-learning framework to detect and classify hard selective sweeps in human populations. *Bioinformatics* **31,** 3946–3952 (2015).
86. Baker, D. J. *et al.* Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. *Nature* **479,** 232–6 (2011).
87. Baker, D. J. *et al.* Opposing roles for p16Ink4a and p19Arf in senescence and ageing caused by BubR1 insufficiency. *Nat. Cell Biol.* **10,** 825–36 (2008).
88. Baker, D. J. *et al.* Naturally occurring p16 Ink4a -positive cells shorten healthy lifespan. *Nature* **530,** 184–189 (2016).
89. Shete, S. *et al.* Genome-wide association study identifies five susceptibility loci for glioma. *Nat. Genet.* **41,** 899–904 (2009).
90. Quaresma, M., Coleman, M. P. & Rachet, B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971–2011: a population-based study. *Lancet* **385,** 1206–1218 (2015).
91. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **21,** 263–265 (2005).
92. Abecasis, G. R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56–65 (2012).
93. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46,** 310–5 (2014).
94. Pasmant, E. *et al.* Characterization of a Germ-Line Deletion, Including the Entire INK4/ARF Locus, in a Melanoma-Neural System Tumor Family: Identification of ANRIL, an Antisense Noncoding RNA Whose Expression Coclusters with ARF. *Cancer Res.* **67,** 3963–3969 (2007).
95. Cunnington, M. S., Santibanez Koref, M., Mayosi, B. M., Burn, J. & Keavney, B. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. *PLoS Genet.* **6,** e1000899 (2010).
96. Pasmant, E., Sabbagh, A., Vidaud, M. & Bièche, I. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J.* **25,** 444–448 (2011).
97. Wrensch, M. *et al.* Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. *Nat. Genet.* **41,** 905–908 (2009).
98. Sharpless, N. E. & DePinho, R. A. How stem cells age and why this makes us grow old. *Nat. Rev. Mol. Cell Biol.* **8,** 703–13 (2007).
99. Jeck, W. R., Siebold, A. P. & Sharpless, N. E. Review: a meta-analysis of GWAS and age-associated diseases. *Aging Cell* **11,** 727–31 (2012).
100. Krishnamurthy, J. *et al.* p16INK4a induces an age-dependent decline in islet regenerative potential. *Nature* **443,** 453–457 (2006).
101. Murray, C. *et al.* GBD 2010: a multi-investigator collaboration for global comparative descriptive epidemiology. *Lancet* **380,** 2055–8 (2012).
102. Murray, C. J. L. *et al.* Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380,** 2197–223 (2012).

103. Global Burden of Disease.
104. Rosenbaum, J. T. Eyeing Macular Degeneration — A Few Inflammatory Remarks. *N. Engl. J. Med.* **367,** 768–770 (2012).
105. Mehryar, M., Farvardin, M., Hosseini, H. & Aslani, M. Potential role of uric acid in the molecular pathogenesis of age-related macular degeneration. *Med. Hypotheses* **66,** 793–795 (2006).
106. Köttgen, A. *et al.* Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat. Genet.* **45,** 145–54 (2013).
107. Ames, B. N., Cathcart, R., Schwiers, E. & Hochstein, P. Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: a hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6858–62 (1981).
108. Zhou, G. *et al.* Haplotype structure and evidence for positive selection at the human IL13 locus. *Mol. Biol. Evol.* **21,** 29–35 (2004).
109. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528,** 499–503 (2015).
110. Markova, N. G. *et al.* Skin cells and tissue are capable of using l-ergothioneine as an integral component of their antioxidant defense system. *Free Radic. Biol. Med.* **46,** 1168–1176 (2009).