Received: 19 August 2019 | Revised: 22 October 2019 | Accepted: 23 October 2019

DOI: 10.1096/fj.201902116R

HYPOTHESES

THE FASEB JOURNAL

Maximum reproductive lifespan correlates with *CD33rSIGLEC* gene number: Implications for NADPH oxidase-derived reactive oxygen species in aging

Naazneen Khan[1,2,3,4,5,6] | **Stuart K. Kim**[7] | **Pascal Gagneux**[1,2,3,4,5,6] | **Laura L. Dugan**[8,9] | **Ajit Varki**[1,2,3,4,5,6]

[1]Glycobiology Research and Training Center, UC San Diego, La Jolla, CA, USA

[2]Center for Academic Research and Training in Anthropogeny, UC San Diego, La Jolla, CA, USA

[3]Department of Medicine, UC San Diego, La Jolla, CA, USA

[4]Department of Pathology, UC San Diego, La Jolla, CA, USA

[5]Department of Anthropology, UC San Diego, La Jolla, CA, USA

[6]Department of Cellular & Molecular Medicine, UC San Diego, La Jolla, CA, USA

[7]Department of Developmental Biology, Stanford University Medical Center, Stanford, CA, USA

[8]Vanderbilt University Medical Center, Medicine-Geriatrics, Nashville, TN, USA

[8]Vanderbilt University Medical Center, Medicine-Geriatrics, Nashville, TN, USA

[9]VA Tennessee Valley Geriatric Research, Education and Clinical Center (GRECC), Nashville, TN, USA

Correspondence
Ajit Varki, UCSD School of Medicine, La Jolla, CA 92093-0687, USA.
Email: a1varki@health.ucsd.edu

Funding information
HHS | NIH | National Institute of General Medical Sciences (NIGMS), Grant/Award Number: R01GM32373

Abstract

Humans and orcas are among the very rare species that have a prolonged post-reproductive lifespan (PRLS), during which the aging process continues. Reactive oxygen species (ROS) derived from mitochondria and from the NADPH oxidase (NOX) enzymes of innate immune cells are known to contribute to aging, with the former thought to be dominant. CD33-related-Siglecs are immune receptors that recognize self-associated-molecular-patterns and modulate NOX-derived-ROS. We herewith demonstrate a strong correlation of lifespan with *CD33rSIGLEC* gene number in 26 species, independent of body weight or phylogeny. The correlation is stronger when considering total *CD33rSIGLEC* gene number rather than those encoding inhibitory and activating subsets, suggesting that lifetime balancing of ROS is important. Combining independent lines of evidence including the short half-life and spontaneous activation of neutrophils, we calculate that even without inter-current inflammation, a major source of lifetime ROS exposure may actually be neutrophil NOX-derived. However, genomes of human supercentenarians (>110 years) do not harbor a significantly higher number of functional *CD33rSIGLEC* genes. Instead, lifespan correlation with *CD33rSIGLEC* gene number was markedly strengthened by excluding the post-reproductive lifespan of humans and orcas ($R^2 = 0.83$; $P < .0001$). Thus, *CD33rSIGLEC* modulation of ROS likely contributes to maximum reproductive lifespan, but other unknown mechanisms could be important to PRLS.

KEYWORDS

CD33rSIGLEC, NADPH-oxidase, prolonged post-reproductive lifespan, reactive oxygen species

Abbreviations: *APOE4*, apolipoprotein e4; CGP, circulating granulocyte pool; FIC, Felsenstein's Independent Contrast; *FOXO3*, forkhead box O-3; *IGF-1*, insulin-like growth factor 1; ITAM, immunoreceptor tyrosine-based activatory motif; ITIMs, immunoreceptor tyrosine-based inhibitory motif; MGP, marginated granulocyte pool; NOX, NADPH oxidase; PGLS, phylogeny generalized least-Squares; PRLS, prolonged post-reproductive lifespan; ROS, reactive oxygen species; SAMPs, self-associated molecular patterns; SHP-1, Src homology domain 2-containing tyrosine phosphatase-1; Sia, sialic acid; Siglecs, sialic acid-binding Ig-like lectins; TGBP, total granulocyte blood pool; TSHZ3, Teashirt Zinc Finger Homeobox 3.

1 | INTRODUCTION

Aging is a complex biological process varying widely between species, being influenced by numerous interacting factors.[1-6] Logically, aging also impacts maximum lifespan, a number that also varies widely between species. One contributor to aging appears to be reactive oxygen species (ROS), originating from two major sources: mitochondria (a relatively constant source in most cell types)[7] and from NADPH oxidase (NOX) enzymes.[8-12] Prominent among the latter is phagolysosomal NOX2, a membrane bound enzyme that generates ROS primarily for pathogen killing,[13,14] and prominently expressed in innate immune cells of the myeloid lineage (predominantly neutrophils, which are typically thought of as more episodic sources of ROS, during inflammation).[9,10,15] While high levels of ROS cause organelle injury, DNA damage, protein misfolding, cell demise, promote aging, and low levels mediate important cell signaling pathways (such as mTOR or Wnt). Indeed, mildly increasing ROS promotes longevity in standard multicellular models of aging (such as *Caenorhabditis elegans*).[16,17] There is now compelling evidence that ROS contribute to aging in many organisms.[18] Notably, calorie restriction, perhaps the most robust method for extending lifespan across species, decreases ROS production during aging in lower organisms as well as in mammals.[19] Also, overexpressing antioxidants such as catalase in mice decreases ROS and extends lifespan.[20] Thus, a fine balance between production and turnover of ROS appears necessary to limit cell injury and extend lifespan.[21,22]

Mitochondria are generally considered the primary source of ROS involved in aging.[6] ROS produced by mitochondria are a byproduct of normal functioning of the electron transport chain and inhibitors of respiratory chain such as rotenone have been shown to suppress the ability of hormones and cytokines to produce ROS in various cell types.[2,23] While mitochondrial ROS can be physiologically regulated, it is estimated that less than 2% of mitochondrial O_2 consumption is converted to the ROS, superoxide. In contrast, the phagolysosomal NOX enzyme expressed by innate immune cells generates high levels of ROS for pathogen killing[13] and converts nearly 100% of consumed oxygen to superoxide. NOX is also a major source of superoxide in synaptosomes with minor contributions from synaptosomal mitochondria which lead to cognitive decline.[24] However, all cells in the body have mitochondria that are constantly producing ROS due to electron leakage. Therefore, it is generally assumed that mitochondria are the major source of total baseline and lifetime ROS production with episodic innate immune inflammation also contributing to lifetime damage and aging.[20,25-27] Moreover, we have recently shown that while neutrophils are indeed quiescent in the blood stream, they immediately begin to produce ROS when they are separated away from the erythrocyte and plasma glycoproteins,[28] which present a high density of sialic acid bearing self-associated molecular patterns (SAMPs).[29] Such separation from SAMPs would also occur in vivo when neutrophils naturally exit the bloodstream, and it is therefore likely that they emit ROS before undergoing apoptosis and clearance by tissue macrophages,[30] even in the absence of active inflammation. These kinds of data support the hypotheses of Finch and colleagues that somatic cell resilience to agents such as ROS determines species differences in longevity during repeated infections and traumatic injuries in the natural environment.[31] Complementary to this hypothesis, species-specific control of ROS may further contribute to differences in lifespan. Of course, other aspects of recurrent and/or chronic innate immune activation may also contribute to "inflammaging."

Control of neutrophil ROS production is at least partly mediated by CD33-related subfamily of Siglecs (Sialic acid-binding Ig-like lectins)[28,32-39] encoded in the *CD33rSIGLEC* gene cluster, and referred to as CD33rSiglecs.[40-44] Consistent with the phylogenetic distribution of sialic acids in the deuterostome lineage of animals, this gene cluster only occurs vertebrates. The CD33rSiglecs recognize endogenous sialylated glycans as SAMPs[29] and modulate myeloid cell lineage reactivity. These type I transmembrane proteins have an N-terminal immunoglobulin (Ig)-like-V-set domain mediating sialic acid (Sia) recognition, followed by a variable number of Ig-like-C-2 type domains, a transmembrane domain, and often, a cytoplasmic tail with one or more immunoreceptor tyrosine-based inhibitory motif (ITIMs). Siglecs with an ITIM motif recruit tyrosine phosphatases such as Src homology domain 2-containing tyrosine phosphatase-1 (SHP-1) dampen cell activation upon engagement of sialic acid-terminated glycans prominent on cell surface glycoproteins and glycolipids, which thus act as SAMPs.[45] The less common activating Siglecs have arginine or lysine residues in their transmembrane domains that instead activate cellular immune responses by recruiting the immunoreceptor tyrosine-based activatory motif (ITAM) containing DAP-12 or DAP-10 adaptors.[46] These inhibitory and activating Siglecs are prominently expressed on innate immune cells, modulating production of ROS[47,48] by phagolysosomal NOX, which generates ROS primarily for pathogen killing,[13,14] but can also have collateral effects on host cells.

Early studies showed that CD33rSiglecs are rapidly evolving and highly variable in gene number among species such as human, chimpanzees, monkeys, and mice.[49] In this regard, we noted a surprising correlation between the number of *CD33rSIGLEC* genes and maximum lifespan in mammalian species, and suggested that this association was due to modulation of oxidative stress arising from innate immune cell activity.[50] Consistent with this hypothesis, inactivation of the Siglec-E, a major *CD33rSIGLEC* gene in mice generated evidence for increased systemic oxidative damage.[50] In addition, impaired expression of the

free-radical scavenging enzyme Gstp1 in these mice leads to higher levels of oxidative adducts of lipids, proteins, and DNA that accelerated aging.

Evolutionary theory predicts selection against decoupling of senescence patterns between somatic and reproductive maintenance, leading to similar aging declines in both fitness traits.[51] Indeed, the classic concept of "antagonistic pleiotropy"[52] suggests that genes that support enhanced reproductive fitness in earlier periods of the lifespan can even be positively selected for even if they accelerate aging in late life success. In this regard, the existence of mid-life reproductive cessation in human females ("menopause") has been an apparent anomaly. This unusual feature has been hypothesized to have evolved in humans because of a "Grandmother" effect, that is, the contributions that post-reproductive females make to the fitness of their children and grandchildren.[53] According to this theory, grand-mothering explains increased adult survival and a longer lifespan and, in turn, favors a longer period of prepubertal development[54] and post-menopausal females who provide alloparental care to grandchildren are suggested to have increased fitness.[55,56] In keeping with this concept, we recently reported that certain gene variants that protect against late-life cognitive decline are derived and human-specific,[57] including a variant of *CD33* (which encodes Siglec-3/CD33).

We here seek to determine whether the suggested correlation between number of *CD33rSIGLEC* genes and maximum lifespan prevails when increasing the dataset of species, study the relative roles of inhibitory and activating Siglecs, ask if other gene families show similar correlations, re-evaluate the relative contribution of mitochondrial and NOX-derived ROS to lifetime exposure, and also address the relative role of *CD33rSIGLEC* genes in prolonged post-reproductive lifespan (PRLS).

2 | MATERIALS AND METHODS

2.1 | Analysis of genomic sequences and gene prediction strategy

Sequences of previously reported human and mouse *CD33rSIGLECs* were retrieved from NCBI (http://www.ncbi.nlm.nih.gov) and MGI (http://www.informatics.jax.org/), respectively. NCBI annotated *CD33rSIGLEC* genes for additional mammalian species were used as references for orthologous gene searching. Additional putative *CD33rSIGLEC* genes were obtained by searching available mammalian genome sequences at UCSC Genome Browser (http://genome.ucsc.edu/), Ensembl (http://www.ensembl.org), NCBI (http://www.ncbi.nlm.nih.gov/gene), and PreEnsembl (for Naked Mole Rat). As *SIGLEC* genes contain introns, we adopted and modified a previously established search strategy.[50,58] First, the candidate genes were subjected to BLAST and BLAT to identify the genomic location of a putative *CD33rSIGLEC* gene in a genome with a previously reported *CD33rSIGLEC* as a query. In addition, gene markers such as VRK3 and KLK were also considered to locate *SIGLEC* gene cluster in newly and poorly annotated mammalian genomes. Second, the orthologous gene was selected based on alignment and phylogenetic closeness with the query *CD33rSIGLEC*. The above gene search strategy was also applied for predicting another 10 randomly selected gene families in mammals, though the criteria used in gene structure evaluation were gene family dependent.

2.2 | Definition of functional genes

Based on previous studies on *CD33rSIGLECs*, several characteristics were considered to define a gene encoding a functional *SIGLEC*.[59] One criterion is that a Siglec protein has the characteristic amino-terminal V-set Ig-like domain capable of binding sialylated glycans. This binding activity requires a conserved arginine residue in the Ig-like V-set domain. The other criterion is that a functional Siglec protein should contain either a cytosolic tail with at least one ITIM motif or a transmembrane domain carrying a positively charged amino acid. The eventually acquired candidate CD33rSiglecs in each species were considered as true orthologs and used in our correlation analysis. Defining a functional gene using our gene prediction approach is not always clear-cut, due to the nature of incomplete genome sequences or genome sequencing errors. Thus, a few additional rules were considered during our prediction process. First, when entire exons of a gene (usually one or two) are missing due to a gap in the genome sequence but ORFs remain undisrupted in the available sequences, we treat the candidate locus as a functional gene. Second, the presence of essential arginine (R) and the cytoplasmic motif (YSEI) in putative orthologous gene were used as markers to identify functional proteins. Additionally, as discussed in the earlier study, genome coverage did not make any difference in the correlation of the number of *CD33rSIGLEC*s and maximum lifespan. There is no trend that higher coverage genomes will have higher number of *CD33rSIGLEC* genes or vice versa. For example, human and mouse genomes have the highest coverage (>12×) out of all mammals and they contain 11 and 4 *CD33rSIGLEC* genes, respectively. Furthermore, in our previous study, we had not included essential Arg mutated *SIGLEC*s because of their inability to recognize Sia. However, in the present study, essential Arg-mutated Siglecs were included because of their baseline activity independent of Sia recognition.

2.3 | Longevity and body weight data

Data regarding maximum lifespan and average adult body weight for mammalian species are obtained from AnAge: the animal aging and longevity (http://genomics.senescence.info/species/).[60] Also, the reproductive age of Orca (Killer whale) was obtained from a recent study.[53]

2.4 | Analysis of functional and null alleles in supercentenarians

To understand the presence and absence of functional *SIGLEC* genes in supercentenarians (age > 110 years) and individual aged ≤ 65 as control, we obtained the information regarding *SIGLEC*s from recently published paper on supercentenarians[61] and compared it with the 34 control genomes obtained from Personal Genomics Project (PGP), Harvard (http://www.personalgenomes.org/).

2.5 | Statistical analysis

Paired Student's *t* test was used for comparisons involving three polymorphic groups (*CD33m*, *SIGLEC-12*, and *SIGLEC-16*) in supercentenarian and control PGP cohort). Lifespan correlation analysis was performed on a linear scale using regression analysis as an add-in Data analysis toolpak in Excel. Prism program (GraphPad, La Jolla, CA) was used for the statistical analyses. Phylogeny generalized least-squares (PGLS) and Felsenstein's Independent Contrast (FIC) analysis were conducted in COMPARE 4.6b (http://www.indiana.edu/~martinsl/compare/) using a degree of freedom of 23, with three (one for calculating contrast and two for estimating the slope and the intercept) subtracted from 26 (the total number of taxa). Phylogenetic regressions controlled for the body mass were run using correlation analysis in the Mesquite software (http://mesquiteproject.org). The model included *CD33rSIGLEC* gene numbers as a response, and maximum lifespan and body mass as covariates.

3 | RESULTS

3.1 | Number of *CD33rSIGLEC* genes correlates with maximum lifespan across 26 species

Our previous study showed a correlation between *CD33rSIGLEC* gene number and maximum lifespan in 14 mammalian species. To assure that this correlation was not due to chance, we have now increased the number of mammalian species to 26 (detailed in Materials and Methods). The genomic localization of Siglecs among 26 mammalian species is shown in Figure 1 (also provided in Supplemental data are sequence files of all identified *CD33rSIGLECs* from all 26 species). Since these Siglecs are categorized based on ITIM or ITAM motifs present in their cytosolic tails, we studied them separately or in combination for their correlation with maximum lifespan. A moderate correlation was observed with *CD33rSIGLECs* containing the ITIM motifs ($R^2 = 0.42$, $P < .05$, Figure 2) and with those containing ITAM motifs ($R^2 = 0.27$, $P < .05$, Figure 2). However, the correlation with maximum lifespan was strongest when considering total *CD33rSIGLEC* gene numbers ($R^2 = 0.58$, $P =< .0001$, Figure 2).

To further test the correlation of maximum lifespan and other gene families, we examined 10 other gene families involved in different functions *viz*. transcription factors (*JUN-c, TRAC*, and *ZNF621*), immune regulation (*IFNα1* and *APOL1*), Cell Signalling (*MAS*, *NPM*), and Metabolic process (*AMY1α, GAA*, and *ORFAC*). None of the other gene families showed strong correlations between gene number and maximum lifespan in 26 mammalian species (Supplemental Figure 1).

3.2 | Correlation between *CD33rSIGLEC* gene number and maximum lifespan is independent of phylogeny or body weight

Closely related species may share similar traits due to their common ancestry and/or body weight, making data from different species not necessarily statistically independent. To control for such effects, we did phylogenetic comparative analysis using PGLS or FIC approaches. The correlation between *CD33rSIGLECs* and longevity remained robust after such phylogenetic correction (Supplemental Figure 2, Supplemental Table 1). Moreover, the correlation also persisted after mathematical correction for body mass represented by average adult body weight, another factor known to correlate with metabolic rate and lifespan (Supplemental Table 2). Taken together, these data indicate that the correlation of the number of *CD33rSIGLEC* genes with maximum lifespan in mammals appears independent of phylogenetic constraints (Pagel Lambda: $R^2 = 0.21$, $df = 24$), effects of genomic location, and evolution of receptors involved in immune responses and body mass.

3.3 | Estimation of adult reproductive lifetime production of neutrophils in humans

So far, we have assumed that the striking correlation of the number of *CD33rSIGLEC* genes with maximum lifespan is

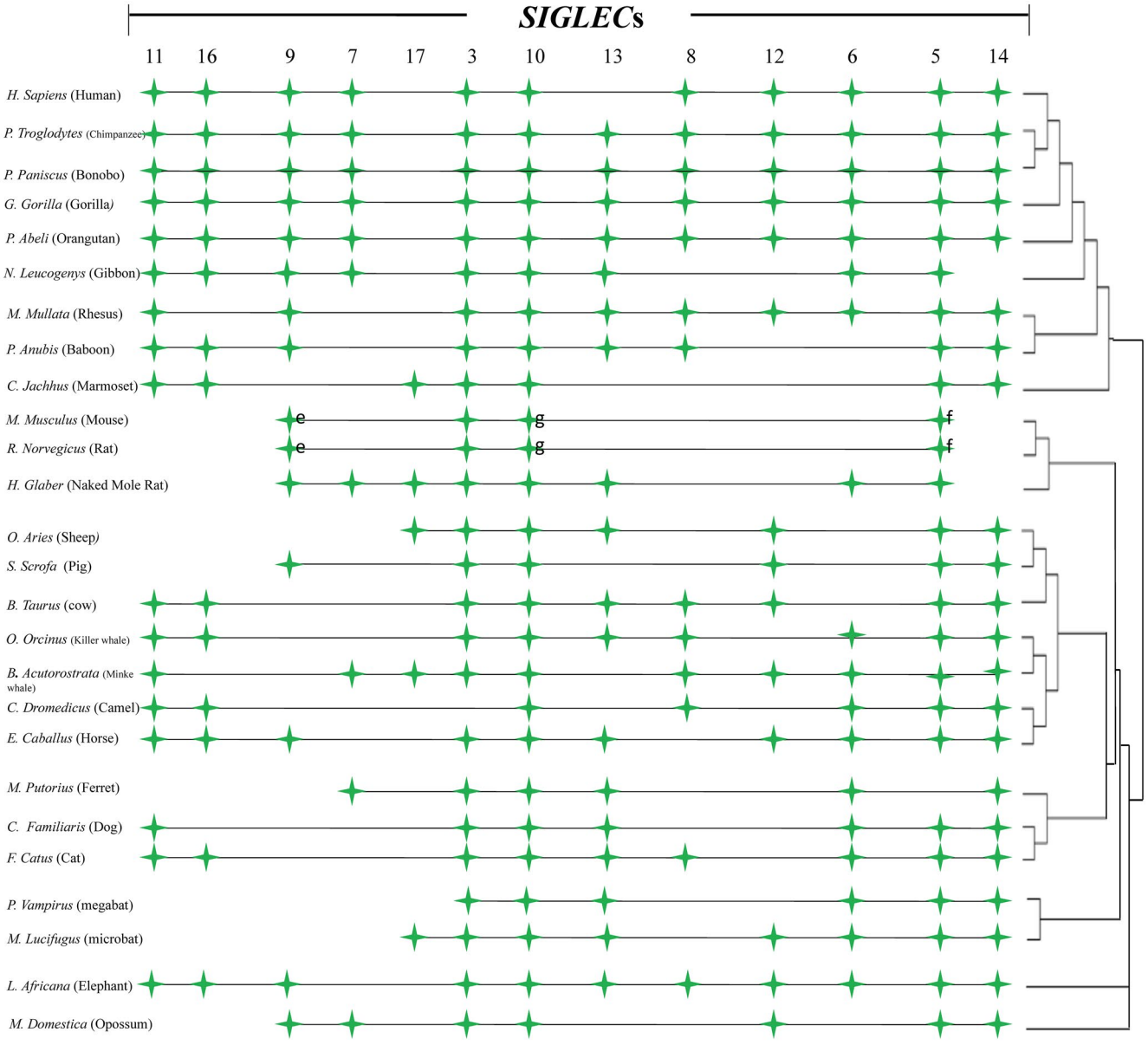

FIGURE 1 Presence and ordering of *CD33rSIGLEC* genes in Genomes of 26 mammalian species. The numbering designation of likely orthologous of *SIGLECs* is indicated at the top and the respective presence and absence of the *SIGLECs* is shown in each species. It should be noted that mouse and rat *SIGLECs* were based on the most likely phylogenetic relatedness with the human *SIGLECs*. The phylogeny (on the right) is not scaled and common names (as well as scientific) of 26 mammalian species can be found in Supplemental Table 3

related to differential modulation of lifetime phagolysosomal ROS production by NOX enzymes and their regulation by Siglecs. However, as detailed in the introduction, lifetime exposure to ROS is often assumed to originate largely from mitochondria, which are present in all cells in the body and represent a constant source. In contrast, ROS produced by phagolysosomal NOX enzymes originates mainly from neutrophils and other myeloid lineage phagocytes (monocytes) and are assumed to mostly originate during acute or chronic immune responses to invasive pathogens.[62,63] However, recent evidence suggests that neutrophils may spontaneously activate when they exit the circulation[64] and escape the inhibitory sialome-based milieu of the bloodstream.[28] Thus, prior to rapid demise, each neutrophil likely releases some ROS, with the amount depending on potential encounters with commensal or pathogenic microbes.

Combining independent lines of evidence including the short half-life and spontaneous activation of neutrophils, we here calculated that even without re-current inflammation, a major source of lifetime ROS exposure may be NOX-derived. Neutrophils are produced in the bone marrow, enter the circulation, and then exit into tissues. Traditional dogma maintains that these cells circulate for only a few hours. Estimated neutrophil half-lives in healthy individuals vary from 6.7 hours

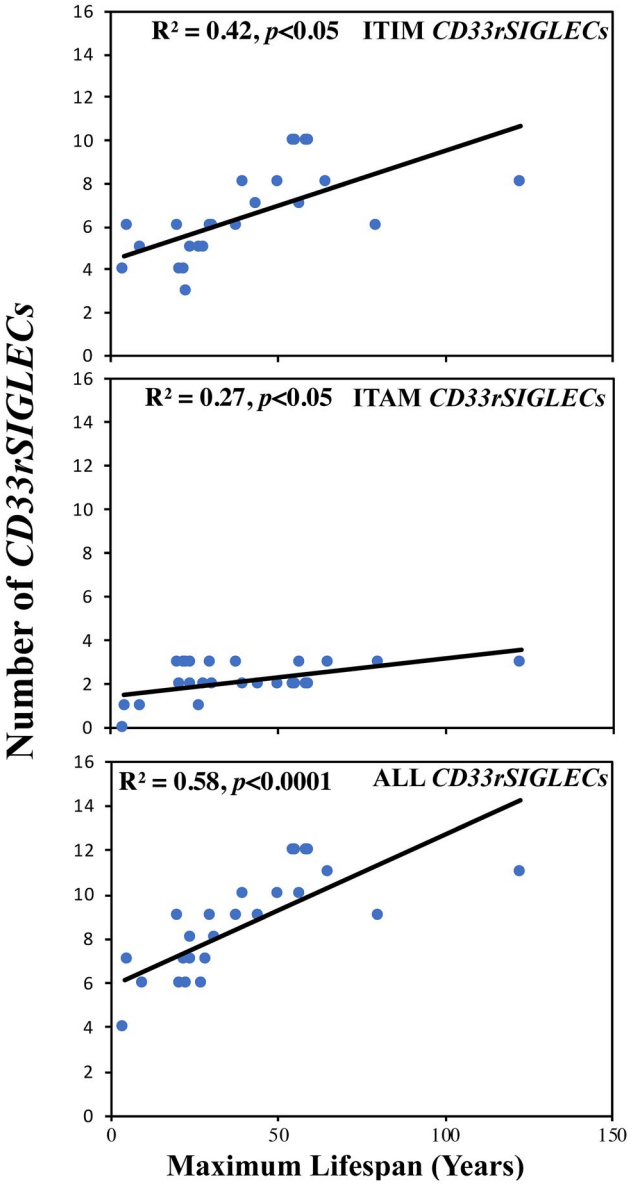

FIGURE 2 Correlation of *CD33rSIGLEC* genes encoding ITIM or ITAM motifs with maximum lifespan (Years). Level of significance (*P* value) was computed using Data analysis toolpak (add-in) in Excel

to 3 days (discussed later). However, the short lifetime values are based on reinfusion of purified neutrophils, which are effectively pre-activated and might be cleared more rapidly. The debate on these different values reached a compromise with a recent estimate of $T^{1/2}$ as 1 day.[65] Assuming this number is reasonably correct, how many neutrophils pass through the body of an adult human each year?

The first step is to calculate the circulating granulocyte pool (CGP) in a healthy individual of ~60 kg weight (Table 1). The average neutrophil count in circulating blood is ~5000 per μL, the typical blood volume is 5000 mL (5×10^6 μL), giving a CGP = $5000 \times 1000 \times 5000 = 25 \times 10^9$ cells. Taking into consideration the approximately equal

TABLE 1 Calculation of blood neutrophil turnover rate during one decade of an individual's life

Parameter	Estimates
Blood volume	5000 mL (A)
Average neutrophil count	5000/μL (B)
Circulating neutrophil pool (CGP) (A × B)	2.5×10^{10} cells (C)
Total neutrophil Blood Pool (TGBP) 2 × C	5×10^{10} cells (D)
Half-life in circulation (h)	1 Day
Neutrophil turnover rate per decade (3650 days)	0.91×10^{14} cells/Decade

Note: Assumes the half-life of neutrophils in circulation is 1 day (See text for discussion).

marginated granulocyte pool (MGP), the total granulocyte blood pool (TGBP) is 5×10^{10} cells. Assuming a half-life of 1day, neutrophil production during an average 30-year adult reproductive lifespan is then $5 \times 10^{10} \times 0.5 \times 30 \times 365 = 2.735 \times 10^{14}$ cells per 30 years (Table 2). Even ignoring multiple infectious episodes during which neutrophil counts and activity may drastically increase, the number of neutrophils that pass through the body of an adult human during an adult reproductive lifetime is ~10 times the number of cells in the body (latest estimate is 3.7×10^{13}).[66] Thus, a major source of ROS may be NOX2-derived, and this may help explain why *CD33rSIGLEC* gene number correlates so well with maximum lifespan.

3.4 | Supercentenarians do not have unusually high number of *CD33rSIGLEC* genes

Based on the findings thus far, we next sought to ask whether extremely long-lived humans harbor higher numbers of *CD33rSIGLEC* genes. We could ask this question because of the polymorphic presence of human-specific *CD33rSIGLECP* pseudogene alleles in human populations in variable frequency, with overall averages of *SIGLEC12P* (58%) *SIGLEC16P* (88%), and *SIGLEC14P* (39%), respectively.[67,68] In addition, there is a polymorphic presence of an alternatively spliced *CD33m* Alzheimer's disease protective allele, also specific to humans.[57] For various technical reasons, the *SIGLEC14P* allele comprised of >14 kb deletion is difficult to determine in standard genome sequences or in the variant call format (VCF) files available from databases. We thus explored the status of *SIGLEC12P*, *SIGLEC16P*, and *CD33m* in supercentenarians and control individuals. The number of CD33rSiglec alleles was correlated between 22 supercentenarian haploid genomes and 68 PGP control genomes, and we found no enrichment of ancestral intact alleles in supercentenarians (Table 2).

TABLE 2 Correlation of number of *CD33rSIGLEC* alleles in supercentenarians (N = 11) and Personal Genomics Project (PGP) controls (N = 34)

	Supercentenarian (22 haploid genomes)	PGP controls (68 haploid genomes)
CD33m	5	16
SIGLEC12P	8	10
SIGLEC16P	7	14

Note: Paired *t* test: $t = 1.557$ $df = 2$, $P = .2597$, Non-significant. Calculation based on *CD33m* (protective allele marker), *SIGLEC12P* (Pseudogene) and *SIGLEC16P* (Pseudogene).

3.5 | *CD33rSIGLEC* gene number correlates best with maximal reproductive lifespan, not total lifespan

The above finding in supercentenarians appears to go against our original hypothesis, which predicted a correlation between total *CD33rSIGLEC* gene numbers and maximum lifespan. However, in carefully reconsidering the data (Figure 3), we noticed that only two species (out of 26) fall far off the line of regression, namely humans and killer whales (orcas) (Figures 2, $R^2 = 0.58$, $P < .0001$). Notably, these are the two species that also have a prolonged PRLS. When we recalculated the regression statistics considering only reproductive lifespan of the above two species (orca and human), humans and killer whale gene numbers now fall on the line of regression, the correlation improves greatly and *P* value becomes highly significant (Figures 2, $R^2 = 0.83$, $P < .0001$).

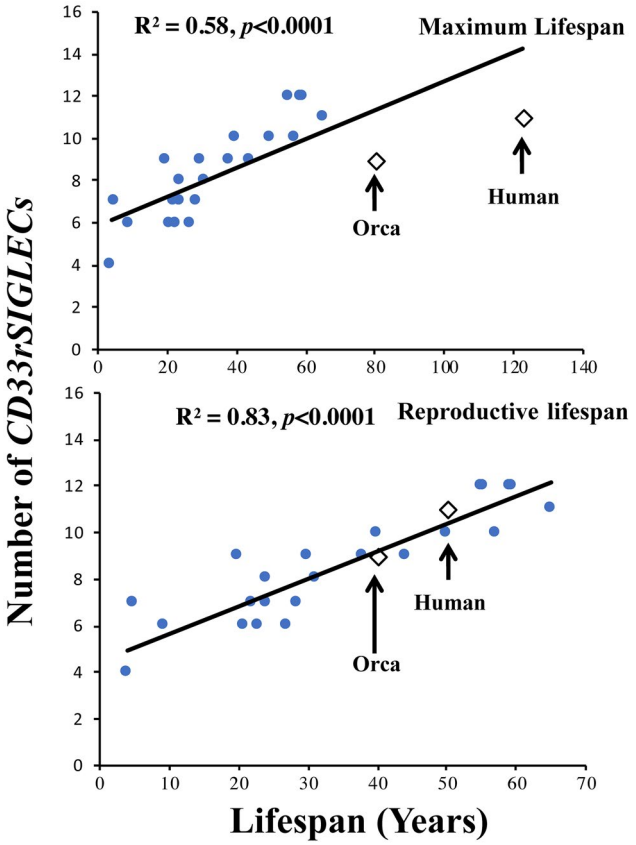

FIGURE 3 Correlation of *CD33rSIGLEC* genes with Maximum or Reproductive lifespan (Years). Level of significance (*P* value) was computed using Data analysis toolpak (add-in) in Excel

4 | DISCUSSION

Since Harman first proposed that ROS are mediators of the aging process,[69] much research has supported this concept. There are of course many mechanisms of aging, and ROS is only one contributor.[6,70] In fact, it is well documented that ROS can be beneficial as well as detrimental, and it is the balance of ROS production that apparently matters (too much is bad, too little is bad). ROS produced by phagolysosomal NOX are prominently produced by professional phagocytes such as neutrophils and monocytes which play a central role in the innate immune response. Furthermore, our discovery of constitutive RBC suppression of neutrophils shows that as soon as neutrophils leave the bloodstream, they are very likely to release ROS, even more so than upon encountering normal commensals at mucosal interfaces.[28] Because of the capacity of NOX2 to produce high levels of ROS constitutively, and the role of *SIGLEC*s in regulating phagocyte NOX2 activity, we suggest that NOX2 is the prominent source of ROS throughout life. While other NOX isoforms such as NOX4 and NOX5 are believed to become important contributors to tissue-specific ROS production in certain disease states, neither isoform has been systematically studied in human aging, so this remains an open question.[11,12]

CD33rSIGLEC genes are expressed on innate immune cells and are variable in number among species, which further reflects the cells variability in ROS production. Of course, other functions of Siglecs might also be associated to their variation in number. For example, bacterial pathogens expressing sialic acids can engage CD33rSiglecs[71,72] and might therefore represent strong driving forces for shaping the evolution of this class of receptors.[29,73] Our earlier study suggested that Siglec gene number of few mammalian species correlated with maximum lifespan and modulates ROS. Here we affirm this association with a larger dataset and found that total *CD33rSIGLEC* number in the genome correlates with maximum lifespan implicating the balanced production and destruction of ROS. However, this correlation becomes much stronger when the unusual PRLSs of humans and orcas was taken out of the analysis, suggesting the role of *CD33rSIGLECs* in regulation of reproductive lifespan.

Note that apart from humans, only two known species (killer whale (*Orcinus orca*) and short-finned pilot whale

(*Globicephala macrorhynchus*)) exhibit prolonged PRLS.[55] Information on post-reproductive survival in long-lived species has been more limited because of the difficulty in obtaining longitudinal data, especially from populations in the wild. While there is evidence of such post-reproductive lifespan in many other species, it is generally short, detectable only in few individuals and occurs mainly in captivity, in association with health promoting measures, veterinary care, and absence of predation.[55] This life history pattern stands in stark contrast to that of other mammalian species, wherein reduced fecundity generally coincides with somatic aging and eventually with death.[55,74] Also as discussed earlier, from the perspective of natural selection during evolution, successful reproduction with transmission of genetic material from individuals to the next generation is the key to "fitness." Thus, the most effective way for an organism to maximize its fitness is to continue to reproduce until somatic aging leads to the end of life. This is indeed the case for the vast majority of mammals, and the time between last parturition and death is related to the overall longevity of most species.[55] Thus, senescence can reduce fertility in late life, and a brief post-reproductive lifespan is a common trait in many mammals.[75] It has been noted that a few marine species such as beluga whale (*Delphinapterus leucas*), narwhal (*Monodon monoceros*), and false killer whale (*Pseudorca crassidens*)[74] do manifest a short but more significant post-reproductive lifespan. In addition, Asian elephants that share some features with humans and these whales in having large brains, being long-lived and having well-defined social networks also share short post-reproductive lifespans.[76] Indeed, in Asian elephants having a nearby grandmother can enhance calf survival even though the grandmother may be still reproducing.[77] However, the extension of life well beyond the end of the reproductive lifespan is extremely rare and is limited only to humans and two species of whales.[55,74,78-80]

Based on genetic variation, it is well documented that comparison of long-lived cohorts and controls showed variation. Analysis of candidate genes such as insulin-like growth factor 1 (*IGF-1*) and forkhead box O-3 (*FOXO3)* have revealed that polymorphisms in these genes are associated with extreme longevity.[81,82] In addition, genome-wide association studies have shown that the apolipoprotein e4 (*APOE4)* haplotype is depleted in centenarians.[83-85] However, more variants that decrease gene function in the Teashirt Zinc Finger Homeobox 3 (*TSHZ3)* were observed in Centenarian and supercentenarian genomes (~110 years) than control genomes.[61] Considering these studies, we found that supercentenarians genomes did not contain an increased number of functional *SIGLEC* alleles in comparison to controls. Although the *CD33* protective allele (also known as CD33m) helps reduce cognitive decline and Alzheimer's disease, the frequency of the variant was not significantly different between supercentenarian and control genomes which supports the conclusion that *CD33rSIGLEC*s are not involved in regulating PRLS, but instead contribute to the regulation of reproductive lifespan. A recent paper showed Older Amish individuals harbor a rare loss-of-function mutation in *SERPINE1* gene that protects against effects of ageing. This mutation affects the function of plasminogen activator inhibitor-1 (PAI-1), which has a vital role in cellular senescence and is expressed at higher levels in senescent cells.[86] This loss of function in "wellderly" individuals could be one of the factors that regulate PRLS. Further investigation is certainly warranted to decipher other factors regulating PRLS and future studies based on post-reproductive lifespan and correlation of genetic variants are required. The naked mole-rat appears resistant to aging with a maximal lifespan of 30 years or more.[87] Despite having such an exceptional longer lifespan, the *CD33rSIGLEC* gene number is only appropriately higher than that of other rodents and falls on the regression line.

We have also now calculated that a very large number of neutrophils are produced and enter tissues during an adult reproductive lifespan, and their NOX enzyme may possibly be a major source of total lifetime ROS along with whole body mitochondria. Siglecs are major regulators of ROS production by phagocytes, perhaps explaining why *SIGLEC* gene numbers so strongly correlated with maximum reproductive lifespan. A relevant question is how much ROS each neutrophil produces after it leaves the circulation and before it disappears. In this regard, our recent work[28] shows that immediately after separation of neutrophils from the heavily sialylated RBCs and plasma proteins that suppress activation by engaging Siglecs, these cells become spontaneously activated. An additional consideration is that, in acute idiopathic or drug-induced agranulocytosis (which involves the sudden selective loss of neutrophils), tissues are rapidly invaded by "normal commensal" organisms.[88] Indeed, prompt recognition of patients with neutropenic fever of any kind is a well-established imperative and the administration of empiric systemic antibiotic therapy is typically used to reduce the risk of severe sepsis.[88,89] These clinical observations suggest that at steady-state neutrophils are likely to be constantly encountering and eliminating commensal organisms that enter too deep into host tissues, likely generating ROS all the time, even in the healthy state.

In summary, the present study shows a strong correlation between *CD33rSIGLEC* gene number and reproductive lifespan further implicating the role of these receptors in balancing oxidative stress and modulate aging patterns and lifespan in mammals. We know of no other gene family that shows such a striking correlation with lifespan and gene number. In view of this fact, and the well-known ROS modulatory role of *SIGLEC*s it is natural for us to focus on this aspect. The lack of correlation with post-reproductive lifespan suggests roles of other genes that are under study. The larger implication is that the study of aging phenomena in humans needs to consider the possibility that the post-reproductive period may be regulated by factors different from those gleaned from studying aging in other animals.

ACKNOWLEDGMENTS

We thank members of the laboratory of AV for helpful discussions and suggestions. This study was supported by a grant from National Institute of Health (R01GM32373) to AV.

CONFLICT OF INTEREST

The authors declare that no conflict of interest.

AUTHOR CONTRIBUTIONS

A. Varki designed, supervised, and wrote the manuscript; P. Gagneux designed, supervised, and edited the manuscript; L.L. Dugan analyzed, wrote, and edited the manuscript; S.K. Kim provided data and edited the manuscript; N. Khan analyzed and wrote the manuscript.

REFERENCES

1. Partridge L. Intervening in ageing to prevent the diseases of ageing. *Trends Endocrinol Metab*. 2014;25:555-557.
2. Kauppila TES, Kauppila JHK, Larsson NG. Mammalian mitochondria and aging: an update. *Cell Metab*. 2017;25:57-71.
3. Sen P, Shah PP, Nativio R, Berger SL. Epigenetic mechanisms of longevity and aging. *Cell*. 2016;166:822-839.
4. Satoh A, Imai SI, Guarente L. The brain, sirtuins, and ageing. *Nat Rev Neurosci*. 2017;18:362-374.
5. López-Otín C, Galluzzi L, Freije JMP, Madeo F, Kroemer G. Metabolic control of longevity. *Cell*. 2016;166:802-821.
6. Campisi J, Kapahi P, Lithgow GJ, Melov S, Newman JC, Verdin E. From discoveries in ageing research to therapeutics for healthy ageing. *Nature*. 2019;571:183-192.
7. Hekimi S, Lapointe J, Wen Y. Taking a "good" look at free radicals in the aging process. *Trends Cell Biol*. 2011;21:569-576.
8. Dugan LL, Quick KL. Reactive oxygen species and aging: evolving questions. *Sci Aging Knowledge Environ*. 2005;2005:pe20.
9. Bedard K, Jaquet V, Krause KH. NOX5: from basic biology to signaling and disease. *Free Radic Biol Med*. 2012;52:725-734.
10. Cachat J, Deffert C, Hugues S, Krause KH. Phagocyte NADPH oxidase and specific immunity. *Clin Sci (Lond)*. 2015;128:635-648.
11. Jha JC, Watson AMD, Mathew G, de Vos LC, Jandeleit-Dahm K. The emerging role of NADPH oxidase NOX5 in vascular disease. *Clin Sci (Lond)*. 2017;131:981-990.
12. Yang Q, Wu FR, Wang JN, et al. Nox4 in renal diseases: an update. *Free Radic Biol Med*. 2018;124:466-472.
13. Brandes RP, Kreuzer J. Vascular NADPH oxidases: molecular mechanisms of activation. *Cardiovasc Res*. 2005;65:16-27.
14. Amulic B, Cazalet C, Hayes GL, Metzler KD, Zychlinsky A. Neutrophil function: from mechanisms to disease. *Annu Rev Immunol*. 2012;30:459-489.
15. Quinn MT, Ammons MC, Deleo FR. The expanding role of NADPH oxidases in health and disease: no longer just agents of death and destruction. *Clin Sci (Lond)*. 2006;111:1-20.
16. D'Autréaux B, Toledano MB. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat Rev Mol Cell Biol*. 2007;8:813-824.
17. Yang W, Hekimi S. A mitochondrial superoxide signal triggers increased longevity in Caenorhabditis elegans. *PLoS Biol*. 2010;8:e1000556.
18. Van Raamsdonk JM, Hekimi S. Superoxide dismutase is dispensable for normal animal lifespan. *Proc Natl Acad Sci U S A*. 2012;109:5785-5790.
19. Masoro EJ. Caloric restriction and aging: an update. *Exp Gerontol*. 2000;35:299-305.
20. Schriner SE, Linford NJ, Martin GM, et al. Extension of murine life span by overexpression of catalase targeted to mitochondria. *Science*. 2005;308:1909-1911.
21. Naviaux RK. Oxidative shielding or oxidative stress. *J Pharmacol Exp Ther*. 2012;342:608-618.
22. Di Meo S, Reed TT, Venditti P, Victor VM. Role of ROS and RNS sources in physiological and pathological conditions. *Oxid Med Cell Longev*. 2016;2016:1245049.
23. Yoon YS, Lee JH, Hwang SC, Choi KS, Yoon G. TGF beta1 induces prolonged mitochondrial ROS generation through decreased complex IV activity with senescent arrest in Mv1Lu cells. *Oncogene*. 2005;24:1895-1903.
24. Ali SS, Young JW, Wallace CK, et al. Initial evidence linking synaptic superoxide production with poor short-term memory in aged mice. *Brain Res*. 2011;1368:65-70.
25. Finkel T, Holbrook NJ. Oxidants, oxidative stress and the biology of ageing. *Nature*. 2000;408:239-247.
26. Balaban RS, Nemoto S, Finkel T. Mitochondria, oxidants, and aging. *Cell*. 2005;120:483-495.
27. Brown GC, Borutaite V. There is no evidence that mitochondria are the main source of reactive oxygen species in mammalian cells. *Mitochondrion*. 2012;12:1-4.
28. Lizcano A, Secundino I, Döhrmann S, et al. Erythrocyte sialoglycoproteins engage Siglec-9 on neutrophils to suppress activation. *Blood*. 2017;129:3100-3110.
29. Varki A. Since there are PAMPs and DAMPs, there must be SAMPs? Glycan "self-associated molecular patterns" dampen innate immunity, but pathogens can mimic them. *Glycobiology*. 2011;21:1121-1124.
30. Mocsai A. Diverse novel functions of neutrophils in immunity, inflammation, and beyond. *J Exp Med*. 2013;210:1283-1299.
31. Finch CE, Morgan TE, Longo VD, de Magalhaes JP. Cell resilience in species life spans: a link to inflammation? *Aging Cell*. 2010;9:519-526.
32. von Gunten S, Yousefi S, Seitz M, et al. Siglec-9 transduces apoptotic and nonapoptotic death signals into neutrophils depending on the proinflammatory cytokine environment. *Blood*. 2005;106:1423-1431.
33. Nutku-Bilir E, Hudson SA, Bochner BS. Interleukin-5 priming of human eosinophils alters siglec-8 mediated apoptosis pathways. *Am J Respir Cell Mol Biol*. 2008;38:121-124.
34. Kano G, Almanan M, Bochner BS, Zimmermann N. Mechanism of Siglec-8-mediated cell death in IL-5-activated eosinophils: role for reactive oxygen species-enhanced MEK/ERK activation. *J Allergy Clin Immunol*. 2013;132:437-445.
35. Claude J, Linnartz-Gerlach B, Kudin AP, Kunz WS, Neumann H. Microglial CD33-related Siglec-E inhibits neurotoxicity by preventing the phagocytosis-associated oxidative burst. *J Neurosci*. 2013;33:18270-18276.
36. McMillan SJ, Sharma RS, Richards HE, Hegde V, Crocker PR. Siglec-E promotes beta2-integrin-dependent NADPH oxidase activation to suppress neutrophil recruitment to the lung. *J Biol Chem*. 2014;289:20370-20376.
37. Favier B. Regulation of neutrophil functions through inhibitory receptors: an emerging paradigm in health and disease. *Immunol Rev*. 2016;273:140-155.

38. Karlstetter M, Kopatz J, Aslanidis A, et al. Polysialic acid blocks mononuclear phagocyte reactivity, inhibits complement activation, and protects from vascular damage in the retina. *EMBO Mol Med*. 2017;9:154-166.
39. Patras KA, Coady A, Olson J, et al. Tamm-Horsfall glycoprotein engages human Siglec-9 to modulate neutrophil activation in the urinary tract. *Immunol Cell Biol*. 2017;95:960-965.
40. Bochner BS, Zimmermann N. Role of siglecs and related glycan-binding proteins in immune responses and immunoregulation. *J Allergy Clin Immunol*. 2015;135:598-608.
41. Adams OJ, Stanczak MA, von Gunten S, Läubli H. Targeting sialic acid-Siglec interactions to reverse immune suppression in cancer. *Glycobiology*. 2018;28:640-647.
42. Varki A, Schnaar RL, Crocker PR. I-type lectins. In: Varki A, Cummings RD, Esko JD, et al., eds. *Essentials of glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2017:450–467.
43. Angata T. Possible influences of endogenous and exogenous ligands on the evolution of human siglecs. *Front Immunol*. 2018;9:2885.
44. Bornhöfft KF, Goldammer T, Rebl A, Galuska SP. Siglecs: a journey through the evolution of sialic acid-binding immunoglobulin-type lectins. *Dev Comp Immunol*. 2018;86:219-231.
45. Walter RB, Raden BW, Zeng R, Hausermann P, Bernstein ID, Cooper JA. ITIM-dependent endocytosis of CD33-related Siglecs: role of intracellular domain, tyrosine phosphorylation, and the tyrosine phosphatases, Shp1 and Shp2. *J Leukoc Biol*. 2008;83:200-211.
46. Blasius AL, Cella M, Maldonado J, Takai T, Colonna M. Siglec-H is an IPC-specific receptor that modulates type I IFN secretion through DAP12. *Blood*. 2006;107:2474-2476.
47. Crocker PR, Varki A. Siglecs in the immune system. *Immunology*. 2001;103:137-145.
48. Gabrilovich DI, Nagaraj S. Myeloid-derived suppressor cells as regulators of the immune system. *Nat Rev Immunol*. 2009;9:162-174.
49. Angata T, Margulies EH, Green ED, Varki A. Large-scale sequencing of the CD33-related Siglec gene cluster in five mammalian species reveals rapid evolution by multiple mechanisms. *Proc Natl Acad Sci U S A*. 2004;101:13251-13256.
50. Schwarz F, Pearce OM, Wang X, et al. Siglec receptors impact mammalian lifespan by modulating oxidative stress. *eLife*. 2015;4:e06184.
51. Hamilton WD. The moulding of senescence by natural selection. *J Theor Biol*. 1966;12:12-45.
52. Gc W. Pleiotropy, natural selection, and the evolution of senescence. *Evolution*. 1957;11:398-411.
53. Croft DP, Johnstone RA, Ellis S, et al. Reproductive conflict and the evolution of menopause in killer whales. *Curr Biol*. 2017;27:298-304.
54. Hawkes K, Smith KR. Brief communication: evaluating grandmother effects. *Am J Phys Anthropol*. 2009;140:173-176.
55. Croft DP, Brent LJ, Franks DW, Cant MA. The evolution of prolonged life after reproduction. *Trends Ecol Evol*. 2015;30(7):407-416.
56. Hawkes K, Coxworth JE. Grandmothers and the evolution of human longevity: a review of findings and future directions. *Evol Anthropol*. 2013;22:294-302.
57. Schwarz F, Springer SA, Altheide TK, Varki NM, Gagneux P, Varki A. Human-specific derived alleles of CD33 and other genes protect against postreproductive cognitive decline. *Proc Natl Acad Sci U S A*. 2016;113:74-79.
58. Shi P, Zhang J. Contrasting modes of evolution between vertebrate sweet/umami receptor genes and bitter receptor genes. *Mol Biol Evol*. 2006;23:292-300.
59. Crocker PR, Clark EA, Filbin M, et al. Siglecs: a family of sialic-acid binding lectins [letter]. *Glycobiology*. 1998;8:v.
60. de Magalhaes JP, Costa J. A database of vertebrate longevity records and their relation to other life-history traits. *J Evol Biol*. 2009;22:1770-1774.
61. Gierman HJ, Fortney K, Roach JC, et al. Whole-genome sequencing of the world's oldest people. *PLoS ONE*. 2014;9:e112430.
62. Robinson JM. Reactive oxygen species in phagocytic leukocytes. *Histochem Cell Biol*. 2008;130:281-297.
63. Finch CE. Evolution in health and medicine Sackler colloquium: Evolution of the human lifespan and diseases of aging: roles of infection, inflammation, and nutrition. *Proc Natl Acad Sci U S A*. 2010;107(Suppl 1):1718-1724.
64. Mayadas TN, Cullere X, Lowell CA. The multifaceted functions of neutrophils. *Annu Rev Pathol*. 2014;9:181-218.
65. Lahoz-Beneytez J, Elemans M, Zhang Y, et al. Human neutrophil kinetics: modeling of stable isotope labeling data supports short blood neutrophil half-lives. *Blood*. 2016;127:3431-3438.
66. Bianconi E, Piovesan A, Facchin F, et al. An estimation of the number of cells in the human body. *Ann Hum Biol*. 2013;40:463-471.
67. Mitra N, Banda K, Altheide TK, et al. SIGLEC12, a human-specific segregating (pseudo)gene, encodes a signaling molecule expressed in prostate carcinomas. *J Biol Chem*. 2011;286:23003-23011.
68. Yamanaka M, Kato Y, Angata T, Narimatsu H. Deletion polymorphism of SIGLEC14 and its functional implications. *Glycobiology*. 2009;19:841-846.
69. Harman D. Aging: a theory based on free radical and radiation chemistry. *J Gerontol*. 1956;11:298-300.
70. Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell*. 2013;153:1194-1217.
71. Carlin AF, Uchiyama S, Chang YC, Lewis AL, Nizet V, Varki A. Molecular mimicry of host sialylated glycans allows a bacterial pathogen to engage neutrophil Siglec-9 and dampen the innate immune response. *Blood*. 2009;113:3333-3336.
72. Chang YC, Nizet V. The interplay between Siglecs and sialylated pathogens. *Glycobiology*. 2014;24:818-825.
73. Varki A, Gagneux P. Multifarious roles of sialic acids in immunity. *Ann N Y Acad Sci*. 2012;1253:16-36.
74. Ellis S, Franks DW, Nattrass S, et al. Postreproductive lifespans are rare in mammals. *Ecol Evol*. 2018;8:2482-2494.
75. Cohen AA. Female post-reproductive lifespan: a general mammalian trait. *Biol Rev Camb Philos Soc*. 2004;79:733-750.
76. Lahdenperä M, Mar KU, Lummaa V. Reproductive cessation and post-reproductive lifespan in Asian elephants and pre-industrial humans. *Front Zool*. 2014;11:54.
77. Lahdenperä M, Mar KU, Lummaa V. Nearby grandmother enhances calf survival and reproduction in Asian elephants. *Sci Rep*. 2016;6:27213.
78. Austad SN. Menopause: an evolutionary perspective. *Exp Gerontol*. 1994;29:255-263.
79. Alberts SC, Altmann J, Brockman DK, et al. Reproductive aging patterns in primates reveal that humans are distinct. *Proc Natl Acad Sci U S A*. 2013;110:13440-13445.
80. Photopoulou T, Ferreira IM, Best PB, Kasuya T, Marsh H. Evidence for a postreproductive phase in female false killer whales Pseudorca crassidens. *Front Zool*. 2017;14:30.

81. Suh Y, Atzmon G, Cho MO, et al. Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proc Natl Acad Sci U S A*. 2008;105:3438-3442.
82. Willcox BJ, Donlon TA, He Q, et al. FOXO3A genotype is strongly associated with human longevity. *Proc Natl Acad Sci U S A*. 2008;105:13987-13992.
83. Nebel A, Kleindorp R, Caliebe A, et al. A genome-wide association study confirms APOE as the major gene influencing survival in long-lived individuals. *Mech Ageing Dev*. 2011;132:324-330.
84. Sebastiani P, Solovieff N, DeWan AT, et al. Genetic signatures of exceptional longevity in humans. *PLoS One*. 2012;7:e29848.
85. Fortney K, Dobriban E, Garagnani P, et al. Genome-wide scan informed by age-related disease identifies loci for exceptional human longevity. *PLoS Genet*. 2015;11:e1005728.
86. Khan SS, Shah SJ, Klyachko E, et al. A null mutation in SERPINE1 protects against biological aging in humans. *Sci Adv*. 2017;3:eaao1617.
87. Ruby JG, Smith M, Buffenstein R. Naked Mole-Rat mortality rates defy gompertzian laws by not increasing with age. *eLife*. 2018;7:e31157.
88. Newburger PE, Dale DC. Evaluation and management of patients with isolated neutropenia. *Semin Hematol*. 2013;50:198-206.
89. White L, Ybarra M. Neutropenic fever. *Hematol Oncol Clin North Am*. 2017;31:981-993.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Khan N, Kim SK, Gagneux P, Dugan LL, Varki A. Maximum reproductive lifespan correlates with *CD33rSIGLEC* gene number: Implications for NADPH oxidase-derived reactive oxygen species in aging. *The FASEB Journal*. 2020;34:1928–1938. https://doi.org/10.1096/fj.201902116R