Selection on Alleles Affecting Human Longevity and Late-Life Disease: The Example of Apolipoprotein E

Fotios Drenos*†, Thomas B. L. Kirkwood

Institute for Ageing and Health, Newcastle University, Newcastle Upon Tyne, Tyne, United Kingdom

Abstract

It is often claimed that genes affecting health in old age, such as cardiovascular and Alzheimer diseases, are beyond the reach of natural selection. We show in a simulation study based on known genetic (apolipoprotein E) and non-genetic risk factors (gender, diet, smoking, alcohol, exercise) that, because there is a statistical distribution of ages at which these genes exert their influence on morbidity and mortality, the effects of selection are in fact non-negligible. A gradual increase with each generation of the e2 and e3 alleles of the gene at the expense of the e4 allele was predicted from the model. The e2 allele frequency was found to increase slightly more rapidly than that for e3, although there was no statistically significant difference between the two. Our result may explain the recent evolutionary history of the epsilon 2, 3 and 4 alleles of the apolipoprotein E gene and has wider relevance for genes affecting human longevity.

Introduction

Evidence points towards the existence of a strong heritable component of human longevity [1,2]. Around a quarter to a third of the variability of lifespan can be attributed to the action of genes [3,4,5,6,7]. One of the best examples of a gene affecting survival in old age is the apolipoprotein E gene APOE. Situated on human chromosome 19 at locus 19q13.32 [http://genome.ucsc.edu/], the APOE gene is 3.6 kb long containing four exons and coding for a 317 amino-acid polypeptide that after cleavage gives rise to a 299 amino-acid long mature protein [8,9]. Apolipoprotein E (APO E) is a member of a diverse family of carrier proteins specializing in lipoprotein particle formation, secretion, transport, binding and metabolism [10]. APO E is synthesized in many different regions of the body such as the liver, brain (primarily astrocytes), skin, macrophages and steroidogenic organs [11].

APO E has three major and more than thirty minor isoforms, the latter being mostly linked with disease. The three common alleles are epsilon (ε) 2, 3 and 4 producing three homozygous (ε2/ε2, ε3/ε3 and ε4/ε4) and three heterozygous (ε2/ε3, ε3/ε4 and ε2/ε4) genotypes [12]. The difference between the alleles lies at two amino acid residues 112 and 158; with the most common allele ε3 having cysteine and arginine at these residues, respectively, while ε2 has cysteine and ε4 arginine at both locations [9]. Carriers of different alleles show differences in the incidence of coronary artery disease (CAD), peripheral atherosclerosis, Alzheimer disease, possibly stroke and even ability to recover from trauma [11,13,14,15,16]. The underlying mechanism for the action of the gene involves an inter-domain interaction between the amino- and carboxyl-terminals of the protein, producing isoform-specific lipoprotein preferences such that E2 and E3 protein isoforms bind preferentially to HDL (the “good” cholesterol), while the E4 isoform shows a preference for VLDL [17]. The APO E2 isoform is defective in binding the LDL receptor, although it retains its ability to interact with LDL-receptor-related protein and other related receptors [11]. A small number of ε2 homozygotes suffer from a condition known as type III hyperlipoproteinaemia, characterized by accumulation of cholesterol-rich remnant lipoproteins due to incomplete catabolism of chylomicrons and VLDL, leading to premature atherosclerosis. The fact that although 90% of the patients are ε2 homozygotes, but only 5% of the ε2 homozygotes suffer from the condition, suggests that further factors are required for the phenotypic expression of the condition [18,19].

Generally in Caucasian populations survival to advanced age is more likely for carriers of the ε2 allele than for ε3 homozygotes and less likely for carriers of the ε4 allele [20,21,22,23,24]. The ε4 allele is strongly associated with increased risk of coronary heart disease (CHD) [14,25,26,27,28], while for the ε2 allele there is some, though weaker, evidence for a protective effect [29,30]. The APOE polymorphism, has also been implicated in Alzheimer’s disease (AD), which is the commonest form of senile dementia [31,32]. With the ε4 allele shown, in a number of studies, to be associated with both familial and sporadic forms of the disease, causing higher incidence and earlier age of onset, and affecting its pathology and rate of progression [33,34]. Normal age-associated cognitive decline, and mild cognitive impairment and the risk of its conversion to AD have also been repeatedly associated with the presence of the ε4 allele [35].

One of the major ideas in the evolutionary theory of ageing [36] is the suggestion that, because the force of natural selection...
declines with age, alleles with deleterious effects seen only at older ages can reach higher frequencies than those that have their effects earlier in life [37]. Therefore, if a gene exerts an effect only after the end of the reproductive phase of the lifespan it has been thought unlikely that it could have been subject to significant direct selection pressure, and this would seem to be the case for some genes affecting late-life diseases such as AD and CHD [38,39]. Nevertheless, the worldwide abundance of ε3, as compared to ε4, suggests that selection has acted upon the ancestral allele [40,41], as compared to ε4 which, from studies in non-human primates appears to be the

Methods

Risk factors

The selection of modifiable risk factors to include in the simulation was based on current knowledge of lifestyle parameters affecting cardiovascular disease (CVD) while keeping in mind the need for independence between the risk factors, simplicity and availability of data. There are two categories of risk factors associated with CVD: non-modifiable risk factors, such as APOE genotype, and gender, and modifiable risk factors, such as smoking, unhealthy diet, lack of physical activity and high alcohol consumption [42]. For the non-modifiable risk factors, the estimated genotype relative risk was estimated from Gerdes et al [29], while the gender difference in CVD was based on the results of Panagiotakos et al [43]. To quantify the effect of diet, we used the five clusters described by Millen et al [44] in the Framingham study (Heart Healthy, Light Eating, Wine and Moderate Eating, High Fat, and Empty Caloric) [44,45]. For alcohol, smoking and exercise, we used the risk estimates from Stampfer et al [46].

Genotype-environment interaction

A number of studies found evidence of interaction between the APOE genotype and each of the modifiable risk factors considered in the model. The selection of modifiable risk factors was based on current knowledge of lifestyle parameters affecting cardiovascular disease (CVD) while keeping in mind the need for independence between the risk factors, simplicity and availability of data. There are two categories of risk factors associated with CVD: non-modifiable risk factors, such as APOE genotype, and gender, and modifiable risk factors, such as smoking, unhealthy diet, lack of physical activity and high alcohol consumption [42]. For the non-modifiable risk factors, the estimated genotype relative risk was estimated from Gerdes et al [29], while the gender difference in CVD was based on the results of Panagiotakos et al [43]. To quantify the effect of diet, we used the five clusters described by Millen et al [44] in the Framingham study (Heart Healthy, Light Eating, Wine and Moderate Eating, High Fat, and Empty Caloric) [44,45]. For alcohol, smoking and exercise, we used the risk estimates from Stampfer et al [46].

Calculated risk

Despite OR overestimating risk for common diseases such as CVD [56,57], many researchers do use it to report their results. To approximate the relative risk (RR) from the OR we used the very popular methodology of Zhang and Yu [58] as being both the simplest and the one that requires the least information for the dataset used. Despite the method's slight inaccuracy in calculating confidence intervals and its inability to account for confounding factors, its balance between simplicity and precision make it an especially useful tool [56,57].

The total risk of an individual could be calculated as the product of the gender relative risk and all four modifiable risk factors, corresponding to the individual's specific genotype. We used a proportional hazard model to estimate the mortality schedule of an individual having relative risk R at a specific age x. According to this, if $\mu_j(x)$ is the hazard of death of genotype j at age x, and $\mu_0(x)$ is the hazard of death of a genotype 2 at the same age x, then $\mu_j(x) = R \mu_0(x)$, where R expresses the relative risk [29].

Simulating evolutionary change

Our simulated populations comprised individuals with randomly assigned genotype, gender and lifestyle parameters from a typical western population. The random draws were done using the MT19937 pseudorandom generator, a variation of the “Mersenne Twister” generator (GSL Reference Manual). Equal number of males and females were generated and subsequently paired as couples randomly. These couples reproduced to age 50 or until either of the individuals died. The reproductive schedule was the same as what is found in contemporary western populations, with a series of binomials distributions describing the chance of a successful birth for each year of life and the probability decreasing with increasing adult age. The age of death was obtained from a Gumbell distribution based on demographic tables for the UK (Office for National Statistics) and a relative risk corresponding to the life style parameters of the simulated individual. Each offspring was given a genotype based on the genotypes of its parents, and all the produced genotypes were pooled into a matrix constituting the initial conditions for the next generation. The process was repeated as many times as the number of generations required. The entire simulation was written in C++ and performed using a 16 CPUs Unix cluster.

Statistical analysis of the results

The Mathematica 4.1 package was used to collect analyse the output of the simulations. To minimize random variation of the results, a set of 60 runs, each comprising 200,000 individuals followed for 50 generations, was used. We summarized the data calculating the mean and standard deviation of all 60 runs for each generation and each allele. A random walk in one dimension was used to obtain the average change per generation, assuming that the change of the allele frequencies was linear and could be described by a simple equation. As long as the alleles do not have a pleiotropic effect balancing the selection for or against them, their frequencies in a stable environment were expected to increase until they were fixed or became extinct in the population. A simple equation describing the change could be written as:

$$x_t = x_{t-1} + d + \sum_i$$

where $x_t$ is the frequency at time or generation t, $x_{t-1}$ is frequency
in the previous generation, \( d \) is the change between two consecutive points in the series and \( \Sigma_e \) is random noise normally distributed. In order to calculate \( d \) we estimated a \( y_{jt} \) such that:

\[
y_{jt} = z_{j+1} - z_j
\]

with simulation run \( j = 1, \ldots, 60 \)

and generation number \( i = 1, \ldots, 49 \)

returning a matrix of 49×60 elements. From this, the mean \( d \) and its 95% confidence intervals (95% CIs) could be calculated. If the 95% CIs are positive then there is evidence for a significant increase of the allele frequency with each generation, while if negative the allele goes towards extinction. In the case that 0 is included within the confidence interval, the change of frequency cannot be considered statistically significant and selection is either extremely weak or due to chance. The data were also transferred into Minitab statistical software to test for statistical significance between and within different variants of the simulation.

**Results**

Starting with APOE genotype frequencies of 1.7% \( e2e2 \), 11.6% \( e2e3 \), 55.8% \( e3e3 \), 1.9% \( e2e4 \), 25.1% \( e3e4 \), and 3.9% \( e4e4 \) and modifiable risk factor distributions as reported in Millen et al \[44\] and Stampfer et al \[46\], and assuming the same population distribution between males and females, the mean life expectancy at birth for the simulated individual was 75.7 years for females with a mean lifespan 0.9 years longer expected lifespan than males. The differences in mortality between APOE genotype classes as fitted in the model can be seen in Figure 1, with \( e2 \) carriers living 1 (OR 0.9) year more, and \( e4 \) carriers living 1.2 (OR 1.13) years less, than the \( e3e3 \) reference genotype. In Figure 2, the curves obtained when the interaction between genotype and diet was considered can be seen, while Table 1 has the calculated relative risk for each combination of diet and APOE genotype. For the force of natural selection on the APOE gene, we observed a positive selection for the \( e2 \) (mean frequency change per generation 3.352×10^{-3}, 95% CI 2.661×10^{-4} to 4.043×10^{-4}) and \( e3 \) (mean 2.785×10^{-4} 95% CI 1.809×10^{-4} to 3.761×10^{-4}) alleles against \( e4 \) (mean = 6.137×10^{-4} 95% CI = 6.946×10^{-4} to 5.327×10^{-4}), with the \( e2 \) frequency increasing slightly more rapidly than that for \( e3 \), although there was no significant difference between them (Figure 3).

Considering that only a fraction of the population will suffer from APOE related diseases, a variant of the model with 30% of the total mortality attributed to CVD was also tested, again, the results pointed to a positive selection of the \( e2 \) and \( e3 \) alleles (\( e2 \) mean 0.846×10^{-4}, 95% CI 0.497×10^{-4} to 1.195×10^{-4}; \( e3 \) mean 0.846×10^{-4}, 95% CI 0.316×10^{-4} to 1.377×10^{-4}) and a negative selection for the \( e4 \) allele (mean = 1.692×10^{-4}, 95% CI = −2.140×10^{-4} to −1.244×10^{-4}) which, as expected, was approximately one third of the change observed previously.

We performed a number of simulation runs with a variety of different starting conditions both to test the sensitivity of the model and analyse the likely changes under different lifestyle choices. The description of all the simulation variants used with their mean life expectancy at birth can be seen in Table 2. The changes in the distribution of the modifiable risk factors explored, except the eradication of smoking, were relatively modest and considered possible to occur, if not already present in certain subgroups of western populations. It is important to note the unexpected decrease in mean lifespan when heavy alcohol consumption is lowered and the slight increase in longevity when non- and very light drinkers decrease. This effect is due to the U-shaped relationship between alcohol consumption and health and the re-arrangement of the population distribution across the categories of risk.

The estimates for the mean frequency change for each APOE allele per generation, under all the different models considered assuming at the same time that only 30% of the population is affected by any APOE related diseases, are presented in Table 3. As

![Figure 1. Differences in the distribution of mortality for the three APOE genotype classes.](https://example.com/figure1.png)

These are estimated using a proportional hazard model as described in the methods. The \( e2 \) carriers have an OR of 0.9 and survive a year more than the \( e3e3 \) carriers. In contrast, \( e4 \) carriers have a decreased lifespan of 1.2 years due to an OR of 1.13 compared to the reference genotype.

doi:10.1371/journal.pone.0010022.g001
Selection on Apolipoprotein E

![Figure 2. Inferred interaction between APOE genotype and diet.](image)

Diet categories are considered on an integer scale from 1 to 5 with Heart Healthy, Light Eating, Wine and Moderate Eating, High Fat, and Empty Calorie clusters [44,45]. We assumed that for a hypothetical diet category 0 there is no distinction in risk between the genotypes.

doi:10.1371/journal.pone.0010022.g002

can be seen in the table, the selection against the e4 allele and the increase in the frequency of the e3 and e2 alleles are robust and remained significant under all the different versions of the simulation. A notable exception was the loss of significance for the e1 allele, which is likely to be under the action of natural selection. According to the results obtained, the e2 and e3 alleles are increasing with each successive generation at the expense of the e4 allele, which is slowly being removed from the population. The selection against the e4 allele was found to be robust, despite changes to the initial conditions of the simulation. We failed to find any significant difference between the increase in frequency of the e2 and e3 alleles, denoting their very similar effects on survival. This accords with meta-analysis studies which have found it difficult to identify any statistically significant risk differences between the carriers of e2 allele and the reference e3e3 genotype, unless a large number of cases is available, mainly due to the low frequency of e2 in the populations considered [15].

Despite the difficulty in observing and measuring the force of natural selection as revealed by change of allele frequencies in specific polymorphisms, especially in humans, alternative methods are available to obtain evidence for its action indirectly. APOE seems to be one of the least variable human genes studied, despite the average neutral mutation rate [59]. Comparing the chimpanzee and human APOE genes, revealed that e4 allele is probably the ancestral allele [59]. Indeed, most great apes carry only the e4 allele, although a change of arginine to threonine at position 61 gives this isoform an affinity similar to the human e3 [40,60,61]. It is unclear when the human e4 allele arose, but a comparison between coding changes in orthologous genes in five species (Mouse, Rat, Dog, Chimpanzee and Human) showed positive selection for APOE along the hominid lineage [62]. Using the haplotypes defining the three genotypes in a coalescent model, Fullerton et al [59] proposed that e3 diverged from the e4 haplotype around 200,000 years ago and that e3 is showing evidence of increasing in frequency, relative to e4. Assuming a constant selective pressure and a generation time of ~20 years, given an e3 frequency of around 0.75, a very simple calculation shows that the required increase of e3 per generation is $7.5 \times 10^{-5}$, close to the order of magnitude predicted from the simulation for the current change of the allele frequency. Using a similar argument, we can further propose that since e2 and e3 alleles seem to be under a similar pressure from natural selection in our simulation and e2 has a lower frequency in human populations, e2 arose later than e3. Again, analysis of APOE haplotypes using a reduced median network revealed that e2 was derived from the e3 allele within the last 80,000 years [59].

Assuming a simple and uniform selection pressure on APOE is rather over simplistic. Current theories suggest a two step evolution of the locus, one to explain the early evolution of the

### Table 1. Calculated relative risk for each combination of diet category as described by Miller et al (2001) and APOE genotype.

| APOE genotype | 22    | 23    | 33    | 34    | 44    |
|---------------|-------|-------|-------|-------|-------|
| Hypothetical diet 0 | 0.944 | 0.944 | 0.944 | 0.944 | 0.944 |
| Heart Healthy | 0.979 | 0.981 | 0.984 | 0.987 | 0.990 |
| Light Eating | 1.083 | 1.093 | 1.104 | 1.116 | 1.129 |
| Wine and Moderate Eating | 1.256 | 1.280 | 1.304 | 1.332 | 1.360 |
| High Fat | 1.499 | 1.541 | 1.584 | 1.633 | 1.683 |
| Empty Calorie | 1.811 | 1.877 | 1.944 | 2.020 | 2.099 |

doi:10.1371/journal.pone.0010022.t001
human alleles and a second step to account for their current global frequencies. Despite the very sparse evidence for both, and their rather conflicting views for the selection pressures applied, the proposed mechanisms are interesting and reveal the importance of APOE in human evolution. Finch and Stanford [63] and Finch and Sapolsky [60] suggested that APOE is a “meat adaptive” gene that permitted increased consumption of animal tissue during hominid evolution while conferring resistance to the associated risks such as hypercholesterolemia and infections, allowing for the extension of human lifespan. In their view, this increase in meat eating provided humans with a solution to seasonal deficits of micronutrients. In addition, relatively safe consumption of nutrient dense meat is believed to have lead to an increase of body size, without restricting activity, and conferred the extra energy required for the cerebral expansion of early humans [64,65]. On the other hand, the frequencies of the three common alleles vary widely between populations of different ethnic backgrounds. In general $e^3$ is globally the most common allele, with Africans having the lowest allele frequency ($\leq 70\%$) compared to Caucasians (70–80%) and East Asians ($\geq 80\%$) [66]. In Europe, and between populations of Caucasian descent, there is a north-to-south gradient of decreasing $e^4$ frequency opposite to the spread of agriculture [67,68]. This has led Corbo and Scacchi [69] to suggest that APOE is a thrifty gene, with the $e^4$ allele being advantageous under seasonal periods of starvation, due to its elevating effect of cholesterol which would otherwise be too low, but detrimental in areas where carbohydrates were readily available, such as those with a long history of agriculture, thus the north-to-south pattern observed in Europe. Doubts have been raised about the lack of seasonal starvation in food producing societies [70], while the hypothesised selection for $e^4$ in pre-agricultural times is opposite to the one suggested by Finch and Sapolsky [60,63].

Although it is clear that variation in the APOE locus has been, and probably still is, under the control of selection, the fact that its main effects are beyond the reproductive lifespan of humans, has posed an interesting puzzle for the mechanisms through which selection is applied. Finch and Sapolsky [60,71] suggested that the spread of $e^3$ in human populations is due to the effect of a mother’s late survival on the fitness of her offspring. The grand-mother hypothesis, as it is known, is centred on the idea that the mother-child food sharing seen among hunter-gatherers may allow post-reproductive grandmothers to enhance their daughters’ fertility, thus elevating their own fitness and increasing the selection for

![Figure 3. Change in frequency of the three APOE alleles with each generation.](image)

**Table 2.** The simulated average lifespan in years of each variant considered in the absence of non APOE related mortality.

| Run      | Parameters                                      | Mean Lifespan |
|----------|-------------------------------------------------|---------------|
| Default  | Default                                        | 75.69         |
| Default 2| Default values with 30% CVD associated mortality| 75.69         |
| Diet 1   | Heart Healthy and Light Eating groups decreased by 20% | 74.93         |
| Diet 2   | High Fat and Empty Calorie groups decreased by 20% | 75.98         |
| Alcohol 1| Light and non-drinkers decreased by 25%        | 75.97         |
| Alcohol 2| Heavy drinkers decrease by 25%                 | 75.54         |
| Exercise 1| Entire population moderately active (2.5h/week) | 76.19         |
| Exercise 2| Most active groups decreased by 25%            | 75.41         |
| Smoking 1| Non-smokers decreased by 25%                   | 75.57         |
| Smoking 2| No smokers in the population                   | 82.95         |

doi:10.1371/journal.pone.0010022.t002
long postmenopausal lifespan [72,73,74]. Finch and Sapolsky [60,71] argued that this evolutionary advantage will have caused selection for an \textit{APOE} allele that will have delayed neuropathology and mortality even if its effects are evident later in life (for counterarguments see [75,76]). Charlesworth [77], on the other hand, proposed that \textit{APOE} is an example of balanced polymorphism with the variation in the locus maintained due to the antagonistic effects of the alleles, where increased late life risk is offset by advantages in younger ages, in accordance with the antagonistic pleiotropy theory of ageing. Martin [76] extended the idea of antagonistic pleiotropy for \textit{APOE} suggesting that the $e^4$ allele could be advantageous in cases of infections with pathogens requiring host lipids for survival. Parasites like \textit{Trypanosome brucei} not able to carry out de novo liposynthesis may depend upon host LDL for acquisition of lipids, with the high affinity of $e^4$ for some lipids actually hindering their uptake by the pathogen. Our model indicates that, although these hypotheses might work at a secondary level, they are not strictly necessary to explain the selection of \textit{APOE} alleles. The direct effect of \textit{APOE} on the mortality or morbidity of the population is sufficient, at least in contemporary Western populations, to produce a selection differential between the three alleles. The simulation predicts that the $e^3$ and $e^2$ alleles are driven towards fixation, a conclusion supported by the Fullerton et al [59] observations for the mutation at site 3937, characterising the $e^3$ allele. This simulation study is considering the pressure of natural selection on a population with demographics equal, or similar, to what is found in contemporary populations. It will be wrong to extrapolate our results in order to interpret the early evolution of \textit{APOE}, since modern populations experience a very different environment compared to our early ancestors. Nevertheless, our basic conclusion that certain genetic polymorphisms, such as \textit{APOE}, exhibiting their main effect later in life can still be under the action of natural selection, still holds. Interestingly, analysis of a Danish cohort of individual born between 1895 and 1899, showed that carriers of the $e^4$ allele had increased early mortality compared to those carrying the $e^2$ allele as suggested by our simulation [78]. It is currently unclear how far back our results can

| Run       | APO E allele | Mean change of allele frequency per generation | 95% confidence intervals for the mean |
|-----------|-------------|-----------------------------------------------|---------------------------------------|
| Default   | $e^2$       | 0.0003352                                     | 0.0002661                             |
|           | $e^3$       | 0.0002785                                     | 0.0001809                             |
|           | $e^4$       | $-0.0006137$                                  | $-0.0006946$                          |
| Default 2 | $e^3$       | 0.0000846                                     | 0.0000497                             |
|           | $e^4$       | $-0.0001692$                                  | $-0.0001244$                          |
| Diet 1    | $e^3$       | 0.0000900                                     | 0.0000547                             |
|           | $e^4$       | $-0.0001798$                                  | $-0.0001340$                          |
| Diet 2    | $e^3$       | 0.0000643                                     | 0.0000130                             |
|           | $e^4$       | $-0.0001453$                                  | $-0.0001000$                          |
| Alcohol 1 | $e^3$       | 0.0000836                                     | 0.0000524                             |
|           | $e^4$       | $-0.0001654$                                  | $-0.0001226$                          |
| Alcohol 2 | $e^3$       | 0.0000588                                     | 0.0000074                             |
|           | $e^4$       | $-0.0001436$                                  | $-0.0000991$                          |
| Exercise 1| $e^3$       | 0.0000895                                     | 0.0000555                             |
|           | $e^4$       | $-0.0001524$                                  | $-0.0001088$                          |
| Exercise 2| $e^3$       | 0.0000637                                     | 0.0000115                             |
|           | $e^4$       | $-0.0001725$                                  | $-0.0001279$                          |
| Smoking 1 | $e^3$       | 0.0000974                                     | 0.0000383                             |
|           | $e^4$       | $-0.0002256$                                  | $-0.0001746$                          |
| Smoking 2 | $e^3$       | 0.0000259                                     | 0.0000057                             |
|           | $e^4$       | $-0.0000665$                                  | $-0.0000393$                          |

doi:10.1371/journal.pone.0010022.t003
be considered as valid, but given the early age (>35 years of age) that the APOE effects on mortality can be seen in the Danish cohort, we believe that this can extend in our recent history.

In our simulation we cannot account for any pleiotropic effects of APOE. We described the effects of the gene on lifespan only through its relation to elevated risk for CVD. APOE has a much wider role than just lipid metabolism, including immunoregulation and susceptibility to infections [60,79,80,81]. We chose not to include these effects within the simulation since reliable quantitative data concerning their impact on lifespan do not exist. APOE, though its role in strophosogenesis, has also been shown to influence fertility in pre-industrial populations [82], old Italians (>80 years of age) [83] and North European males [84], with the e2 carriers having less children. The small number of e2 subjects together with the many cultural and socioeconomic factors affecting current human reproduction makes it difficult to say if this effect is indeed real and relevant in the model used. A limitation of our model was the lack of data for the interaction between the APOE genotypes and each of the modifiable risk factors considered, which forced us to use a very approximate method to calculate the unknown interactions. Further epidemiological work is required to elucidate the precise quantitative relationship of the APOE gene with the parameters of the model so that the predictions can be made more accurate. While, for simplicity, we assumed that the rest of the risk factors are independent of each other, although it has been recognised that these tend to cluster, to some degree, in individuals of high risk [85].

Despite some limitations, our results indicate that although the main effect of APOE is seen after the end of the reproductive lifespan of humans, the relatively rare events of early mortality or morbidity is enough to produce selection against the e4 allele. Considering these findings in the light of the evolutionary theory of ageing and longevity, particularly as regards the ‘mutation accumulation’ of late acting deleterious alleles, we can identify an important perspective on this concept. Since most longevity-associated genes, such as APOE, will cause a distribution around a mean for the corresponding mortality, evolution will rarely, if ever, manage to push deleterious gene action entirely beyond the reach of selection, although as envisaged by Medawar there will be a continual selection pressure to postpone such action to later and later ages. The gradual postponement of the detrimental gene effect will take place simultaneously in a wide array of mortality associated genes. Any single mortality distribution exposed significantly more often to the action of selection will be moved towards older ages faster, until it reaches a balance with all the other detrimental genotypes. This process will continue to drive the longevity of the organism until the age-related mortality becomes non-significant in relation to the age-independent mortality. In this respect the evolution of longevity, although based on selection of individual genes, is more of an evolutionarily coordinated transfiguration of the whole genome to achieve the maximum lifespan in the given conditions as hypothesized by Hamilton [56].

Author Contributions
Conceived and designed the experiments: FD TBLK. Performed the experiments: FD. Analyzed the data: FD. Contributed reagents/materials/analysis tools: FD. Wrote the paper: FD TBLK.

References
1. Perls TT, Bubrick E, Wager CG, Vigg J, Kruglyak L (1996) Siblings of centenarians live longer. Lancet 351: 1560–1560.
2. Hjelmsberg JV, Iachine I, Skytte A, Vaupel JW, McGuire M, et al. (2006) Genetic influence on human lifespan and longevity. Human Genetics 119: 319–321.
3. Hansdottir AM, McGuire M, Holm NV, Sorensen TIA, Harvald B, et al. (1996) The heritability of human longevity: A population-based study of 2872 Danish twin pairs born 1870–1900. Human Genetics 97: 319–323.
4. Ljungquist B, Berg S, Lanke J, McClean GE, Pedersen NL (1998) The effect of genetic factors for longevity: A comparison of identical and fraternal twins in the Swedish twin registry. Journals of Gerontology Series A - Biological Sciences and Medical Sciences 53: M441-M446.
5. Cournia U, Legay JM, Schachter F (2000) Evidence of sex-linked effects on the inheritance of human longevity: a population-based study in the Valerie valley (French Jura), 18–20th centuries. Proceedings of the Royal Society of London Series B-Biological Sciences 267: 1021–1025.
6. Godmunksson H, Guðbjartsóson DF, Kong A, Guðbjartsóson Þ, Friðge M, et al. (2005) Inheritance of human longevity in Iceland. European Journal of Human Genetics 8: 743–749.
7. Skytte A, Pedersen NL, Kaprio J, Stazi MA, Hjelmsberg JV, et al. (2003) Longevity studies in GenomEUtwin. Twin Research 6: 448–454.
8. Wernette-Hammond CE (1989) Glycosylation of human apolipoprotein E. The carbohydrate attachment site is threonine 194. The Journal of Biological chemistry 264: 9094–9101.
9. Hagberg JM, Wihlund KR, Ferrell RE (2000) APO E gene and gene-environment effects on plasma lipoprotein- lipid levels. Physiological Genomics 4: 101–108.
10. van Boekxmeer FM (1994) Apoe and Ace genes - Impact on human longevity. Nature Genetics 4: 4–5.
11. Smith JD (2002) Apolipoproteins and aging: emerging mechanisms. Ageing Research Reviews 1: 345–365.
12. Kall MC Jr, Weisgraber KH, Mahley RW (1982) Human apolipoprotein E. The complete amino acid sequence. Journal of Biological Chemistry 257: 4171–4178.
13. Utermann G, Hardewig A, Zimmer F (1984) APOLIPOPROTEIN-E. Human Genetics 65: 237–241.
14. Song YQ, Stampfer MJ, Liu SM (2004) Meta-analysis: Apolipoprotein E genotypes and risk for coronary heart disease. Annals of Internal Medicine 141: 137–147.
15. Bennet AM, DR Angelantonio E, Ye Z, Wendey F, Dahlia A, et al. (2007) Association of apolipoprotein E genotypes with lipid levels and coronary risk. Jama-Journal of the American Medical Association 298: 1300–1311.
16. Zhou WD, Xu D, Peng XQ, Zhang QH, Jia JP, et al. (2008) Meta-analysis of APOE4 allele and outcome after traumatic brain injury. Journal of Neurotrauma 25: 279–290.
17. Dong LM, Wilson C, Wardell MR, Simmons T, Mahley RW, et al. (1994) Human Apolipoprotein-E. The Finnish Centenarians Study. Journals of lipid Research 39: 658–669.
18. Heinemann P, van der Sman-de Beer F, Moghaddam PH, Huijts P, Stalenhoef AFH, et al. (2009) The expression of type III hyperlipoproteinemia: involvement of lipolysis genes. European Journal of Human Genetics 17: 620–628.
19. Schachter F, Fauredelanef L, Guenot F, Rouger H, Frogel P, et al. (1994) Genetic associations with human longevity at the Apoe and Ace loci. Nature Genetics 6: 29–32.
20. Blanche H, Cabanne L, Salbatiou M, Thomas G (2001) A study of French centenarians: are ACE and APOE associated with longevity? Comptes Rendus De l’Academie Des Sciences Serie III-Sciences De La Vie-Life Sciences 324: 129–135.
21. Rea IM, Mc Dowell I, McMaster D, Smye M, Stout R, et al. (2001) Apolipoprotein E alleles in nonagenarian subjects in the Belfast Elderly Longitudinal Free-living ageing Study (BELFAST). Mechanisms of Ageing and Development 122: 1367–1372.
22. Friesen GB, Lohuja J, Geroldi C, Trabucchi M (2001) Longevity and the epsilon 2 allele of apolipoprotein E: The Finnish Centenarians Study. Journals of Gerontology Series A - Biological Sciences and Medical Sciences 56: M75–M78.
23. Frisoni GB, Louhija J, Geroldi C, Trabucchi M (2001) Longevity and the epsilon 2 allele of apolipoprotein E: The Finnish Centenarians Study. Journals of Gerontology Series A - Biological Sciences and Medical Sciences 56: M75–M78.
24. Lewis SJ, Brunner EJ (2004) Methodological problems in genetic association studies of longevity - the apolipoprotein E gene as an example. International Journal of Epidemiology 33: 962–970.
25. Wilson PWFMD, Myers RHP, Larson MGS, Ordovas JMP, Wolf PAMD, et al. (1994) Apolipoprotein E Alleles, Dyslipidemia, and Coronary Heart Disease: The Framingham Offspring Study. Jama-Journal of the American Medical Association 272: 1666–1671.
Selection on Apolipoprotein E

26. Laloue C, Schaefer EJ, Cupples LA, Wilson PWF, Levy D, et al. (2001) Apolipoprotein E genotype and cardiovascular disease in the Framingham Heart Study. Atherosclerosis 154: 529–537.

27. Stengard JH, Pelikanen J, Elmfeldt D, Nisinen A, Sing CF (1996) Genotypes with apolipoprotein epsilon 4 allele predict cardiovascular events: a longitudinal study of elderly Finnish men. Human Genetics 97: 677–684.

28. Wilson PWF, Schaefer EJ, Larson MG, Orvos JM (1996) Apolipoprotein E alleles and risk of coronary disease - A meta-analysis. Arteriosclerosis, Thrombosis and Vascular Biology 16: 1250–1255.

29. Gerdes LU, Jeune B, Ranberg KA, Nybo H, Vaugel JB (2000) Estimation of apolipoprotein E genotype-specific relative mortality risks from the distribution of genotypes in centenarians and middle-aged men: Apolipoprotein E genotype is a ‘frailty gene’, not a ‘longevity gene’. Genetic Epidemiology 19: 202–210.

30. Kolovesi G, Yamakoskouri N, Hatzivasiliou M, Malaks J, Daskalova D, et al. (2002) Association of apolipoprotein E polymorphism with myocardial infarction in Greek patients with coronary artery disease. Current Medical Research and Opinion 18: 110–114.

31. Strittmatter WJ, Saunders AM, Schmitz D, Pericakvance MA, et al. (1993) APOLIPOPROTEIN-E - HIGH-AVVIDITY BINDING TO BETA-AMYLOID AND INCREASED FREQUENCY OF TYPE-4 ALLELE IN LATE-ONSET FAMILIAL ALZHEIMER-DEISEASE. Proceedings of the National Academy of Sciences of the United States of America 90: 1977–1981.

32. Saunders AM, Strittmatter WJ, Schmechel D, Georgeheyslof PHS, Pericakvance MA, et al. (1993) ASSOCIATION OF APOLIPOPROTEIN-E ALLELE EPSILON-4 WITH LATE-ONSET FAMILIAL AND SPORADIC ALZHRHEIMERS-DISEASE. Neurology 43: 1467–1472.

33. Roses AD (1997) Apolipoprotein E, a gene with complex biological interactions in aging: review of the literature. Journal of Biological Chemistry 272: 170–183.

34. Huang YD (2006) Apolipoprotein E and Alzheimer disease. Neurology 66: 677–684.

35. de Carli CMD, Miller BLMD, Swan GEP, Reed TP, Wolf PAMD, et al. (2001) Association of apolipoprotein E polymorphism with myocardial infarction of genotypes in centenarians and middle-aged men: Apolipoprotein E gene is a ‘frailty gene’, not a ‘longevity gene’. Genetic Epidemiology 19: 202–210.

36. Charlesworth B (1994) Evolution in Age-structured Populations. Cambridge: Cambridge University Press.

37. Hawkes K, Grandmothers and the evolution of human longevity; 2002 Apr. 24–26. Washington, D.C.: Amer Inst Nutrition, 389–398.

38. Mahley RW, Rall SC (2000) Apolipoprotein E: Far more than a lipid transport protein. Annual Review of Genomics and Human Genetics 1: 507–537.

39. Hatch SL, Yue H, Schaefer EJ, Larson MG, Orvos JM (2000) Apolipoprotein E phenotype determines the effect of alcohol on blood pressure in middle-aged men. Atherosclerosis 154: 529–537.

40. Fullerton SM, Clark AG, Weiss KM, Nickerson DA, Taylor SL, et al. (2000) Association of apolipoprotein E polymorphism with myocardial infarction of genotypes in centenarians and middle-aged men: Apolipoprotein E gene is a ‘frailty gene’, not a ‘longevity gene’. Genetic Epidemiology 19: 202–210.

41. Finch CE, Finch CE (2000) Alzheimer’s disease, the role of heritability, and the evolution of aging in humans. Quarterly Review of Biology 75: 3–50.

42. Corbo RM, Scachar S, Murrell I, Mulas G, Alfano G (1995) APOLIPOPROTEIN-E POLYMORPHISM IN ITALIAN INDIANS-RELATIONSHIP BETWEEN ALLELE EPSILON-4 AND CORONARY HEART DISEASE. Atherosclerosis 112: 85–90.

43. Finch CE, Stanford CB (2004) ‘Thrifty genotypes and thrifty phenotypes; 2004 Jul 05–08. Dublin, IRELAND: Cabi Publishing, 128–129.

44. Finch CE, Finch CE (2000) Alzheimer’s disease and some speculations about the evolution of its modifiers. Alzheimer’s Disease: A Compendium of Current Theories. pp 99–103.

45. Hawkes K (2004) Human longevity - The grandmother effect. Nature 428: 120–129.

46. Hawkes K (2004) Grandmothers and the evolution of human longevity; 2002 Apr. Buffalo,NY: Wiley-Liss, 380–400.

47. Hawkes K, O’Connell JF, Jones NGB, Alvarez H, Charnov EL (1998) Grandmothering, menopause, and the evolution of human life histories. Proceedings of the National Academy of Sciences of the United States of America 95: 1336–1339.

48. Finch CE, Finch CE (2000) Alzheimer’s disease and some speculations about the evolution of its modifiers. Alzheimer’s Disease: A Compendium of Current Theories. pp 99–103.

49. Finch CE, Finch CE (2000) Alzheimer’s disease and some speculations about the evolution of its modifiers. Alzheimer’s Disease: A Compendium of Current Theories. pp 99–103.

50. Finch CE, Finch CE (2000) Alzheimer’s disease and some speculations about the evolution of its modifiers. Alzheimer’s Disease: A Compendium of Current Theories. pp 99–103.

51. Finch CE, Finch CE (2000) Alzheimer’s disease and some speculations about the evolution of its modifiers. Alzheimer’s Disease: A Compendium of Current Theories. pp 99–103.
84. Gerdes LU, Gerdes C, Hansen PS, Klausen IC, Faergeman O (1996). Are men carrying the apolipoprotein epsilon 4- or epsilon 2 allele less fertile than epsilon 3 epsilon 3 genotypes? Human Genetics 98: 239–242.

85. Twisk JWR, Kemper HCG, van Mechelen W, Post GB (2001). Clustering of risk factors for coronary heart disease: The longitudinal relationship with lifestyle. Annals of Epidemiology 11: 157–165.