

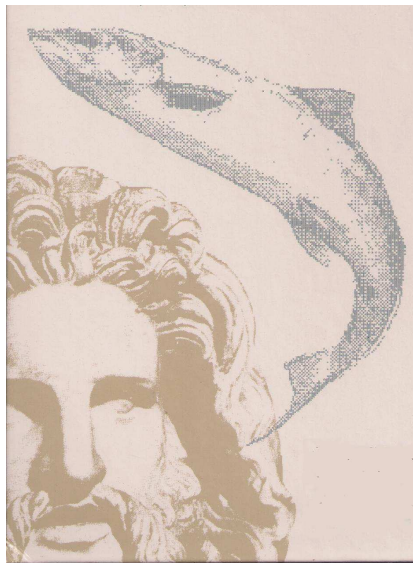
Alleles for Longevity

Kenneth W. Wachter

Departments of Demography and Statistics
University of California at Berkeley
Demography Brown-Bag Luncheon Talk

1 September 2010

Alleles for Longevity: Estimating Demographic Effects



An update on research presented on 7 June 2010 at the annual conference of the Berkeley Population Center.

- 1 Genes, GWAS, Worms, and FOX
- 2 Two Questions for Demographers .
- 3 German Cohort Lifetables from Brass Logits
- 4 Six Equations in Six Unknowns
- 5 Bias
- 6 Dominance and Hazard Structure

Thanks to the Miller Institute for Basic Research in Science and to the National Institute on Aging (Grant AG022500-06A1).

1. Genes, GWAS, Worms, and FOX

Volumes from the National Research Council Committee on Population,
courtesy of Richard Suzman and the National Institute on Aging:

1997 Between Zeus and the Salmon

2001 Cells and Surveys

2008 Biosocial Surveys

GWAS = Genome Wide Association Studies

HRS = Health and Retirement Study

(Expert meeting 23 and 24 September 2010)

The Elegant Worm *C. Elegans*

The nematode worm
Caenorhabditis elegans



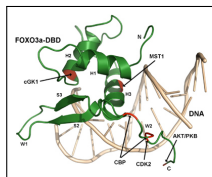
Cynthia Kenyon (2010) “The Genetics of Ageing”, Nature, 464: 504–512.

Insulin Signalling Pathways

Human Forkhead Box 03A Gene: “FOX03A”

Bradley J. Willcox et al. (2008) “FOXO3A genotype is strongly associated with human longevity”, PNAS 105:13987–13992.

The Forkhead Box Transcription Factor



Association of *FOXO3A* variation with human longevity confirmed in German centenarians

Friederike Flachsbart^a, Amke Caliebe^b, Rabea Kleindorp^a, H el ene Blanch e^c, Huberta von Eller-Eberstein^d, Susanna Nikolaus^d, Stefan Schreiber^{a,1,2}, and Almut Nebel^{a,1}

^aInstitute of Clinical Molecular Biology, ^bInstitute of Medical Informatics and Statistics, and ^cPopgen Biobank, Christian-Albrechts-University, 24105 Kiel, Germany; and ^dFondation Jean Dausset, Centre d'Etude du Polymorphisme Humain, 75010 Paris, France

Edited by Cynthia J. Kenyon, University of California, San Francisco, CA, and approved December 31, 2008 (received for review September 25, 2008)

The human forkhead box O3A gene (*FOXO3A*) encodes an evolutionarily conserved key regulator of the insulin-IGF1 signaling pathway that is known to influence metabolism and lifespan in model organisms. A recent study described 3 SNPs in the *FOXO3A* gene that were statistically significantly associated with longevity in a discovery sample of long-lived men of Japanese ancestry [Willcox et al. (2008) *Proc Natl Acad Sci USA* 105:13987–13992]. However, this finding required replication in an independent population. Here, we have investigated 16 known *FOXO3A* SNPs in an extensive collection of 1,762 German centenarians/nonagenarians and younger controls and provide evidence that polymorphisms in this gene were indeed associated with the ability to attain exceptional old age. The *FOXO3A* association was considerably stronger in centenarians than in nonagenarians, highlighting the importance of centenarians for genetic longevity research. Our study extended the initial finding observed in Japanese men to women and indicates that both genders were likely to be equally affected by variation in *FOXO3A*. Replication in a French centenarian sample generated a trend that supported the previous results. Our findings confirmed the initial discovery in the Japanese sample and indicate *FOXO3A* as a susceptibility gene for prolonged survival in humans.

aging | forkhead box O3A | genetic association study | long-lived individuals

Life expectancy in humans is influenced by various environmental and genetic factors. Approximately 25–32% of the overall variation in adult lifespan is accounted for by genetic differences that become particularly important for survival after the age of 60 (1–5). The mechanisms influencing lifespan have been intensively studied in *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, or *Drosophila melanogaster*, and hundreds of genetic variants that lead to life extension in model systems have been identified (6–8). The success in finding lifespan-control genes in lower organisms has also motivated efforts to search for corresponding genes in humans. However, to date variation in only 1 gene, which codes for apolipoprotein E (APOE), has been found to be consistently associated with survival in various populations. Although numerous case-control candidate studies have been performed and associations of the longevity phenotype with biologically plausible genes have been described

notypes in a discovery sample of long-lived Americans of Japanese ancestry. However, this finding required replication in an independent population. Here, we have investigated 16 known SNPs, which capture the majority of the variation in *FOXO3A* via its common haplotypes, in an extensive collection of 1,762 German centenarians, nonagenarians, and younger controls and provide evidence that polymorphisms in this gene are indeed associated with the ability to attain exceptional old age. Our findings confirmed the initial discovery in the Japanese sample and thus support *FOXO3A* as a susceptibility gene for prolonged survival in humans.

Results

In the present study, 16 polymorphisms in *FOXO3A* were analyzed for association with the human longevity phenotype (Tables 1 and 2). The tested SNPs are spaced across the entire gene region, including the promoter (Fig. 1) and capture the majority of its allelic variation by haplotype tagging. All SNPs were in Hardy-Weinberg equilibrium (HWE) in the control population. For the association analyses, we applied an established longevity study design (13, 14) by comparing German long-lived individuals (LLI; subset A; $n = 1,031$; aged 95–110 years) and a centenarian subset (subset B; $n = 388$) to appropriately matched younger controls ($n = 731$; aged 60–75 years). All markers were subjected to allelic case-control comparisons (CCA) by using the entire LLI sample (subset A) and the centenarian subset (subset B). For subset A, single-marker analysis revealed 4 SNPs with nominally significant P_{CCA} values (Table 1). For the centenarians (subset B), 11 SNPs showed significant association (Table 2). Although subset B is smaller in size and therefore expected to have less power than the overall LLI sample, the significance level was more pronounced in the centenarians and revealed a stronger effect as reflected in the odds ratios (ORs) (Tables 1 and 2). The 3 top-ranking *FOXO3A* markers in subset B (rs3800231, rs9400239, and rs479744) passed correction for multiple testing (Bonferroni-adjusted significance threshold = 0.0016; for 2×16 tests). Because this adjustment did not take into account the strong linkage disequilibrium (LD) between the investigated markers (Fig. 1), the obtained threshold must be regarded as conservative. The results from the comparison of the genotypic data (CCG) are presented as additional information but they were not included in the initial

“Minor allele frequency distribution of rs2802288 in Germans by age groups”.
Table 6 of Flachsbart et al. (2009)

Age	n	Frequency
60-75	731	0.385
95-99	631	0.402
100-104	362	0.441
105-110	21	0.524

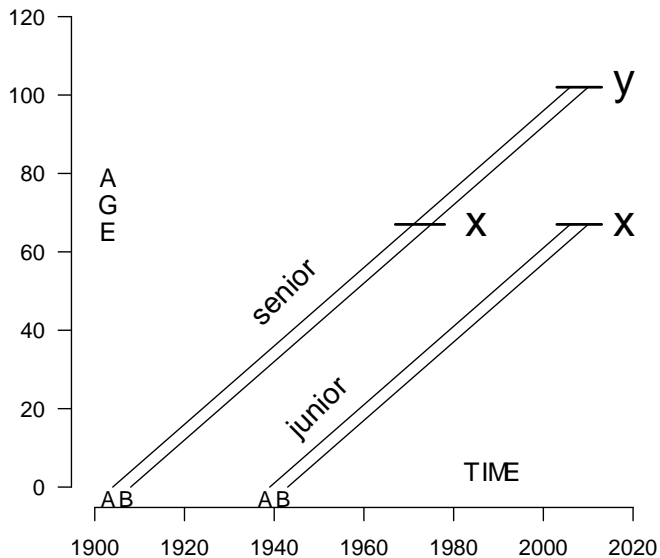
2. Two Questions for Demographers

- 1 Is there serious bias from birth cohort confounding?
- 2 Can we distinguish proportional from additive hazards?

Answers

- 1 No.
- 2 Almost.

Alleles for Longevity with Birth Cohort Confounding



Additive: (hazard for Bs) = η + (hazard for As)

Proportional: (hazard for Bs) = $(1 + \zeta) \times$ (hazard for As)

The hazard is minus the slope of the logarithm of survivorship.

- B, A, W : Carriers of allele B ; wild type A ; whole cohort W .
- s, j : Senior cohort s born around 1906;
Junior cohort j born around 1941.
- x, y : ages x (around 67) and y (around 102).
- ℓ and p : lifetable survivorship ℓ with unit radix;
proportion p of alleles within cohort.

$\ell_y(As)$ is the unknown survivorship to age y of carriers of A in the senior cohort.

$p_x(Bj)$ is the observed proportion at age x of carriers of B in the junior cohort.

- Ratios of survivorships equal ratios of allele proportions, since odds ratios do not depend on marginals (Karl Pearson, 1904):

$$\rho = \frac{p_y(B)}{p_x(B)} / \frac{p_y(A)}{p_x(A)} = \frac{\ell_y(B)}{\ell_x(B)} / \frac{\ell_y(A)}{\ell_x(A)}$$

- With proportional effects on hazards, the unobserved survivorship to x for the senior cohort differs from the observed survivorship to x for the junior cohort, yielding biased estimates of odds ratios and effect sizes.
- With additive effects on hazards, under our key assumption that allele frequencies at birth are the same across cohorts, cohort differences cancel out.

$$\begin{aligned}\rho &= \frac{p_y(B)}{p_x(B)} / \frac{p_y(A)}{p_x(A)} \\ &= \frac{p_0(B)\ell_y(B)/\ell_y(W)}{p_0(B)\ell_x(B)/\ell_x(W)} / \frac{p_0(A)\ell_y(A)/\ell_y(W)}{p_0(A)\ell_x(A)/\ell_x(W)} \\ &= \frac{\ell_y(B)}{\ell_x(B)} / \frac{\ell_y(A)}{\ell_x(A)}\end{aligned}$$

Odds ratios do not depend on marginals: Karl Pearson (1904).

As before, we have

Additive: (hazard for Bs) = η + (hazard for As)

Proportional: (hazard for Bs) = $(1 + \zeta) \times$ (hazard for As)

$$\eta = \frac{-1}{y-x} \log(\rho)$$

$$\zeta = \frac{-1}{\log(\ell_x(A)/\ell_y(A))} \log(\rho)$$

With proportional hazards,

$$p_x(A) = \frac{p_0(A)\ell_x(A)}{p_0(A)\ell_x(A) + p_0(B)\ell_x(A)(\ell_x(A))^\zeta}$$

When cohort survivorships differ, the unobserved value $p_x(A, s)$ for the senior cohort is not the same as the observed value $p_x(A, j)$ for the junior cohort. Bias occurs despite our key assumption that the frequencies of A and B at birth are the same regardless of cohort, then with additive hazards the proportion $p_x(A)$ is the same for the senior as for the junior cohort.

With additive hazards, under our key assumption, cohort differences cancel out:

$$\ell_x(B) = \ell_x(A)e^{-\eta x}$$

With additive hazards we have

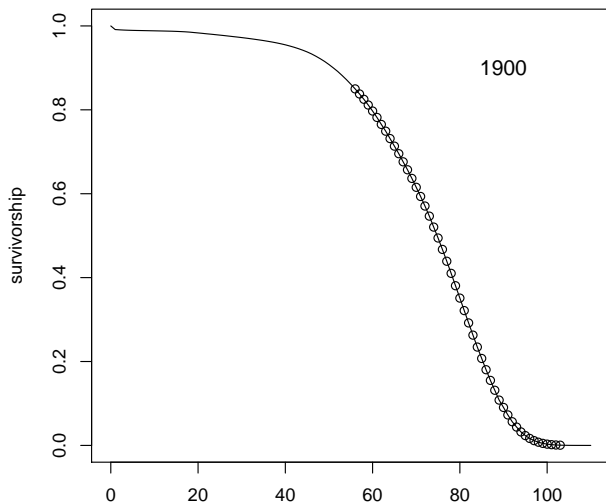
$$p_x(A) = \frac{p_0(A)\ell_x(A)}{p_0(A)\ell_x(A) + p_0(B)\ell_x(A)e^{-\eta x}}$$

Note that $\ell_x(A)$ cancels between numerator and denominator.

3. German Cohort Lifetables from Brass Logits

- The Human Mortality Database only has mortality rates for our senior cohort for ages 56 through 103 and for our junior cohort for ages 26 through 73. (I use the five-year-wide cohort mortality rates).
- There are no l_x values.
- For Brass Logits, we need l_x values.
- Solution:
 - Choose a trial value of l_{56} and compute l_x from 56 to 103.
 - Fit a Brass logit lifetable, using the German 2008 period lifetable as the Brass standard.
 - Compare the predicted value of l_{56} to the trial value.
 - Choose the (unique) fitted table which makes trial value match the prediction.
 - Repeat with the junior cohort.

Fitted Senior Whole-Cohort Survivorships



4. Six Equations in Six Unknowns

- Two equations can be solved for the two unknowns (1) proportion of B's at birth and (2) effect size when we already have values for the survivorships for wild-type A's. We match the predicted ratio of B's to A's to the observed ratios for the two ages x and y .
- Three equations let us obtain the unknown survivorships for wild-type A's from the observed whole-cohort survivorships, when we already have values for the proportion of B's at birth and the effect size. We need three survivorships: (1) to age x in the junior cohort; (2) to age x in the senior cohort; (3) to age y in the senior cohort;
- One equation expresses the unknown bias in the odds ratio in terms of the other five unknowns.

$$\frac{(p_0) \psi \psi^\zeta}{(1 - p_0) \psi} = \frac{p_x(Bj)}{p_x(Aj)} \quad (1)$$

$$\frac{(p_0) \theta \theta^\zeta}{(1 - p_0) \theta} = \frac{p_y(Bs)}{p_y(As)} \quad (2)$$

$$\phi = \ell_x(Ws) / (1 - p_0 + p_0 \phi^\zeta) \quad (3)$$

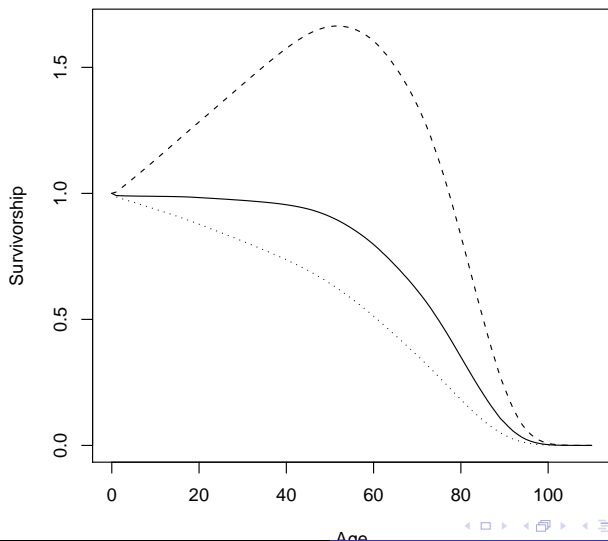
$$\psi = \ell_x(Wj) / (1 - p_0 + p_0 \psi^\zeta) \quad (4)$$

$$\theta = \ell_y(Ws) / (1 - p_0 + p_0 \theta^\zeta) \quad (5)$$

$$\rho = \left[\frac{\psi^\zeta}{\phi^\zeta} \right] \left[\frac{p_y(Bs)}{p_y(As)} \right] / \left[\frac{p_x(Bj)}{p_x(Aj)} \right] \quad (6)$$

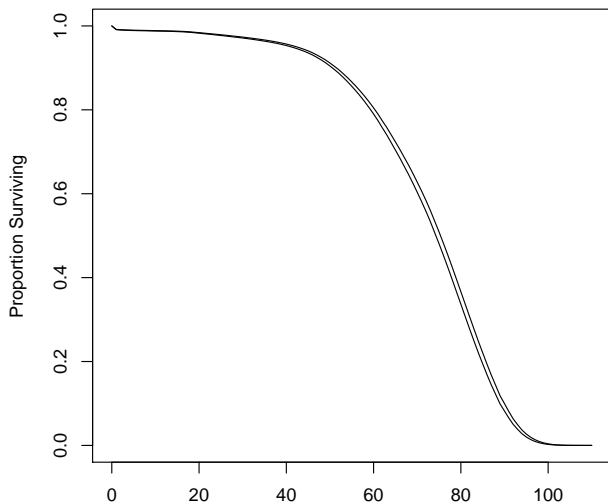
- η or ζ is the effect size in the adopted model; hazards for carriers of B equal hazards for carriers of A plus η in the additive model and times $1 + \zeta$ in the proportional model;
- p_0 equals $p_0(Bs) = p_0(Bj)$ is the common value of the proportion of carriers of B at birth assumed to be the same in the senior and junior cohorts;
- ϕ equals $\ell_x(As)$ probability of surviving from birth to age x of carriers of A in the senior cohort;
- ψ equals $\ell_x(Aj)$, the probability of surviving from birth to age x of carriers of A in the junior cohort;
- θ equals $\ell_y(As)$ the probability of surviving from birth to y of carriers of A in the senior cohort;
- ρ is the ratio of probabilities of surviving from x to y for carriers of B in the numerator compared to carriers of A in the denominator in the senior cohort;

Fitting Survivorships with Additive Hazards



- Proportional hazards give effect $\zeta = -.0779$.
- Additive hazards give effect $\eta = -.0157$.
- Bias from using period data (from junior cohort) for age group 60 to 75 in place of true senior cohort data is negligible in this case.
- With $p_x(Bj)$ equal to the observed value of 0.3851, we have $p_x(Bs) = 0.3869$, about a tenth of a standard error different.
- From the equations, for a 5% bias in the odds ratio, senior cohort survivorship would have to be only 0.517 of junior cohort survivorship with effect sizes as measured in this study.

No Bias is Good News and Also Bad News



“Minor allele frequency distribution of rs2802288 in Germans by age groups”.

Table 6 of Flachsbart et al. (2009)

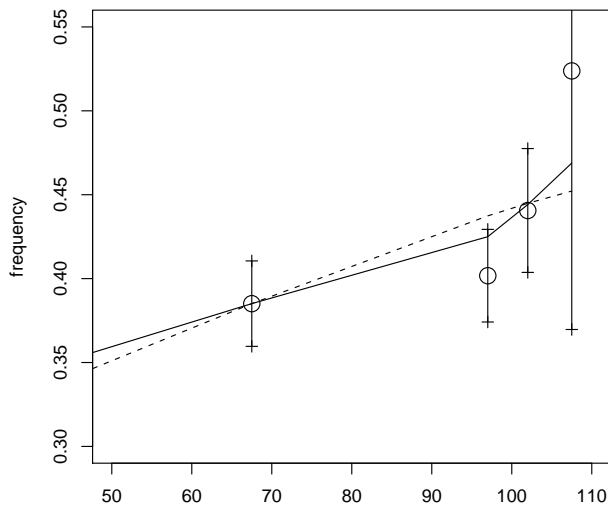
Age	n	Frequency	Carriers*
60-75	731	0.385	0.6224
95-99	631	0.402	0.6259
100-104	362	0.441	0.7127
105-110	21	0.524	0.8095

*Estimates by KWW of proportion of carriers from matching P_{CCG} test statistics quoted by Flachsbart et al.

Instead, we change the equations

- Assume Hardy-Weinberg Equilibrium at Birth for bb , ba , and aa ;
- Project each genotype. Then aggregate to proportion of alleles.
- Find parameters that fit proportions of alleles.

Comparing Proportional and Additive Hazards



What does the comparison show?

- Observed values by age are circles with 95
- Both dotted proportional fits and dashed additive fits have been forced by our estimation method to go through the observed value for 60 – 75 and the aggregate cell combining 100 – 104 with 105 – 110.
- Additive effect has to be fudged with age cutoff of 50 years.
- In a simulation with 1000 cases, letting the proportions of alleles vary around their observed values with binomial distributions, fitting effects to the simulated values, and comparing predictions for the nonagenarians, in 92% of cases the proportional hazard fit was better than the additive hazard fit.

- Birth cohort confounding looks less serious than it might be.
- Current sample sizes are approaching adequacy for inferences about age-specific genetic effects on survival
- The minor allele of rs2802288 in FOXO3A looks like an allele gene with small early-age effects and a rapid onset of effects at late ages.
- Alleles with effects of this kind, if numerous, would very important for mathematical formulations of the evolutionary theory of senescence.