Journal of Gerontology: BIOLOGICAL SCIENCES
1997, Vol. 52A, No. 1, B39–B47

Copyright 1997 by The Gerontological Society of America

Candidate Biomarkers of Aging: Age-Sensitive Indices of Immune and Muscle Function Covary in Genetically Heterogeneous Mice

Richard A. Miller,[1,2,3] Fred Bookstein,[2] Jack Van der Meulen,[2]
Sasha Engle,[1] Julie Kim,[1] Lisa Mullins,[1] and John Faulkner[2,4]

Departments of [1]Pathology and [4]Physiology, University of Michigan School of Medicine.
[2]Institute of Gerontology, University of Michigan.
[3]Department of Veterans Affairs Medical Center, Ann Arbor.

A longitudinal experiment was designed to test the hypothesis that individual mice differ in their aging rate and to validate candidate biomarkers proposed to measure the rate of aging. Mice were bred as the genetically heterogeneous progeny of a cross between CB6F1 mothers and C3D2F1 fathers. Half of the mice were fed ad libitum (AL group), and the other half were subjected to 60% calorie restriction (CR group). Each mouse was tested at about 9 months of age using age-sensitive tests of immune status, and then again at about 12 months of age using age-sensitive tests of muscle function. The data were then analyzed using the method of partial least squares to determine the combinations of test weights that maximize the covariance of the weighted sum of immune measures with the weighted sum of muscle function measures. Both AL and CR mice exhibited a statistically significant relation between the immune status tests and the muscle function tests. Maximal covariance was obtained with a set of weighting coefficients consistent with our working hypothesis: mice with high levels of CD4 memory T cells (which increase with age) also had relatively low levels of muscle strength and endurance. Low strength was associated with low CD8 cells in the AL mice, with high numbers of CD8 memory cells in the CR mice and with low CD3 cells in both diet groups. The partial least squares method generates composite indices of immune status and muscle function that can be evaluated as biomarkers of aging rate in these mice. Further work will be needed to assess whether these tests predict either longevity or the trajectory of change in other age-sensitive molecular and physiological traits.

THE development of a method for estimating biological age would have important theoretical and practical implications for gerontology. Mice, dogs, and human beings are presumed to age at different rates, because the signs of aging — biochemical and cellular changes, physiological declines, and exponential increases in the probability of disease — are typically seen in 2-year-old mice, 10-year-old dogs, and 70-year-old human beings. Despite the clear differences in aging rate across species, whether members of a given species age at different rates is not clear. Some authors argue plausibly that each individual will show an idiosyncratic pattern of physiological changes as he or she grows older, influenced by a unique combination of genetic susceptibilities, environmental influences, and stochastic events (Adelman, 1987; Costa and McCrae, 1995). According to this view it is not possible to assign a single "biological age" to any individual, because any two individuals will differ in a multitude of age-sensitive traits. Members of this philosophical school (Masoro, 1995) argue that aging is best viewed as a collection of independent processes that operate together, in parallel, to convert young mice and people into old ones.

An alternate perspective, stated in extreme form, would maintain that aging is a single physiological process that is coupled, more or less loosely, to an entire diverse range of age-dependent outcomes. This "single process" model makes the claim that age-associated changes in cell proliferation, collagen cross-linking, muscle function, protective immunity, disease risk, and a wide range of other phenotypes are influenced by a rate-setting mechanism whose nature is not yet known, and whose pace varies across and perhaps also within species. In this model age-sensitive traits are influenced both by the aging process and also by factors, including genetic predispositions and environmental variations, that may be unrelated to aging. This causal complexity presents a challenge: can one detect, experimentally, the "signal" of aging rate amid the "noise" generated by other factors that influence age-sensitive outcomes?

A critical test of the single process hypothesis would be to determine to what extent age-sensitive traits in one functional domain, such as muscle function, covary with traits in a second domain, such as immune status. The testable prediction of the single process hypothesis is that individual animals, within a given species, that show especially severe or especially early change in one form of senescence, ought also to exhibit severe or premature change in many other age-dependent systems. We have tested this hypothesis using the method of partial least squares (PLS) which has been exploited extensively in psychology and morphometrics (Bookstein, 1991; Streissguth et al., 1993), but has not to date been applied to gerontological questions, including the problem of biomarker characterization and validation.

Methods

Mice. — This study used genetically heterogeneous mice, bred as the progeny of CB6F1 females and C3D2F1 males. The F1 hybrid breeding stock was purchased from The Jackson Laboratories (Bar Harbor, ME). A "four-way cross" breeding system generated a population of mice in which no two individual animals were genetically identical, but whose characteristics — as a population — can be reproduced at any sample size at any time in any laboratory with access to the F1 hybrid parental or inbred grandparental strains. The mice were housed one per cage in a specific pathogen-free colony using microisolator cages to minimize the risk of cage-to-cage spread of infectious agents. Specific pathogen-free status was verified every 3 months using a procedure in which sentinel mice were first exposed to spent bedding from the study mice and then tested serologically for antibodies to Sendai virus, Mycoplasma pulmonis, and murine coronavirus (mouse hepatitis virus). Anal tape tests for pinworm were also performed on the sentinel mice. In addition, once each year sera from sentinel mice were tested by a commercial laboratory for a more extensive panel of viral antigens, including lymphocytic choriomeningitis, ectromelia, minute virus of mice (MVM), reovirus-3, and six other less common pathogenic viruses. All such tests were negative throughout the course of this study.

In the ad libitum-fed (AL) group, mice were given free access to food and water. Mice in the caloric-restricted (CR) group were given 60% of the food consumed by a sample of 5–7 gender-matched mice in the AL group. The diets used were PMI Autoclavable Lab Diets purchased from Bionetics, Inc. (Jefferson, AR). The CR diet was formulated so as to provide adequate levels of micronutrients to the CR mice despite their lower caloric intake. This calorie-restriction regime was imposed stepwise, with restriction to the level of 90% imposed at 2 months; 75% at 3 months; and 60% at 4 months. Food consumption in the AL group was monitored until the mice were 12 months old, but the monitoring was then terminated because no change in food consumption was observed between the 3rd and 12th month.

To minimize potential effects of seasonal variation, mice were entered into the study in staggered cohorts, with 25% of the population entered in January, April, July, and October of 1994. Each quarterly cohort contained an equal number of males and females, and an equal number of AL and CR mice. Blood was drawn from each mouse at approximately 8 months of age (range, 7.5–9.1 months) for immune phenotyping. Each mouse was then tested for muscle function at approximately 12 months of age (range, 11–12.5 months). Immune status tests and muscle function tests were staggered in this fashion to minimize the effects of each test battery on the outcome of the other test set. Of the 128 mice entered into the study, only one animal died prior to 12 months of age (i.e., prior to the completion of the first round of muscle function tests).

Immune function testing. — Each mouse was bled from the tail vein into a tube containing 15 μl of heparin (225 μg/ml). The blood sample (typically 400 μl) was then diluted by addition of 1.5 volumes of sterile phosphate-buffered saline (PBS) and then centrifuged at 4 °C for 5 min at 3000 rpm. Sixty percent of the supernatant was then removed and stored for later biochemical analysis. The remaining blood sample was then mixed with 1.5 volumes of Hanks' balanced salt solution (HBSS) (Sigma) containing 0.2% bovine serum albumin (BSA) (Sigma). The samples were then divided into eight equal aliquots, each containing cells equivalent to approximately 40 μl of blood. Six of these aliquots were used in this study. Aliquot 1 was stained with fluorescein isothiocyanate (FITC)-conjugated anti-CD3 antibody and mixed with a known number of phycoerythrin (PE)-conjugated latex beads, to allow an estimate of the number of CD3-positive T cells per unit of blood volume. Aliquot 2 was stained with FITC-anti-CD3 and PE-anti-CD4 to measure the percentage of CD3 cells that were also CD4-positive. Aliquot 3 was treated as for aliquot 2, except for the use of PE-anti-CD8 in place of PE-anti-CD4. Aliquot 4 was treated with FITC-anti-CD44 and PE-anti-CD4 to estimate the proportion of CD4 cells that express the memory cell marker CD44. Aliquot 5 was treated in the same way as aliquot 4, except that FITC-anti-CD8 was used in place of FITC-anti-CD4. Aliquot 6 was treated with FITC-anti-CD45RB and PE-anti-CD4 to estimate the proportion of CD4 cells that expressed the naïve cell marker CD45RB.

All FITC and PE antibodies were purchased from PharMingen (San Diego, CA), and were diluted 1:100 in cold HBSS/BSA. PE-conjugated beads (Flow Cytometry Standards, San Juan, PR) were diluted to 4 × 10^5 beads/ml in cold HBSS/BSA. Each sample contained 100 μl of the cell suspension and 50 μl of the appropriate FITC antibody, followed immediately by 50 μl of PE-conjugated beads or the appropriate PE antibody. Samples were incubated at 4 °C for 30 min with gentle shaking every 5–10 min. Samples were then centrifuged at 3000 rpm, washed once with HBSS/BSA, washed once with HBSS alone, and resuspended in PBS. To remove red blood cells, samples were transferred to 2 ml of FACS lysing solution (Becton Dickinson, San Jose, CA) and incubated for 5–10 min at room temperature. Samples were then centrifuged at 3000 rpm, resuspended in 200 μl of 0.1% paraformaldehyde, and kept in the dark at 4 °C until analysis. Stained samples were analyzed using a Becton Dickinson FACScan flow cytometer, using forward and side scatter gates to restrict analysis to lymphocytes. Ten thousand gated cells were analyzed for each sample. An IBM-compatible PC with PC-LYSYS software was used for data analysis.

Muscle function testing. — The mice were anesthetized with a 1/80 dilution of Avertin (99% 2,2,2-tribromoethanol in *tert*-amyl alcohol). Supplemental doses were administered as required to maintain a level of anesthesia that prevented response to peripheral tactile stimuli. When anesthetized, the mice were placed on a plexiglass platform of a "shoe" apparatus (Ashton-Miller et al., 1992). The left knee was secured to the platform and the foot was attached to the shoe with a Velcro strap. The shoe was mounted on the shaft of a servomotor. A microcomputer controlled the servomotor for rotations of the ankle joint that permitted adjustments in muscle length. Strain gauges mounted on the shaft of the servomotor monitored the torque produced by

the dorsiflexor muscle group. Needle electrodes were placed adjacent to the peroneal nerve, and the dorsiflexor muscles were stimulated maximally with the dorsiflexor muscle group adjusted to optimal length for force development (L_o). The maximum isometric tetanic force (F_0) was then determined from the plateau of the frequency-force curve. The sustained force (F_s) and the fatigue index (FI) were determined by stimulating the dorsiflexor muscle group with repeated isometric tetanic contractions for a period of 2 min (Burke et al., 1973). With the muscle group placed at an average L_o for the total muscle group, the peroneal nerve was stimulated at 80 Hz for 200 ms every 660 ms which produced a total of 182 contractions. The force of the initial contraction (F_i) and that of the final contraction (F_f) of the repeated isometric contraction protocol were measured. F_s was calculated as F_i multiplied by the duty cycle (Brooks and Faulkner, 1991). The fatigue index was calculated as $(F_f/F_i) \times 100$. Following the completion of the muscle function tests, the anesthetized mice were removed from the apparatus, allowed to recover, and then returned to their cages.

F_0 and F_s are a function of the total fiber cross-sectional area (Brooks and Faulkner, 1991); but data on cross-sectional area were not available. Consequently, F_0 and the F_s were normalized by the body mass (M_b) and designated as F_0/M_b and F_s/M_b, respectively. The fatigue index is already normalized by the division of F_f by F_i.

Statistical analysis: overview. — PLS constitutes a family of methods for analyzing sets, or blocks, of multivariate data. In the current case there are two such blocks, one consisting of immune status variables obtained from 9-month-old mice, and the other consisting of muscle function tests performed on the same set of mice at 12 months of age. The analysis begins with a matrix of ordinary correlation coefficients, each of which summarizes the correlation between one of the immune measures and one of the muscle measures. The goal of the analysis is to generate two sets of weighting factors — one for muscle tests and one for immune tests — that summarize the muscle and immune data in a useful way. Specifically, the weights are chosen in such a way that the weighted sum of the immune outcomes, for a given mouse, has maximum covariance with the weighted sum of the muscle outcomes. The weighted sum of these values is referred to as a *latent variable*: in the current case one latent variable is a composite summary of immune status, and another latent variable is a composite summary of muscle function. The individual weighting factors are referred to as *saliences*. Specific immune tests that have high salience receive highest weights — positive or negative — in the computation of the immune latent variable. The set of weights can then be used to calculate, for each individual mouse, a composite immune latent variable and a composite muscle latent variable. The covariance relationship between the immune latent variable and the muscle latent variable can be tested for statistical significance to see if the pattern is likely to represent chance alone. The pattern of salience factors can also be interpreted to see if they support, refute, or suggest hypotheses of biological interest.

The two sets of weighting factors just described are actually the first in a series of such pairs that can extract potentially meaningful latent variables from the cross-block correlation matrix. After the first pair is computed as described above, one can compute a second pair that explains as much as possible of the remaining variation, improving the least-squares fit to the matrix of correlation coefficients as much as possible over the fit generated by the first pair of latent variables. In some cases, including the within-diet analyses here, the second pair of latent variables accounts for very little additional correlation, and can thus safely be ignored.

Statistical analysis: methodology. — The analysis begins with two lists of variables, X_1 through X_m and Y_1 through Y_n. In the present case the X's are the six immune status tests, and the Y's are the five muscle function assays plus M_b. R is the correlation matrix of the X's by the Y's, with m rows and n columns, such that R_{ij} is the correlation between X_i and Y_j. The calculation then produces a vector A containing m coefficients A_i, one for each X, and a vector B containing n coefficients B_j, one for each Y. The elements of each vector, A or B, are scaled so that their squares sum to 1. The elements are chosen to meet the following criterion. Consider the latent variable score

$$LV_X = \sum_{i=1}^{m} A_i X_i .$$

(In this calculation the X values are expressed as z-scores, i.e., they have been rescaled to a mean of zero and a variance of 1, before the arithmetic is carried out.) Similarly the latent variable for the Y values is

$$LV_Y = \sum_{j=1}^{n} B_j Y_j .$$

The elements of the vectors A and B are chosen such that the covariance (d) between LV_X and LV_Y is a maximum for any pair of such linear combinations for which the coefficient vectors A and B both sum in square to 1.

Note that the PLS method optimizes the covariance, not the correlation, between the scores. The more familiar canonical correlations analysis optimizes correlation instead. Our interest, however, is in what the variables of each block might have in common for predicting to the other block, and the canonical correlations approach is not designed to answer this sort of question. Indeed, PLS can be thought of (Wold, 1989) as a hybrid of regression and factor analysis that attempts to find "factors" common to all the regressions between the blocks.

The elements A_i of the vector A are proportional to the correlations of the corresponding X-block variable X_i with the latent variable LV_Y representing the B's, i.e., the other block in the analysis. Thus each A_i is the salience of the variable X_i for the latent variable representing the Y block. The analogous conditions apply to the B_j values in relation to LV_X.

In the technical language of multivariate statistics, we can characterize this analysis by the least-squares property of the singular-value decomposition (Mardia et al., 1979). The outer product AB^t, the $m \times n$ matrix whose ijth entry is A_iB_j, is, after multiplication by the covariance d, the best (least-squares) fit of all such matrices to the correlation matrix R between the blocks. The goodness of fit of this approxima-

tion serves as a figure of merit for the overall PLS analysis. It is characterized as "the fraction of summed squared correlation explained" in the model R = $dAB' + error$.

While the tactics of PLS are not yet a standard component of any statistical package, the analysis is easily executed in any package that includes a singular value decomposition routine. We used version 3.3 of the Splus statistical package (StatSoft, Inc., Tulsa, OK). In terms of the usual algebraic quantities returned by singular-value software, the vectors A and B are the first left and right singular vectors of the cross-block correlation matrix R, and the covariance d is the first singular value.

Differences in immune and muscle function indices between the AL and CR groups, and between males and females, were assessed using a two-tailed Student's t-test.

Results

Gender and diet effects on immune and muscle function tests. — Immune status tests (at 8 months of age) and muscle function tests (at 12 months of age) were performed on a group of 127 mice, of which 63 were CR and 64 fed an AL diet. Some tests were spoiled by technical failures, but technically acceptable results were obtained on between 105 and 123 mice for each of the six immune and five muscle function assays. Table 1 summarizes the test battery, showing mean values and standard errors for the entire group of mice and for the AL and CR groups separately. Since F_0 and F_s are expected to vary with M_b, Table 1 also includes calculations of F_0/M_b and F_s/M_b as well as the fatigue index. Table 1 also shows the probability, assessed by a two-tailed Student's t-test, that the observed differences between AL and CR mice would arise by chance sampling from a single homogeneous pool of animals. (The CD3 scores are very long-tailed. Assessed by a nonparametric Mann–Whitney test, the difference between the diet groups in CD3 approaches statistical significance. None of the other assessments are changed in this way by abandoning the assumptions of the t-test.)

Compared with AL controls, CR mice had higher levels of CD4 T cells and lower levels of CD8 cells in their peripheral blood. AL mice were, as expected, heavier than CR mice at 12 months of age. The AL mice generated higher levels of F_0, but the CR mice had higher F_0/M_b. There was no intergroup difference in F_s, but F_s/M_b was almost 70% higher in CR than in AL mice, and fatigue index was also significantly greater in CR than in AL mice.

We also examined the data for differences between males and females. Males tended to be heavier and to generate significantly greater F_0 than females; they also, for reasons that are obscure, tended to have higher numbers of virgin CD4 T cells (CD4v) (data not shown). None of the other immune or muscle function tests showed a significant difference between males and females.

Correlations among age-sensitive indices of immune and muscle function. — We calculated the correlation coefficients for each immune variable against each muscle variable, treating body mass as a member of the muscle block for this analysis. Separate cross-block correlation matrices were computed for the AL and CR mice; they are shown in Table 2. Each correlation coefficient is based upon a sample of 49–61 individual mice. As is standard for PLS analysis, the correlation coefficients are based on a slightly different sample for each pair of values, because each set of data contains a slightly different set of missing values. For samples of about 50, correlation coefficients of ±.36 are significant at the level of .01. We do not highlight these significance levels because the analysis deals with the pattern of the entire matrix and not merely the significant subset.

We then used the PLS method to calculate latent variables relating muscle function to immune status markers. The analysis was carried out separately for AL mice and for CR mice. The results are shown in Table 3, which displays the salience of each of the immune and muscle tests, proportional to the correlation of the test with the salience-weighted average of the values in the other test block.

The significance of a PLS analysis is evaluated using a permutation test (Bookstein et al., 1990; Good, 1994) to compare the observed first singular value to values that are generated when the association between immune and muscle measures is annihilated by random permutation. These tests

Table 1. Mean Scores on Tests of Immune Status and Muscle Function in AL and CR Mice

	Ad libitum-Fed (AL)		Calorie-Restricted (CR)		Both		
Test	Mean ± *SD*	*n*	Mean ± *SD*	*n*	Mean ± *SD*	*n*	*p*-value*
CD3 (cells/μl)	1138 ± 730	52	970 ± 884	53	1053 ± 812	105	.293
CD4 (% of CD3)	66 ± 7	56	70 ± 7	58	68 ± 7	114	.001
CD8 (% of CD3)	30 ± 7	60	26 ± 5	61	28 ± 6	121	.001
CD4m (% of CD4)	25 ± 11	60	26 ± 12	63	25 ± 12	123	.602
CD8m (% of CD8)	27 ± 11	63	26 ± 15	60	27 ± 13	123	.644
CD4v (% of CD4)	69 ± 19	57	69 ± 14	55	69 ± 17	112	.933
F_0 (Newtons)	3.27 ± 0.64	62	2.66 ± 0.61	61	2.97 ± 0.69	123	.001
F_s (Newtons)	0.37 ± 0.11	61	0.36 ± 0.10	61	0.37 ± 0.11	122	.582
Fatigue index (FI) (%)	45 ± 9	61	54 ± 10	61	49 ± 10	122	.001
M_b (grams)	38 ± 9	62	21 ± 2	61	30 ± 10	123	.001
F_0/M_b (Newtons/kg)	91 ± 24	62	125 ± 26	61	107 ± 30	123	.001
F_s/M_b (Newtons/kg)	10.2 ± 3.9	61	17.0 ± 4.9	61	13.6 ± 5.6	122	.001

*Significance test for a two-tailed Student's t-test comparing AL to CR mice.

Table 2. Cross-Block Correlation Matrices for AL and CR Mice

Assay	F_0	F_s	Fatigue Index (FI)	M_b	F_0/M_b	F_s/M_b
AL mice						
CD3	0.19	0.33	0.13	−0.15	0.29	0.38
CD4	0.15	−0.07	−0.09	0.01	0.16	−0.03
CD8	0.16	0.30	0.13	−0.04	0.12	0.26
CD4m	−0.32	−0.22	0.07	−0.07	−0.24	−0.19
CD8m	−0.07	0.02	0.12	0.00	−0.10	−0.01
CD4v	0.02	−0.03	−0.01	−0.32	0.24	0.14
CR mice						
CD3	0.30	0.19	−0.05	0.24	0.24	0.12
CD4	0.05	−0.08	−0.29	0.09	0.04	−0.09
CD8	−0.08	0.03	0.14	−0.28	0.03	0.12
CD4m	−0.60	−0.42	0.23	−0.21	−0.61	−0.37
CD8m	−0.21	−0.14	0.16	0.18	−0.34	−0.22
CD4v	0.11	0.17	0.18	−0.28	0.24	0.26

Notes: n = 49–61 mice for each AL correlation, and n = 52–61 for each CR correlation.

Table 3. Saliences for AL and CR Mice

Immune Assay	Salience	Muscle Assay	Salience
AL mice only			
CD3	−0.68	F_0	−0.42
CD4	−0.08	F_s	−0.50
CD8	−0.45	*FI*	−0.10
CD4m	0.49	*M*	0.17
CD8m	0.06	F_0/M	−0.49
CD4v	−0.26	F_S/M	−0.54
CR mice only			
CD3	−0.35	F_0	−0.54
CD4	−0.03	F_S	−0.39
CD8	0.02	*FI*	0.18
CD4m	0.82	*M*	−0.11
CD8m	0.36	F_0/M	−0.60
CD4v	−0.23	F_S/M	−0.38

are not affected by any aspect of correlation structure within the blocks separately. For the AL mice, the observed first singular value is .926. After 500 permutations, the 24th largest of the first singular values was .921. The observed relationship for the AL mice is thus just significant at the $p \leq .05$ criterion. For the CR mice, the observed first singular value is 1.29. In 500 permutations, the third largest of the first singular values was 1.28, and the relationship is thus significant for the CR mice at $p < .005$. The second singular values for AL and CR mice were, respectively, .40 and .65. Each explains less than half as much covariance as the first singular vectors. We will not consider them further until we have a longer series of waves of this longitudinal data set.

That the first singular value is significant implies that there is some interpretable pattern in the matrices of correlation coefficients shown in Table 2. The role of the salience coefficients shown in Table 3 is to guide the interpretation. For the AL mice, the three largest components of the immune latent variable are, in size order, (−)CD3, CD4m, and (−)CD8. This immune variable is apparently a contrast between CD4m and the combination of CD3 and CD8, with the other immune tests having lower salience. The muscle latent variable is not a contrast but rather a composite of the two force measurements, F_0 and F_s, together with their mass-adjusted versions F_0/M and F_s/M. This pair of singular vectors corresponds to a fairly clear pattern within the correlation matrix itself (Table 2, top): correlations between CD4m and both F_0 and F_s are negative, while correlations between CD3 and CD8 and both F_0 and F_s are positive.

The salience pattern for the CR mice (Table 3, bottom) is similar in several respects to that seen for the AL mice. The immune latent variable is largely a contrast between CD4m and CD8m against (−)CD3. Thus for the CR mice the salience of CD8 values is much lower than for AL mice, and the salience of CD8m much more prominent. The muscle latent variable for CR mice is very similar to that calculated for AL mice, with strong parallel contributions from F_0, F_s, F_0/M_b, and F_s/M_b, but comparatively little effect of fatigue index or M_b. The pattern in the correlation matrix (Table 2, bottom) is correspondingly straightforward, showing negative correlations between CD4m and CD8m against F_0, F_s, and their mass-adjusted variants, along with positive correlations between CD3 and each of the measures of muscle force.

Figure 1 shows scatterplots relating the first immune latent variable (vertical axis) to the first muscle latent variable for each mouse in the study, with AL and CR mice plotted in separate panels. In each case there is a clear relationship between the immune and muscle latent variables. The covariances are .78 and 1.14, respectively, for the AL and CR groups. These covariance values are somewhat less than the singular values because missing values have been substituted by means in the computation of scores for each mouse. The correlation coefficients are $r = .45$ for the AL mice, and $r = .51$ for the CR mice. The correlation coefficient for the AL mice is only slightly above the highest of the individual AL correlation coefficient, and the CR correlation is slightly worse than the highest of the individual CR values. This should not be thought of as "little gain" (as in multiple regression approaches), since the subject of discussion is not optimization of the correlation but instead the discovery of patterns in the correlation structure as a whole. The scatterplot for the AL mice (Figure 1, Left) shows that the relation between the two latent variables is well summarized by the correlation coefficient. For the CR mice (Figure 1, Right), there is one joint outlier, and the correlation coefficient of .51 is thus an overestimate; the correlation coefficient for the data with the outlier omitted (.40) provides a better summary of the CR data set as a whole.

DISCUSSION

Our working model involves three linked suppositions. First, we hypothesize that there is a physical process, of unknown nature, that influences the rate of change in a wide range of traits that change in adults as they grow older. We further propose that the rate of this process differs among members of a species. Lastly, we postulate that the effects of these differences in rate of aging are strong enough that they can be detected despite the confounding effects of other

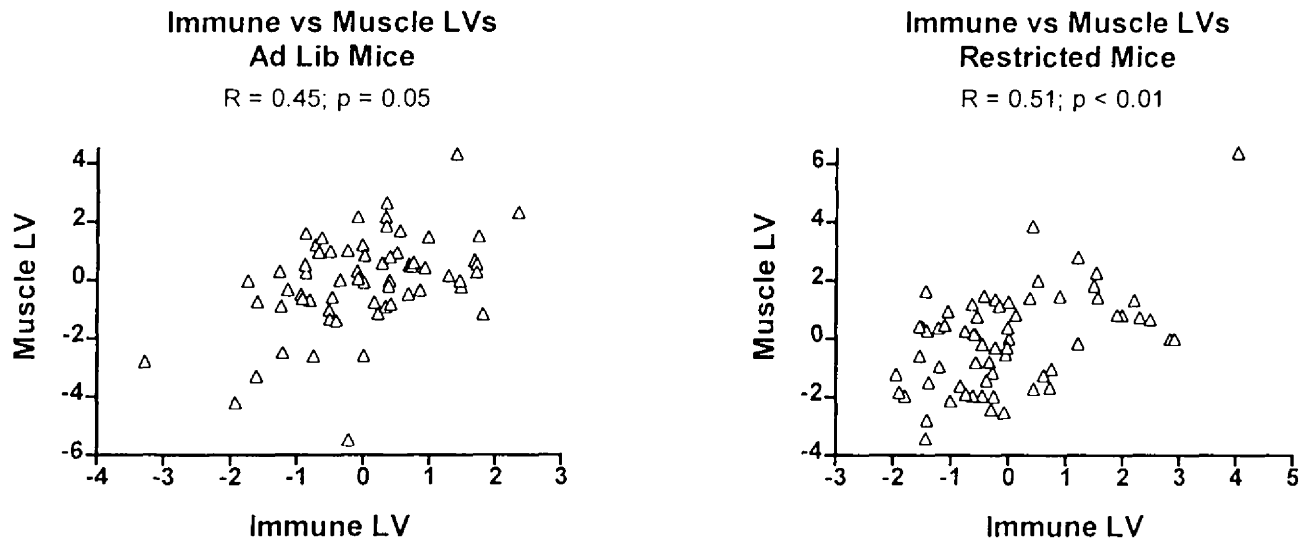

Figure 1. Scatterplot of immune latent variable (LV) scores versus muscle LV scores for ad lib-fed mice (left) and for caloric-restricted mice (right). Each point represents one animal.

causes of individual variation. This model suggests a set of criteria that could be used to evaluate the potential usefulness of age-sensitive traits as candidate biomarkers of the rate of aging. (*1*) The candidate biomarker, or combinations of biomarkers, ought to exhibit measurable correlations with other biomarkers, even those that reflect changes in organ systems thought to be physiologically independent of one another, such as the immune system and the skeletal musculature. (*2*) It should provide information about the underlying rate of aging that varies within a species and thus provide an ordering of individual members of a cohort that is stable across the years or months of adult life. (*3*) The trait or combination of traits ought to have prognostic value, and differ among individuals destined to exhibit varying risks of late-life disease and thus varying life spans. This criterion requires that candidate biomarkers be innocuous, with no detectable effects on life span, disease incidence, or even tests carried out later in the life course. (*4*) Ideally, the trait should also discriminate groups subjected to treatments known to affect life span (e.g., caloric restriction), although effects of the intervention that affect the assays in ways unrelated to aging rate may make this criterion difficult to meet. Caloric restriction will, for example, lead to alterations in muscle size and body mass that strongly influence the outcome of some muscle function tests independently of the CR effect on aging.

Our experimental design will eventually provide data relevant to each of these test criteria, but the current analysis is relevant only to the first issue. Our test battery included a number of outcomes, such as CD4m, CD8m, F_0, F_s, fatigue index, F_0/M_b, and F_S/M_b, that have been shown in cross-sectional studies to differ between young and old adults (Brooks and Faulkner, 1988, 1994; Lerner et al., 1989; Ernst et al., 1990; Faulkner et al., 1995). If indeed individual differences in aging rate influence both immune status and muscle function, we would predict negative correlations between CD4m and CD8m against all indices of strength and fatigue. These expectations are largely supported by the data on the CR mice, except for fatigue index. The AL mice show the predicted correlations for the CD4m data, but not for the CD8m surface markers. CD4v values are expected to vary inversely with CD4m counts, but are more difficult to measure accurately; this technical constraint may have led to the relatively low correlations for CD4v against the strength indices and the relatively low salience of CD4v in the PLS calculations.

Cross-sectional data on age-associated changes in other immune markers, CD3, CD4, and CD8, have been inconsistent. There have been some reports of age-dependent change in CD4 and CD8 cell numbers or proportions, but the magnitude and even direction of the change differ from study to study. In our AL group, low levels of CD8 cells are associated with low levels of muscle force and sustained force. Indeed, studies of mouse spleen (Komuro et al., 1990) and human peripheral blood (Moody et al., 1981; Walker et al., 1990; Utsuyama et al., 1992) have found significant age-associated declines in the relative proportion of CD8 T cells compared to CD4 cells. In contrast, however, four other groups have observed significant changes in the opposite direction in the CD4/CD8 ratio (Boersma et al., 1985; Grossmann et al., 1991; Kirschmann and Murasko, 1992; Callahan et al., 1993), while 10 groups have reported an absence of any systematic shift in CD4 and CD8 cells in aging mice and humans (see Miller, 1995, for citations).

The question of age-dependent changes in T-cell number has also been a subject of experimental controversy. Significant increases in splenic T-cell numbers have been reported to occur in CBA/H and AB6F_1 mice (Stutman, 1972; Patel, 1981). In contrast, age-dependent decreases in mouse splenic T-cell numbers have also been reported in DBA/2 and CBA/H-T6 mice (Kay et al., 1979). All the published work on age-related changes in mouse T-cell numbers has been based on studies of splenic T cells of genetically homogeneous strains and may therefore have only limited

relevance to peripheral blood analyses in heterogeneous mice. A decline in peripheral blood lymphocyte count has been shown to predict 1-year mortality in apparently healthy humans (Bender et al., 1986), but the available data did not discriminate between blood T cells and other blood lymphocyte populations. Our data show a positive association between CD3 T-cell number and several measures of muscle force in both AL and CR mice. It will be of interest to see whether this association persists to later ages and whether mice with high scores on these tests are relatively resistant to other varieties of age-associated decline.

For all mammals, and particularly human beings, skeletal muscles show decreases in both mass and force development between adulthood and old age (Brooks and Faulkner, 1994; Faulkner et al., 1995). For human beings, muscle mass decreases by 36–40% between the ages of 30 and 80 years, and force development decreases at least an equivalent amount (Grimby and Saltin, 1983). More precise studies of the skeletal muscles of young, adult, and old mice (Brooks and Faulkner, 1988) and rats (Carlson and Faulkner, 1988) have demonstrated that, with aging, strength decreases even more rapidly than muscle mass. Consequently, with no change in fiber length, the specific F_0 (kN/m^2) of the muscle decreases 20–25%. The atrophy and weakness observed in skeletal muscles of old animals is similar to the atrophy and weakness associated with physical inactivity in animals of any age (Brooks and Faulkner, 1994; Faulkner et al., 1995). Despite the similarities, the atrophy and weakness associated with old age constitute an intrinsic, irreversible decline that occurs even in the presence of continued high levels of daily physical activity. For masters and professional sportsmen and sportswomen, the declines in performance are already evident in the thirties (Faulkner et al., 1995). The muscles of old people can be trained, but beyond the thirties training will not return the muscle mass, strength, and power to the values that could have been attained by training of muscles by the same individuals in their twenties.

In addition to atrophy and weakness, muscles of old animals are also more easily fatigued than those in young and adult animals (Faulkner and Brooks, 1995). Brooks and Faulkner demonstrated with a test of sustained power that muscles in old (24-month) mice had a sustained power only 55% of the value for young (4-month) mice. Even adult (12-month) mice showed an impairment with values for sustained power 75% of the value for the young mice. With aging, the greater deficit for fatigability than for strength or power appears to be attributable to an inability to maintain a high energy balance for either force or power (Dudley and Fleck, 1984). The failure to maintain a high energy balance could result from a decrease in the capacities of the oxidative enzyme systems, concentration of mitochondria, or the ability of the circulation to deliver oxygen (Brooks and Faulkner, 1991).

The PLS approach provides a useful way to assess the significance and implications of relationships within sets of potentially linked outcome measures. The permutation tests show that the relationship between the immune and muscle tests is unlikely to have been observed by chance alone, for either the AL or CR mice. The calculated saliences (Table 3) suggest that, for both AL and CR mice, the relationship between the immune and muscle outcomes is based almost equally on the F_0 and F_s measures, with or without adjustment for body mass. For the AL mice, CD3, CD4m, and CD8 assays contribute the most to the relationship with muscle strength, while for the CR mice CD4m, CD8m, and CD3 are most influential. Thus the PLS analysis provides support for the proposition that individual differences in immune status may be related to differences in muscle strength and endurance via pathways that are consistent with gerontological predictions. The PLS data also suggest that CD3 and CD8 cell counts deserve to be examined as potential biomarkers of aging.

Suggestions that individual differences in the rate of aging could be monitored by assays of age-sensitive traits are a long-standing tradition in biomedical gerontology (Ludwig and Smoke, 1980; Reff and Schneider, 1982; Harrison and Archer, 1983; Baker and Sprott, 1988). Powerful objections have been raised to this idea. The objections typically take two forms: *(1)* complaints that the statistical arguments used in biomarker research are prone to artifact, and *(2)* assertions that age-sensitive traits are in fact not timed by a single process, let alone one that is measurable by assays of the traits themselves (Adelman, 1987; Costa and McCrae, 1988, 1995; Partridge and Harvey, 1993).

On the question of statistical methodology, we note that attempts to combine measures of age-sensitive traits into more powerful predictors of longevity, i.e., as composite biomarkers, have typically relied on variants of multiple regression or canonical correlation analysis. These analyses have been criticized on technical grounds (Costa and McCrae, 1988, 1995), in our view justifiably. We prefer the PLS approach to any of the assortment of alternative multivariate analyses that have been suggested for similar questions. The statistical technique of canonical correlations analysis, for instance, optimizes the correlation between the two latent variables, rather than their covariance; but it is no longer the case that the coefficients of the linear combination represent the importance of the contributions of individual variables to the cross-block prediction. In a canonical correlations analysis of highly correlated blocks like these, variables measuring "the same thing" will be assigned coefficients that vary wildly in value and perhaps even in sign; we can learn little about actual explanatory power from such coefficients. Furthermore, the analysis depends on the structure of the within-block correlation matrices, which is not the case for PLS. Nor would it be appropriate to pool all 12 of these variables in one single factor analysis and then interpret the patterns of loadings on the components that emerge. Those estimates would likewise depend on the within-block correlation structure rather than attending primarily to the structure of cross-block prediction that is our primary concern.

Future reports from this study will incorporate additional rounds of immune measurement (at ages 15 and 21 months) and muscle measurement (18 and 24 months). The only competing technique that can likewise handle arbitrary numbers of blocks in this way is the technique of structural equations modeling (SEM), a composite of regression and factor analysis techniques popular in several subfields of the psychological sciences. The computations of PLS are not far

from those of structural-equations analysis of the same data sets when the SEM models are constrained to ignore within-block factor structure and to estimate by least squares rather than maximum likelihood. The PLS approach is preferable to SEM mainly because PLS supplies the explicit latent variable scores that enter into scatterplots such as Figure 1 and because PLS allows us to test the dependence of these scores on age at death. The PLS analysis does not attempt to predict longevity by any measured variable by either simple or multiple regression. It will, however, help to determine if what the immune and muscle test batteries measure *in common* is associated with longevity. PLS is the only method we know of that allows us to approach this gerontological challenge in its most straightforward form.

The second objection, arguing that aging is not a single process with a single rate subject to individual variation, seems to us a serious one, but one that is amenable to experimental test. In support of the single process idea, we note demonstrations that single gene differences in invertebrates (Johnson, 1990; Van Voorhies, 1992; Kenyon et al., 1993) can lead to intra-species differences in age-sensitive traits, including mortality risk. We note also that an environmental effect, calorie restriction, can retard a panoply of age-sensitive phenotypes (Weindruch and Walford, 1988). Nonetheless, the single process model is deservedly controversial, and testing it is a central focus of the current study. The model predicts, among other things, that individuals whose aging rate is high compared to others in the population will have relatively extreme phenotypes in many age-sensitive traits, including tests of systems for which other sources of linked change seem unlikely. Thus, for example, it seems unlikely, based on present knowledge of pathophysiology, that muscle weakness will lead to alterations in CD4 memory T-cell numbers, and equally unlikely that alterations in T-cell subsets will lead to diminutions in muscle strength or endurance. Thus, evidence to suggest correlated variation in age-sensitive measures of muscle and immune status, such as we present in Table 3 and Figure 1, is most parsimoniously explained by a model in which variations in both muscle and immune status are regulated by a third factor, hypothetically the aging rate. We can also predict that other age-sensitive traits, such as collagen cross-linking, protein oxidation levels, and measures of function in other physiologic systems, will prove to covary with the immune and muscle assays that are salient in the current data set. To address this idea, we are providing samples of tissue from these mice to several collaborators with interest and expertise in these test systems.

One long-term goal of this study is to see if combinations of age-sensitive indices can predict life span and risk of disease incidence. The latent variables derived in the current study relate immune status to muscle function without regard to longevity, but analogous methods can be used to calculate latent variables that exhibit maximal covariance with individual life span. The input variables for this calculation may include both traits like those used in the current work and also rates of change in these traits across the life span. Our model of the aging process predicts that combinations of variables that show strong cross-block covariance (as in Figure 1 and Table 3) will also contribute strongly to latent variables that estimate life span; this prediction will be tested as members of this cohort grow older and begin to die.

The use of life span as the central test for intergroup differences in aging rate has a long tradition in experimental aging research, but in our view only for want of a better metric. Life span is influenced by many factors other than biological aging. These factors include genetic predisposition for specific diseases, blatant or subtle differences in environmental variables, and interactions between genetic and nongenetic influences. In statistical terms, the life span measurements contain only a very few bits of information (in the limit, one bit: alive or dead), and are thus in principle less well suited to quantitative analysis than indices that integrate a wide range of age-sensitive measurements. Biomarkers, if they are to be more than merely risk factors, must be shown to provide useful prognoses not merely for life span and disease incidence, but also for a broad spectrum of physiological and biochemical changes characteristic of the aging process itself. The current study does not yet meet this standard. Additional waves of measurement will allow us to assess the degree to which our candidate biomarkers provide stable indices of relative immune and muscle status over the life span. Additional channels of measurement — e.g., tests of collagen crosslinking, protein oxidation, retrovirus recombination, serum interleukin 6 levels, etc. — will allow us to see if immune or muscle latent variables covary with other indices of age-dependent change. Life span and necropsy data will let us test the hypothesis that combinations of immune and muscle tests predict remaining longevity and/or susceptibility to specific late-life illnesses. This ongoing longitudinal program may thus provide a test of whether the PLS method, applied to a system of partially correlated immune and muscle outcomes, can reveal latent variables related to interindividual differences in immune, muscle, and potentially whole animal aging rates.

ACKNOWLEDGMENTS

This research was supported by National Institute on Aging Grants AG-11067, AG-08808, and AG-11037. We acknowledge as well the expert assistance in animal husbandry provided by Ann Luciano and Sherry Swalwell, and the technical assistance of Jennifer McGrath. We thank Dr. Kevin Flurkey for information on age-associated changes in T-cell numbers in mice.

Address correspondence to Dr. Richard A. Miller, The Geriatrics Center/University of Michigan, 1150 West Medical Center Drive, Room 6240, Ann Arbor, MI 48109-0642. E-mail: millerr@umich.edu

REFERENCES

Adelman, R.C. Editorial: Biomarkers of aging. Exp. Gerontol. 22:227–229; 1987.

Ashton-Miller J.A.; He, Y.; Kadherisan, V.A.; McCubbrey, D.A.; Faulkner, J.A. An apparatus to measure in vivo biochemical behavior of dorsi- and plantarflexors of mouse ankle. J. Appl. Physiol. 72:1205–1211; 1992.

Baker, G.T.; Sprott, R.L. Biomarkers of aging. Exp. Gerontol. 23:223–239; 1988.

Bender, B.S.; Nagel, J.E.; Adler, W.H.; Andres, R. Absolute peripheral blood lymphocyte count and subsequent mortality of elderly men: the Baltimore Longitudinal Study of Aging. J. Am. Geriatr. Soc. 34:649–654; 1986.

Boersma, W.J.A.; Steinmeier, F.A.; Haaijman, J.J. Age-related changes in the relative numbers of Thy-1 and Lyt-2 bearing peripheral blood

lymphocytes in mice: a longitudinal approach. Cell. Immunol. 93:417–430; 1985.

Bookstein, F.L. Morphometric tools for landmark data. New York: Cambridge University Press, 1991.

Bookstein, F.L.; Sampson, P.; Streissguth, A.; Barr, H. Measuring "dose" and "response" with multivariate data using partial least squares techniques. Commun. Stat. Theory Methods 19:765–804; 1990.

Brooks, S.V.; Faulkner, J.A. Contractile properties of skeletal muscles from young, adult and aged mice. J. Physiol. (London) 404:71–82; 1988.

Brooks, S.V.; Faulkner, J.A. Maximum and sustained power of extensor digitorum longus muscles from young, adult, and old mice. J. Gerontol. Biol. Sci. 46:B28–B33; 1991.

Brooks, S.V.; Faulkner, J.A. Skeltetal muscle weakness in old age: underlying mechanisms. Med. Sci. Sports Exercise 26:432–439; 1994.

Burke, R.E., Levine, D.N.; Tsairis, P.; Zajac, F.E. Physiological types and histochemical profiles in motor units of cat gastrocnemius. J. Physiol. (London) 234:723–748; 1973.

Callahan, J.E.; Kappler, J.W.; Marrack, P. Unexpected expansions of CD8-bearing cells in old mice. J. Immunol. 151:6657–6669; 1993.

Carlson, B.M.; Faulkner, J.A. Reinnervation of long-term denervated rat muscle freely grafted into an innervated limb. Exp. Neurol. 102:50–56; 1988.

Costa, P.T.; McCrae, R.R. Measures and markers of biological aging: "a great clamoring . . . of fleeting significance." Arch. Gerontol. Geriatr. 7:211–214; 1988.

Costa, P.T.; McCrae, R.R. Design and analysis of aging studies. In: Masoro, E.J., ed. Handbook of physiology. Section 11: Aging. New York: Oxford University Press, 1995: 25–36.

Dudley, G.A.; Fleck, S.J. Metabolite changes in aged muscle during stimulation. J. Gerontol. 39:183–186; 1984.

Ernst, D.N.; Hobbs, M.V.; Torbett, B.E.; Glasebrook, A.L.; Rehse, M.A.; Bottomly, K.; Hayakawa, K.; Hardy, R.R.; Weigle, W.O. Differences in the expression profiles of CD45RB, Pgp-1, and 3G11 membrane antigens and in the patterns of lymphokine secretion by splenic CD4+ T cells from young and aged mice. J. Immunol. 145:1295–1302; 1990.

Faulkner, J.A.; Brooks, S.V. Muscle fatigue in old animals: unique aspects of fatigue in elderly humans. New York: Plemum Press, 1995.

Faulkner, J.A.; Brooks, S.V.; Zerba, E. Muscle atrophy and weakness with aging: contraction-induced injury as an underlying mechanism. J. Gerontol. (Special Issue) 50A:124–129; 1995.

Good, P. Permutation tests. New York: Springer-Verlag, 1994.

Grimby, G.; Saltin, B. The aging muscle. Clin. Physiol. 3:209–218; 1983.

Grossmann, A.; Maggio-Price, L.; Jinneman, J.C.; Rabinovitch, P.S. Influence of aging on intracellular free calcium and proliferation of mouse T-cell subsets from various lymphoid organs. Cell. Immunol. 135:118–131; 1991.

Harrison, D.E.; Archer, J.R. Physiological assays for biological age in mice: relationship of collagen, renal function, and longevity. Exp. Aging Res. 9:245–251; 1983.

Johnson, T.E. Increased life-span of age-1 mutants in *Caenorhabditis elegans* and lower Gompertz rate of aging. Science 249:908–912; 1990.

Kay, M.M.B.; Mendoza, J.; Diven, J.; Denton, T.; Union, N.; Lajiness, M. Age-related changes in the immune system of mice of eight medium and long-lived strains and hybrids. I. Organ, cellular and activity changes. Mech. Ageing Dev. 11:295–346; 1979.

Kenyon, C.; Chang, J.; Gensch, E.; Rudner, A.; Tabtiang, R. A *C. elegans* mutant that lives twice as long as wild type. Nature 366:461–464; 1993.

Kirschmann, D.A.; Murasko, D.M. Splenic and inguinal lymph node T cells of aged mice respond differently to polyclonal and antigen-specific stimuli. Cell. Immunol. 139:426–437; 1992.

Komuro, T.; Sano, K.; Asano, Y.; Tada, T. Analysis of age-related degeneracy of T-cell repertoire: localized functional failure in CD8+ T cells. Scand. J. Immunol. 32:545–553; 1990.

Lerner, A.; Yamada, T.; Miller, R.A. PGP-1hi T lymphocytes accumulate with age in mice and respond poorly to concanavalin A. Eur. J. Immunol. 19:977–982; 1989.

Ludwig, F.C.; Smoke, M.E. The measurement of biological age. Exp. Aging Res. 6:497–522; 1980.

Mardia, K.V.; Kent, J.T.; Bibby, J.M. Multivariate analysis. London: Academic Press, 1979.

Masoro, E.J. Aging: Current concepts. In: Masoro, E.J., ed. Handbook of physiology. Section 11: Aging. New York: Oxford University Press, 1995: 3–21.

Miller, R.A. Immune system. In: Masoro, E.J., ed. Handbook of physiology. Section 11: Aging. New York: Oxford University Press, 1995: 555–590.

Moody, C.E.; Innes, J.B.; Staiano-Coico, L.; Incefy, G.S.; Thaler, H.T.; Weksler, M.E. Lymphocyte transformation induced by autologous cells. XI. The effect of age on the autologous mixed lymphocyte reaction. Immunology 44:431–438; 1981.

Partridge, L.; Harvey, P.H. Methuselah among nematodes. Nature 366:404–405; 1993.

Patel, P.J. Aging and antimicrobial immunity: impaired production of mediator T cells as a basis for the decreased resistance of senescent mice to Listeriosis. J. Exp. Med. 154:821–831; 1981.

Reff, M.E.; Schneider, E.L. Biological markers of aging. Washington, DC: U.S. Department of Health and Human Services, 1982.

Streissguth, A.P.; Bookstein, F.; Sampson, P.; Barr, H. The enduring effects of prenatal alcohol exposure on child development. Ann Arbor: University of Michigan Press, 1993.

Stutman, O. Lymphocyte subpopulations in NZB mice: deficit of thymus-dependent lymphocytes. J. Immunol. 109:602–611; 1972.

Utsuyama, M.; Hirokawa, K.; Kurashima, C.; Fukayama, M.; Inamatsu, T.; Suzuki, K.; Hashimoto, W.; Sato, K. Differential age-change in the numbers of CD4+CD45RA+ and CD4+CD29+ T cell subsets in human peripheral blood. Mech. Ageing Dev. 63:57–68; 1992.

Van Voorhies, W.A. Production of sperm reduces nematode lifespan. Nature 360:456–458; 1992.

Walker, D.; Gauchat, J.F.; De Weck, A.L.; Stadler, B.M. Analysis of leukocyte markers in elderly individuals. Aging Immunol. Infect. Dis. 2:31–43; 1990.

Weindruch, R.; Walford, R.L. The retardation of aging and disease by dietary restriction. Springfield, IL: Charles C Thomas, 1988.

Wold, H.O.A. Introduction. In: Wold, H., ed. Theoretical empiricism: a general rationale for scientific model building. New York: Paragon House, 1989: vii–xi.

Received February 19, 1996
Accepted July 11, 1996