SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Human and Mouse Cohorts. Several cohorts of humans and mice were utilized in studies using genetic, cognitive, and other experimental approaches (Table S1).

Human Participants. For the discovery cohort, Cohort 1, stringent inclusion criteria were implemented for normal aging individuals; thus, replicate Cohorts 2 and 3 were analyzed with a parallel approach that restricted analyses to cognitively normal adults between 52–85 years of age with Mini-Mental State Exam scores of 28 or higher (see Table S2 for individual selection criteria for each population). All individuals studied, except 3 (2 Asian-Americans and 1 Latino), were Caucasian.

Human Neuropsychological Testing. Individuals were recruited for healthy aging studies at three centers. Studies were approved by committees on human research, and subjects provided written informed consent before participating. The tests used in Cohort 1 were suited for assessing cognitively normal individuals because they lack ceiling or floor effects and probe cognitive domains that are vulnerable in human aging (Drag and Bieliauskas, 2010; Gunstad et al., 2006). These included: phonemic fluency (number of D-words in 1 min; language and phonemic generation), category fluency (number of animals in 1 min; language and semantic generation), digit span backward (working memory), modified Trails B (completion time for set shifting task; attention and working memory), California Verbal Learning Test-II (CVLT-II; number of words correctly recalled at 20 min; verbal episodic memory), Benson Figure delay trial (score on free recall of drawing at 10–15 min; a modified version of the Rey-Osterrieth Figure test; spatial episodic memory), and Stroop inhibition (number correct in 60 sec; response inhibition). For reviews of tests, see (Kramer et al., 2003; Pa et al., 2010). For the replication Cohorts 2 and 3, tests that included cognitive domains assessed in the discovery cohort were evaluated (Table S4 for summary and references). Test responses were combined into a global composite score to evaluate cognitive function in a broad manner and to decrease the risk of a type I error by minimizing multiple comparisons. For each individual, test responses were standardized and then averaged resulting in a composite score. This score reflects the number of standard deviations above or below the global average (composite Z-score). Higher composite Z-scores

reflect better global cognition. Global composite Z-scores for some individuals reflect 5–6 test scores, instead of 7, due to random omissions during neuropsychological testing. Baseline neuropsychological evaluations were examined in 706 out of 718 individuals; inclusion of measures from a follow-up exam of 12 individuals did not contribute variance to composite Z-scores (ANOVA, p=0.59). Prior to calculating composite scores in the discovery cohort, we performed a principle component analysis (not shown) on the data. Our analysis revealed that each test contributes unique variation to the overall data. For meta-analysis, equal variance among cohorts was tested and confirmed. Then, global composite Z-scores from Cohorts 1–3 were combined and analyzed.

Human Genotyping. For Cohorts 1 and 3, genotyping for the KL-VS variant (rs9536314 for F352V and rs9527025 for C370S) was run as part of a 29-plex single nucleotide polymorphism (SNP) genotyping assay using Sequenom iPLEX Technology (Sequenom, San Diego, CA). The following PCR primer pairs were each used to amplify both F352V (rs9536314) and C370S (rs9527025) for sequencing: set 1 (forward, 5'-CACTCAGGGAGGTCAGGTGT-3'; reverse, 5'-CCTGAGACAAACCAGCCATT-3') and set 2 (forward, 5'-CACTGTGGGGTGACCTACCT-3'; reverse, 5'-CCTGAGACAAACCAGCCATT-3'). All PCR reactions and subsequent steps were performed as per the manufacturer's instructions. Spectro chip® arrays were spotted using a Nanodispenser, according to the manufacturer's instructions (Sequenom, San Diego, CA). A Sequenom MALDI-TOF spectrophotometer was used to read the array. The SpectroAquire and MassARRAY Typer Software packages (Sequenom, San Diego, CA) were used for interpretation and Typer analyzer V3.4.0.18 was used to review and analyze all data. For Cohort 2, DNA was genotyped on the Affymetrix Genechip 6.0 platform. Data underwent quality control (QC) analysis with PLINK software (Purcell et al., 2007). Standard QC measures were applied to these data (genotype success rate >95%, genotype-derived gender concordant with reported gender, excess inter/intra-heterozygosity) and SNPs (HWE $p > 0.001$; MAF > 0.05, genotype call rate > 0.95; misshap test > $1x10^{-9}$). EIGENSTRAT (Price et al., 2006) was then used to identify and remove population outliers based on default parameters. Imputation was performed using MACH software (version 1.0.16a) (Scott et al., 2007) and HapMap release 22 CEU (build 36) (Frazer et al., 2007). An imputation confidence score (INFO) of > 0.3 was used for QC.

Human Serum Measurements. Serum from morning fasting blood samples of individuals in Cohort 1 was collected and stored at –80°C. Soluble α-klotho, the major circulating form of klotho, was measured using a solid-phase sandwich enzyme-linked immunosorbent assay (Immuno-Biological Laboratories, Takasaki, Japan) (Yamazaki et al., 2010) according to the manufacturer's instructions. Serum was diluted 4 fold with the supplied Enzyme Immunoassay buffer. A standard curve was established by serial dilution of recombinant human soluble α-klotho protein. Standards and repeated samples were included on each plate to control for inter-plate variability. Diluted serum was loaded in duplicate onto a plate pre-coated with affinity-purified anti-human klotho (67G3) mouse IgG monoclonal antibody and incubated for 1 h at room temperature (RT). Plates were washed 7 times with washing buffer, followed by the addition of 100 μL HRP-conjugated anti-human klotho (91F1) mouse IgG monoclonal antibody for 30 min at RT. Plates were washed 9 times, and the reaction was visualized by the addition of 100 μL of chromogenic substrate (TMB) for 30 min at RT. The reaction was stopped with 100 μL of 1 N H_2SO_4, and the absorbance at 450 nm was measured on a Spectramax 190 plate reader (Molecular Devices, Sunnyvale, CA). α-Klotho levels were calculated with SoftMax Pro software.

Mice and drug treatment. All mice, except those used to determine longevity, were on an incipient congenic C57BL/6J background (N6–7: B6;C3H). Mice used for the longevity study were on a C3H/B6 background. Mice were kept on a 12-h light/dark cycle with *ad libitum* access to food (Picolab Rodent Diet 20, Labdiet) and water. The standard housing group was five mice per cage except for single housing during water maze studies. All cognitive and behavioral studies were carried out during the light cycle. Nontransgenic (NTG) C57BL/6 mice were crossed with hemizygous klotho (KL) transgenic mice (Line 46) (Kuro-o et al., 1997), which express mouse klotho ubiquitously under the EF-1α promoter. All studies were conducted in a blinded manner on age-matched and sex-balanced littermate offspring from matings between NTG and hemizygous KL mice, unless indicated otherwise. Mice were analyzed in multiple cohorts, detailed in Table S1. Ifenprodil (Tocris) was dissolved in distilled water and Ro 25-6981 (Tocris) in normal saline at 0.5 mg/ml. Drugs were injected intraperitoneally (i.p.) 30 min (ifenprodil) or 10 min (Ro 25-6981) before behavioral testing of mice at 5 mg/kg or 7.5 mg/kg as indicated. All animal studies were approved by the Institutional

Animal Care and Use Committee of the University of California, San Francisco and conducted in compliance with NIH guidelines.

Morris water maze. The water maze pool (diameter, 122 cm) for mice contained white, opaque water (21°± 1°C) with a square, 14-cm^2 platform submerged 2 cm below the surface. Prior to hidden platform training, mice underwent two pre-training trials by swimming through a channel to mount a hidden platform. During hidden platform training, the platform location remained consistent and the drop location varied between trials. Mice received two training sessions, consisting of two trials each, daily for five days. The maximum time allowed per trial was 60 s. For the probe trial, the platform was removed and the mice were allowed 60 s to swim. Following probe testing, mice were tested for their ability to find the platform when marked with a visible cue (15-cm pole placed on the platform) in two training sessions.

Y-maze. After acclimation to the room for 30 min prior to testing, mice were placed in one arm of the Y-maze (three identical arms, 120° apart) and allowed to explore freely for 3 min. Arm entries were recorded by the experimenter, and an alternation was counted any time the mouse entered each of the three arms in successive arm entries; chance alternation was 22%. The Y-maze apparatus was cleaned with 70% alcohol between testing sessions. Percent alternations was calculated.

Contextual fear conditioning. Mice were tested for their ability to remember the context in which they received a foot shock using a trace fear conditioning paradigm. All mice were initially acclimated to the testing chamber for 12 min 24 h before training. During training, mice were placed into the chamber for 12 min. First, baseline freezing activity was recorded for 4 min. Then, a series of 4 cycles (100 s each) of the following was presented: 20-s silence, 20-s tone, 18-s silence, 2-s 0.30-mA foot shock, and 40-s silence. Twenty-four hours later, mice were returned to the same chamber for 3 min without receiving shocks, and the time they spent freezing was recorded for measurement of context memory. The chamber was cleaned with 70% ethanol between mice. A subgroup of mice was tested for context-independent cued conditioning 2 h after context testing. For this purpose, mice were placed into a new dark chamber with a new odor (2% acetic acid), and the time they spent freezing was recorded. First,

baseline freezing was recorded for 3 min. Then, a series of 4 cycles (100 s each) of the following was presented: 20-s silence, 20-s tone and 60-s silence.

Open field. Total activity of mice in the open field was measured with an automated Flex-Field/Open Field Photobeam Activity System (San Diego Instruments, San Diego, CA). Mice were acclimated to the testing room for 30 min before testing. Mice were tested in a clear plastic chamber (41 x 30 cm) for 10 min, with two photobeam arrays measuring movements. The open field was cleaned with 70% alcohol between testing sessions.

Elevated plus maze. After habituation to dim lighting in the testing room, mice were placed at the center of the apparatus (Hamilton-Kinder, Poway, CA) at the junction between open and closed arms of the maze and allowed to explore for 5 min. The maze was cleaned with 70% alcohol between testing sessions. Time spent in open arms was calculated.

Protein extraction. Total lysates were obtained by homogenizing and sonicating microdissected hippocampus from mice in ice-cold lysis buffer (1X PBS pH 7.4, 1 mM DTT, 0.5 mM EDTA, 0.5% Triton, 0.1 M phenylmethyl sulfonyl fluoride (PMSF), protease inhibitor mixture (Roche), and phosphatase inhibitors 2 & 3 (Sigma)). Following centrifugation at 9,400 x *g* for 10 min at 4° C, supernatant from each sample was collected for measurement of protein concentrations and western blot analyses.

Synaptic membrane fractionation. Separation of synaptic membrane fractions was performed as described (Goebel-Goody et al., 2009; Li et al., 2011) with minor modifications. This protocol enriches for post-synaptic density (PSD)-associated synaptic membranes and excludes intracellular membrane sources (Goebel-Goody et al., 2009). Brains were perfused with 0.9% normal saline, and hippocampus or cortex was dissected and homogenized in 500 µl ice-cold sucrose homogenization buffer containing (in mM): 320 sucrose, 10 Tris (pH 7.5), 1 Na_3VO_4, 5 NaF, 1 EDTA, and 1 EGTA. Homogenization was performed in a glass–grinding vessel using a rotating Teflon pestle (2500 rpm) with at least 20 passes to create a Dounce homogenate. The homogenate was centrifuged at 1000 x *g* for 10 min to remove nuclei and incompletely homogenized material (P1). The resulting supernatant (S1) was spun at 10,000 x *g* for

15 min to obtain a P2 pellet. The P2 was subsequently resuspended in 167 µl sucrose buffer using a motorized pestle mixing/grinding rod (Kontes) directly in the microfuge tube with 30 pulses. The P2 was then subjected to detergent extraction by adding 8 volumes of Triton X-100 buffer (final Triton concentration: 0.5% v/v) containing the following (in mM): 10 Tris (pH 7.5), 1 Na_3VO_4, 5 NaF, 1 EDTA, and 1 EGTA. The suspension was incubated at 4° C for 20 min with gentle rotation and then centrifuged at 32,000 x *g* for 20 min. We operationally defined the pellet as the PSD fraction and the supernatant as the non-PSD (pre-synaptic and peri-synaptic) fraction. The pellet was resuspended in 200 µl sucrose buffer, and proteins were solubilized by the addition of SDS (final concentration 1% v/v), sonication, and heating at 100° C for 5 min. Because the non-PSD fraction was too dilute for biochemical analysis, non-PSD proteins were concentrated by acetone precipitation. Eight volumes of 100% acetone were added to the non-PSD fraction, and the mixture was gently vortexed and incubated at -20° C overnight. The concentrated protein precipitate was then collected by a 15-min spin at 3000 x *g*. The acetone supernatant was decanted and the precipitate was centrifuged at 3000 x *g* for 1 min to remove any excess acetone. The precipitate was air-dried at room temperature for at least 15 min and then resuspended in 200 µl sucrose buffer. Proteins were solubilized by the addition of SDS (final concentration 1% v/v), sonication, and heating at 100° C for 5 min. Protein concentrations were determined by Bradford assay.

Immunoblotting. For detection of klotho and actin in mouse hippocampal homogenates, 15 µg of protein was loaded into each well of a 4–12% gradient SDS–PAGE gel. Gels were transferred to nitrocellulose membranes and immunoblotted with antibodies against klotho (KM2076, 1:1000, TransGenic, Japan) or actin (1:3000, Sigma-Aldrich). For each experiment, the loading control (actin) was quantified independently to ensure it did not differ among groups (data not shown). For detection of GluN2B, synaptophysin, and PSD-95 in synaptic membrane fractions, 5 µg of protein was loaded into each well of a 4-20% Tris-HCl gradient SDS-PAGE gel. Gels were transferred to nitrocellulose membranes and immunoblotted with antibodies against GluN2B (06-600, 1:1000, Millipore), GluN2A (05-901R, 1:1000, Millipore), GluN2C (ab110, 1:500, Abcam), GluN1 (AB9864R, 1:1000, Millipore), GluR1 (MAB2263, 1:1000, Millipore), GluR2 (AB1768, 1:4000, Millipore), synaptophysin (MAB5258, 1:1000, Millipore), or PSD-95 (610496,

1:500, BD Biosciences). Quantifications were performed as described (Palop et al., 2003).

Quantitative Reverse Transcription (RT)-PCR. Total RNA was isolated from dissected hippocampus with RNeasy Mini kits (Qiagen, Valencia, CA). Total RNA was reverse transcribed with random hexamers and oligo(dT) primers. The expression level of *GluN2B* relative to *GAPDH* was determined by SYBR green dye chemistry and an ABI Prism 7900 sequence detector (Applied Biosystems, Foster City, CA), as recommended by the manufacturer. The quality of primers and amplification reactions was verified by analysis of dissociation curves, slopes of standard curves, and reactions without RT. The following primers were used: mouse *GluN2B* (forward, 5'–TGCGCTCTCCCTTAATCTG–3'; reverse, 5'–TGCGCTCTCCCTTAATCTG–3'; and mouse *GAPDH* (forward, 5'–GGGAAGCCCATCACCATCTT–3'; reverse,5'–GCCTTCTCCATGGTGGTGAA–3').

Immunohistochemistry. Immunohistochemistry for FOS was performed as described (Palop et al., 2003) on floating 30-μm sliding microtome sections. Primary anti-FOS rabbit antibody (1:5000; EMD, Germany) was detected with biotinylated donkey anti-rabbit. Sections were then incubated with an avidin-biotin complex (Vector, CA) before development with 3,3'-diaminobenzidine tetrahydrochloride (Sigma-Aldrich, MO). Images were acquired with a digital microscope (Axiocam, Carl Zeiss). Relative numbers of FOS-IR granule cells were determined by counting FOS-IR cells in the granular layer of the dentate gyrus in every 10th serial coronal section throughout the rostrocaudal extent of the granular layer.

Acute brain slices. Brains of 3–4-month-old mice were quickly removed and immersed in ice-cold cutting solution containing (in mM) 234 sucrose, 2.5 KCl, 1.25 NaH_2PO_4, 10 $MgSO_4$, 26 $NaCO_3$, 11 glucose and 1.3 ascorbic acid, and oxygenated with 95% O_2/5% CO_2. Transverse slices of 400 μm were cut on a Leica VS100 vibroslicer (Leica, Germany) and incubated at 32° C for 30 min in an interface incubation chamber (Automated Scientific, CA). Following a 1-h incubation at room temperature, slices were transferred to a submerged recording chamber and continuously perfused with oxygenated artificial cerebrospinal fluid (aCSF) containing (in mM) 126 NaCl, 2.5 KCl,

1.25 NaH_2PO_4, 1 $MgSO_4$, 26 $NaCO_3$, 10 glucose and 2 $CaCl_2$ at 3ml/min (31–32°C). Recording began after slices had equilibrated for 10–20 min.

Electrophysiology. Acute transverse slices (400 μm) were prepared from mice. Brains of 3–4-month-old mice were quickly removed and immersed in ice-cold cutting solution containing (in mM) 234 sucrose, 2.5 KCl, 1.25 NaH_2PO_4, 10 $MgSO_4$, 26 $NaCO_3$, 11 glucose and 1.3 ascorbic acid, and oxygenated with 95% O_2/5% CO_2. Transverse slices of 400 μm were cut on a Leica VS100 vibroslicer (Leica, Solms, Germany) and incubated at 32° C for 30 min in an interface incubation chamber (Automated Scientific, Berkeley CA). Following a 1-h incubation at room temperature, slices were transferred to a submerged recording chamber and continuously perfused with oxygenated artificial cerebrospinal fluid (aCSF) containing (in mM) 126 NaCl, 2.5 KCl, 1.25 NaH_2PO_4, 1 $MgSO_4$, 26 $NaCO_3$, 10 glucose and 2 $CaCl_2$ at 3ml/min (31–32°C). Recording began after slices had equilibrated for 10–20 min.

For field potential recordings in acute hippocampal slices, a glass pipette Ag/AgCl electrode (2–3 MΩ) filled with aCSF was placed halfway between the end of the granule cell layer and the vertex of the two blades of the dentate gyrus, ~50 μm from the granule cell layer of the dorsal blade to stimulate the medial perforant pathway. Field EPSPs were recorded with glass electrodes (2–3 MΩ) filled with aCSF and placed at the same distance from the granule cell layer but 300–350 μm closer to CA3. Both electrodes were advanced into the slice 20 μm at a time until a maximal response was recorded. Test stimuli of 50-μs duration and varying intensity (10–100 μA) were delivered from a Master 8 stimulator (A.M.P.I, Israel). Stimulus was delivered at 0.033 Hz. Stimulation of the medial perforant path was confirmed by paired-pulse depression at the beginning of each experiment. A 15-min stable baseline of fEPSP was established (at an intensity to evoke 30–35% of the maximal response) before induction of LTP. The means of every three fEPSPs were used for final analysis and plotting. LTP was induced by theta-burst protocol (four trains delivered every 20 s, each train containing 10 bursts at 5 Hz, and each burst containing four pulses at 200 Hz), while the stimulus intensity was elevated by increasing the pulse width to 100 μs. Bicuculline (2.5 μM, Tocris) was included in aCSF to facilitate induction of LTP.

To measure AMPA/NMDA receptor EPSC ratios, the perforant pathway was stimulated with a concentric bipolar platinum/iridium electrode ~200 μm from the patch pipette, and evoked whole-cell currents were recorded from visually-identified dentate

granule cells in the presence of 100 μM picrotoxin (Sigma, St Louis, MO) and 5 μM bicuculline (Tocris, Bristol, United Kingdom). Internal pipette solution contained (in mM) 120 $CsMeSO_3$, 4 NaCl, 2 $MgCl_2$, 10 HEPES, 10 BAPTA, 5 MgATP, and 0.3 Na_3GTP. The value of AMPAR-mediated EPSCs was measured at the peak of the trace (holding potential -70 mV) and the value of NMDA receptor-mediated EPSCs was measured 80–100 ms after the peak of the trace (+40 mV), when the AMPA receptor component had completely subsided. The decay time constant (τ) was generated by fitting the decaying phase of isolated NMDA receptor EPSCs to a single exponential equation.

For measurement of spontaneous EPSCs (sEPSCs), the brains were sliced in solution containing (in mM) 93 N-Methyl-D-glucamine (NMDG), 2.5 KCl, 1.2 NaH_2PO_4, 30 $NaHCO_3$, 10 $MgSO_4$, 0.5 $CaCl_2$, 20 HEPES, 25 glucose, 5 L-ascorbic acid, 2 thiourea, 3 sodium pyruvate and 12 N-acetyl-cysteine (pH 7.4). The slices were incubated in the same solution at 37^{o}C for 15 minutes before they were transferred to and held in recovery solution containing (in mM) 92 NaCl, 2.5 KCl, 1.2 NaH_2PO_4, 30 $NaCO_3$, 2 $MgSO_4$, 2 $CaCl_2$, 20 HEPES, 25 glucose, 5 L-ascorbic acid, 2 thiourea, 3 sodium pyruvate and 12 N-acetyl-cysteine at room temperature until recording. sEPSC were recorded from dentate granule cells at -70mV in the presence of 100 μM picrotoxin and 50 μM D-APV (Tocris). The recorded sEPSCs were analyzed by clampfit 10.2 (Axon Instruments, Union City, CA).

Glass pipette electrodes were pulled from borosilicate glass capillaries (WPI Inc, Sarasota, FL) on a horizontal microelectrode puller (P-97, Sutter Instruments, Novato, CA). Electrophysiological recordings were filtered at 2 kHz, digitized at 20 kHz by a Multiclamp 700A amplifier (Axon Instruments) and acquired with a Digidata-1322A digitizer and WinLTP program (WinLTP Inc, University of Bristol, United Kingdom). Offline analysis was performed with WinLTP and Original pro 8.0 (Origin Labs).

Statistical Analyses. Experimenters were blinded to the genotypes of humans and the genotypes and treatment of mice. Statistical analyses were performed using GraphPad Prism (version 5.0) for t-tests, log-rank tests (for survival analyses), and repeated measures ANOVA. R (nlme package) (R Development Core Team, 2011) was used for mixed model ANOVAs, post-hoc tests, principle component analysis, linear models, and power analyses. Differences between two means were assessed by t-tests. For human studies, linear models including age, sex, and education, KL-VS carrier status (0 or 1 allele), and relevant interactions of these terms were fit, with or without *APOE* ε4 carrier

status, to identify statistically significant predictors of cognitive performance and klotho levels found in serum. Statistical significance of each variable was assessed using linear regression t-tests (equivalent to a type IV or III ANOVA). Power analyses were performed to determine the sensitivity of our data to detect significant changes when data showed trends. For mouse studies, exclusion criteria (greater than 2 standard deviations above or below the mean) were defined *a priori* to ensure unbiased exclusion of outliers. For Morris water maze hidden-platform training analyses, LTP, and NMDA receptor EPSC decay analyses, a mixed model ANOVA (factors: genotype and day, time, or intensity) including effects of repeated measures was used as described (Young et al., 2009). This model accounted for correlated responses from repeated measures from each mouse. Day (for water maze training) and time (for LTP and NMDA receptor EPSC decay) were treated as continuous variables. Unless indicated otherwise, t-tests used in mouse studies were one-tailed since we hypothesized that klotho improved cognitive or related measures. ANOVAs were followed by post-hoc tests. We corrected for the multiple comparisons of post-hoc tests with the Bonferroni-Holm (step-wise Bonferroni) procedure (Holm, 1979) to control for family-wise error rate at a level of α=0.05. Survival analyses were performed with log-rank tests. Proportional hazard testing (Grambsch and Therneau, 1994) was performed on survival curves using R to determine whether genotype effects on survival were age-dependent. Error bars represent SEM. Null hypotheses were rejected below a p-value of 0.05.

SUPPLEMENTAL REFERENCES:

Bennett, D.A., Schneider, J.A., Buchman, A.S., Barnes, L.L., Boyle, P.A., and Wilson, R.S. (2012). Overview and findings from the rush Memory and Aging Project. Curr Alzheimer Res *9*, 646–663.

Drag, L.L., and Bieliauskas, L.A. (2010). Contemporary review 2009: Cognitive aging. J Geriatr Psychiatry Neurol *23*, 75–93.

Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M.*, et al.* (2007). A second generation human haplotype map of over 3.1 million SNPs. Nature *449*, 851–861.

Goebel-Goody, S.M., Davies, K.D., Alvestad Linger, R.M., Freund, R.K., and Browning, M.D. (2009). Phospho-regulation of synaptic and extrasynaptic N-methyl-d-aspartate receptors in adult hippocampal slices. Neuroscience *158*, 1446–1459.

Grambsch, P.M., and Therneau, T.M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. Biometrika *81*, 515–526.

Gunstad, J., Paul, R.H., Brickman, A.M., Cohen, R.A., Arns, M., Roe, D., Lawrence, J.J., and Gordon, E. (2006). Patterns of cognitive performance in middle-aged and older adults: A cluster analytic examination. J Geriatr Psychiatry Neurol *19*, 59–64.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scand J Statist *6*, 65–70.

Kramer, J.H., Jurik, J., Sha, S.J., Rankin, K.P., Rosen, H.J., Johnson, J.K., and Miller, B.L. (2003). Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. Cogn Behav Neurol *16*, 211–218.

Kuro-o, M., Matsumura, Y., Aizawa, H., Kawaguchi, H., Suga, T., Utsugi, T., Ohyama, Y., Kurabayashi, M., Kaname, T., Kume, E.*, et al.* (1997). Mutation of the mouse klotho gene leads to a syndrome resembling ageing. Nature *390*, 45–51.

Li, S., Jin, M., Koeglsperger, T., Shepardson, N.E., Shankar, G.M., and Selkoe, D.J. (2011). Soluble Aβ oligomers inhibit long-term potentiation through a mechanism involving excessive activation of extrasynaptic NR2B-containing NMDA receptors. J Neurosci *31*, 6627–6638.

Pa, J., Possin, K.L., Wilson, S.M., Quitania, L.C., Kramer, J.H., Boxer, A.L., Weiner, M.W., and Johnson, J.K. (2010). Gray matter correlates of set-shifting among neurodegenerative disease, mild cognitive impairment, and healthy older adults. J Int Neuropsychol Soc *16*, 640–650.

Palop, J.J., Jones, B., Kekonius, L., Chin, J., Yu, G.-Q., Raber, J., Masliah, E., and Mucke, L. (2003). Neuronal depletion of calcium-dependent proteins in the dentate gyrus is tightly linked to Alzheimer's disease-related cognitive deficits. Proc Natl Acad Sci USA *100*, 9572–9577.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet *38*, 904–909.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J.*, et al.* (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet *81*, 559–575.

R Development Core Team (2011). R: A language and environment for statistical computing (Vienna, Austria, R Foundation for Statistical Computing).

Scott, L.J., Mohlke, K.L., Bonnycastle, L.L., Willer, C.J., Li, Y., Duren, W.L., Erdos, M.R., Stringham, H.M., Chines, P.S., Jackson, A.U.*, et al.* (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science *316*, 1341–1345.

Semba, R.D., Cappola, A.R., Sun, K., Bandinelli, S., Dalal, M., Crasto, C., Guralnik, J.M., and Ferrucci, L. (2011). Plasma klotho and mortality risk in older community-dwelling adults. J Gerontol A Biol Sci Med Sci *66*, 794–800.

Small, G.W., Kepe, V., Ercoli, L.M., Siddarth, P., Bookheimer, S.Y., Miller, K.J., Lavretsky, H., Burggren, A.C., Cole, G.M., Vinters, H.V.*, et al.* (2006). PET of brain amyloid and tau in mild cognitive impairment. N Engl J Med *355*, 2652–2663.

Small, G.W., Siddarth, P., Kepe, V., Ercoli, L.M., Burggren, A.C., Bookheimer, S.Y., Miller, K.J., Kim, J., Lavretsky, H., Huang, S.C.*, et al.* (2012). Prediction of cognitive decline by positron emission tomography of brain amyloid and tau. Arch Neurol *69*, 215–222.

Wilson, R.S., Beckett, L.A., Barnes, L.L., Schneider, J.A., Bach, J., Evans, D.A., and Bennett, D.A. (2002). Individual differences in rates of change in cognitive abilities of older persons. Psychol Aging *17*, 179–193.

Yamazaki, Y., Imura, A., Urakawa, I., Shimada, T., Murakami, J., Aono, Y., Hasegawa, H., Yamashita, T., Nakatani, K., Saito, Y.*, et al.* (2010). Establishment of sandwich ELISA for soluble alpha-Klotho measurement: Age-dependent change of soluble alpha-Klotho levels in healthy subjects. Biochem Biophys Res Commun *398*, 513–518.

Young, M.E., Clark, M.H., Goffus, A., and Hoane, M.R. (2009). Mixed effects modeling Morris water maze data: Advantages and cautionary notes. Learn Motiv *40*, 160–177.