# Hippocampal Gene Expression Meta-Analysis Identifies Aging and Age-Associated Spatial Learning Impairment (ASLI) Genes and Pathways

# Raihan K. Uddin\*, Shiva M. Singh\*

Molecular Genetics Unit, Department of Biology, University of Western Ontario, London, Ontario, Canada

#### Abstract

A number of gene expression microarray studies have been carried out in the past, which studied aging and age-associated spatial learning impairment (ASLI) in the hippocampus in animal models, with varying results. Data from such studies were never integrated to identify the most significant ASLI genes and to understand their effect. In this study we integrated these data involving rats using meta-analysis. Our results show that proper removal of batch effects from microarray data generated from different laboratories is necessary before integrating them for meta-analysis. Our meta-analysis has identified a number of significant differentially expressed genes across age or across ASLI. These genes affect many key functions in the aged compared to the young rats, which include viability of neurons, cell-to-cell signalling and interaction, migration of cells, neuronal growth, and synaptic plasticity. These functional changes due to the altered gene expression may manifest into various neurodegenerative diseases and disorders, some of which leading into syndromic memory impairments. While other aging related molecular changes can result into altered synaptic plasticity simply causing normal aging related non-syndromic learning or spatial learning impairments such as ASLI.

Citation: Uddin RK, Singh SM (2013) Hippocampal Gene Expression Meta-Analysis Identifies Aging and Age-Associated Spatial Learning Impairment (ASLI) Genes and Pathways. PLoS ONE 8(7): e69768. doi:10.1371/journal.pone.0069768

Editor: Manlio Vinciguerra, University College London, United Kingdom

Received April 15, 2013; Accepted June 14, 2013; Published July 18, 2013

**Copyright:** © 2013 Uddin, Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors acknowledge support from the Natural Sciences and Engineering Research Council grant and the University of Western Ontario. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

\* E-mail: ruddin@uwo.ca (RKU); ssingh@uwo.ca (SMS)

## Introduction

Aging and age-associated cognitive impairments are complex and multifactorial and involve both genetic as well as environmental determinants. Both in humans and in animal models the process of normal aging often results in cognitive decline, with or without the presence of any aging related neurological disorders. Disorders related to cognitive impairments range from non-syndromic benign senescent forgetfulness to the syndromic memory loss that characterizes Alzheimer's disease [1,2,3]. These manifestations are highly heterogeneous and individual, family, and population specific. They continue to increase with the current trend in longevity in most populations [4,5,6]. As such they are emerging as a major societal challenge. Attempts in the last decade to gain insight into aging and age-associated learning impairments have been aided by advances in genome-wide methods and technologies, particularly gene expression involving microarrays. Further, the hippocampus in the brain is integral to memory function including spatial memory both in humans and in rodents [7,8]. It is greatly affected by aging, and is among the first to be affected during dementia [9,10,11,12]. The microarray technology has been used widely, more specifically, to understand the gene expression changes related to aging and age-associated memory impairments in the hippocampus in humans [13] using post-mortem tissues [14] and in animal models such as rodents after behavioural training [4,9,15]. Results show that learning

induces a complex reprogramming of gene expression, which is also affected by the aging processes. Moreover, the results of the individual studies are heterogeneous and often difficult to interpret. They often highlight different gene sets and pathways, have limited conclusions, and do not consider their broader implications that may go beyond individual experiments. It is therefore desirable to integrate results from these studies towards a consensus view of the genes affected and the molecular mechanisms underlying brain aging and age-associated learning impairments. This is now possible because of the availability of considerable amount of original microarray data in the public microarray data repositories, as well as the availability of improved statistical analytical methods. This study focuses on age-associated spatial learning impairment (ASLI). It uses original results from all available microarray gene expression data involving ASLI in rats using an inverse-variance meta-analysis approach [16]. The results establish that a large number of genes are differentially expressed across age and across spatial learning impairment. More importantly, they allow identification of pertinent lists of aging and ASLI related genes. Further, the follow up analysis has offered a novel insight into the underlying molecular pathways associated with aging and age-related non-syndromic memory impairments such as ASLI.

#### Methods

#### Data Selection

In order to reduce heterogeneity among studies selected for this meta-analysis we followed a conservative data selection process. We focused on datasets generated from carefully designed behavioural studies involving hippocampus dependent ASLI in Fischer 344 strain of male rats (*Rattus norvegicus*) as assessed by the Morris Water Maze (MWM). We mainly used the GEO (http://www.ncbi.nlm.nih.gov/geo/) and the ArrayExpress (http://www.ebi.ac.uk/array express/) microarray data repositories to search for microarray gene expression datasets using the keyword "memory and brain". We also used the PubMed literature database to search for relevant studies (Figure 1). Affymetrix raw data (CEL files) for the selected studies were either directly downloaded from the GEO website or obtained through personal communication with the original authors.

#### Data Preprocessing

All arrays were first assessed for image quality using the dChip software [17] (http://biosun1.harvard.edu/complab/dchip/). Minor contaminations present in a few of the arrays were corrected using the built in image gradient correction algorithm in dChip. The data quality was assessed using the RNA degradation ratios, relative log expression (RLE), and normalized unscaled standard errors (NUSE) plots using the simpleaffy and affyPLM packages in Bioconductor (http://www.bioconductor.org/) following standard procedures [18]. Arrays with bad quality e.g. variable background brightness, uneven hybridization, etc., or arrays having greater than 15% array outlier values were excluded. Within-study normalization and expression measurement were performed using the RMA methods [19] with default options in the affy package in R [20]. Within-study batch correction was performed using the Empirical Bayes method also known as the ComBat [21], which has been shown to produce better results than other comparable methods [22,23]. Array hybridization dates were retrieved from CEL files and used as processing batches to perform batch



**Figure 1. Data selection process.** Search in the public microarray data repositories identified 38 microarray datasets involving cognitive impairments. We excluded 19 datasets that were either not relevant to our study or were not associated with any publication. We excluded 14 more studies as they involved different learning paradigms, test conditions, and outcomes in mice. We finally selected five studies that dealt with hippocampus dependent age-associated spatial learning in rats. doi:10.1371/journal.pone.0069768.g001

correction. Age and spatial learning impairment were used as covariates.

#### Data Integration

A common probe-set file that contains best matching pairs of probe-sets representing the same gene in the two chip types i.e. rgu34a and rae230a was downloaded from the Affymetrix website (www.affymetrix.com). Applying the common file, probe-sets from all studies belonging to the two different chip types were merged into three categories as follows: i) rg\_exclu, probe-sets exclusive to the rgu34a chip type, ii) all5\_com, probe-sets common among all five studies, and iii) rae\_exclu, probe-sets exclusive to the rae230a chip type. Each probe-set specific data and their analysis outcome from all studies were combined in two ways (Figure 2): a) effect size integration, which combined the preprocessed data first before any analysis.

a) Effect size (ES) integration. Meta-analysis, which combines the results of independent but related studies in a relatively inexpensive way, has the ability to increase the statistical power to obtain a more precise estimate of gene expression differences. Though there are many ways to combine the results across studies, platform, and species [23,24,25], combining effect sizes using an inverse-variance method [26,27] is considered to be the most comprehensive approach for meta-analysis of gene expression microarrays [24]. Therefore, we estimated effect sizes on the within-study batch-corrected data using the random effect size (ES) model as follows. First, for each probe-set, study-specific sample sizes, mean expression measures, and standard deviations were computed for each comparison. In order to understand the effect of age and spatial learning impairment, data were analyzed in two ways, e.g. by comparing samples across age (aged vs. young, AY) and across learning impairment (aged-impaired vs. agedunimpaired, IU), respectively. Next, the meta package in R (http:// cran.r-project.org/web/packages/meta/meta. pdf) was used to calculate each study-specific standardized mean difference (SMD) (Cohen's d) for each probe-set, and later, probe-set SMDs for all studies in each category (e.g. rg\_exclu, all5\_com, and rae\_exclu) were pooled utilising Hedges' adjusted g [16] to obtain the final random ES for each probe-set. Effect size values for all probe-sets from all three categories were then combined together, annotated, and summarized. Duplicate probe-sets and multiple probe-sets annotated to the same gene were summarized by keeping the probe-set with the lowest p-value (of the z-value) for the gene [28]. Uninformative probe-sets were filtered out by removing probe-sets whose expression values had a coefficient of variation of zero across all arrays and probe-sets with a p-value (of the ES z-value) greater than 0.1. The p-values of the treatment effect for all probesets were adjusted for Benjamini and Hochberg (BH) multiple testing correction [29] in R.

**b) Direct data integration.** This was done by a cross-study and cross-platform normalization process by first combining data separately for each category (e.g. rg\_exclu, all5\_com, and rae\_exclu) and then adjusting data across all studies. For each category, data were adjusted similarly as within-study batch correction, however, considering individual studies as separate batches. Next, differential expression (DE) analysis was performed by comparing the data in two ways as above e.g. AY and IU using the *limma* software package [30]. Significantly differentially expressed genes from all three categories were combined together, annotated, and summarized as described above. Duplicate and multiple probe-sets issues and multiple testing corrections were also handled similar to the ES analysis.

### Functional and Pathway Analysis

Functional and pathway analysis was performed mainly using the Ingenuity Pathway Analysis (IPA) software (http://www. ingenuity.com). Data sets containing identifiers of significant (pvalue  $\leq 0.05$ ) differentially expressed genes from AY or IU comparisons with their corresponding ES estimates (as foldchange values) and p-values were used as input. Identifiers that were successfully mapped to their corresponding objects in the IPA knowledge base were considered for functional, network, and canonical pathway analysis.

For functional analysis the mapped identifiers that were associated with biological functions and/or diseases in the IPA Knowledge Base were considered. Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function and/or disease assigned to the data set is due to chance alone. The expression levels (up- or downregulation) for all of the input genes in each function annotation category were compared with the information stored for those genes in the IPA Knowledge Base to predict whether the expression patterns correspond to the activation state (decreased or increased) for that category.

For network analysis the mapped identifiers were overlaid onto a global molecular network developed from information contained in the IPA Knowledge Base. Networks of network eligible molecules were then algorithmically generated based on their connectivity. Next, the functional analysis of a network identified the biological functions and/or diseases that were most significant to the molecules in the network based on the association of the network molecules with the biological functions and/or diseases in the Ingenuity Knowledge Base. Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function and/or disease assigned to that network is due to chance alone.

Canonical pathways analysis identified the pathways from the IPA library of canonical pathways that were most significant to the gene lists. All the mapped identifiers from the data set that were associated with a canonical pathway in the Ingenuity Knowledge Base were considered for the analysis. The significance of the association between the data set and the canonical pathway was measured in two ways: a) a ratio of the number of molecules from the data set that map to the pathway divided by the total number of molecules that map to the canonical pathway is displayed. b) Fisher's exact test was used to calculate a p-value determining the probability that the association between the genes in the dataset and the canonical pathway is explained by chance alone.

### Results

#### Data Selection, Preprocessing, and Integration

Search in the public microarray data repositories reveals that there is a large body of microarray data available involving cognitive impairments (Figure 1). Review of the resulting articles reveals that the goals of these studies are varied and include different learning paradigms, test conditions, subjects, and tissue types. After careful examinations of these datasets and following suitable data selection guidelines (see method), we identified five individual studies consisting of a total of 287 arrays (one animal per assay), which used two different Affymetrix chip types, RG\_U34a and RAE230a (Table 1). The data represented young and aged rats that were learning unimpaired and aged rats that were learning impaired from a set of results published during 2003 to 2009. The selected datasets referred to as BL [31], B7 [4], R7 [32], B8 [15], and K9 [33] in this study allowed us to assess a combined gene expression changes related to aging, as well as



**Figure 2. A summary of the meta-analysis workflow.** Five individual studies (BL, B7, R7, B8, and K9) were selected for this meta-analysis. The studies involved two different array platforms, Affymetrix RG-U34a and RAE-230a. Following preprocessing, data were integrated across studies and across array platforms and analyzed in two ways: meta-analysis using random effect size model and differential expression analysis using the *limma* software. Top significant genes were used to identify enriched functions and pathways and to construct knowledge based gene regulatory networks using the lngenuity Pathway Analysis software.

doi:10.1371/journal.pone.0069768.g002

ASLI in rats across multiple studies. These studies investigated spatial learning tasks in young (3–6 months old) and aged (24–26 months old) animals using the MWM as the training and assessment protocol. The BL and K9 studies were similar in design where only the unimpaired young and impaired aged animals were considered for comparison. The B7, R7, and B8 studies were similar in design where both young and aged groups had impaired and unimpaired animals as well as additional controls, e.g. cage controls, stress controls, and controls for visual impairment. A total of 265 arrays were finally selected following quality assessment (see methods, data preprocessing).

Hierarchical clustering analysis with normalized data shows that batch effects are clearly evident in all studies even after normalization (see Figures 3 and 4 for some representative results). Arrays that were hybridized on the same date as a batch are clustered together in the dendrogram. We used an Empirical Baves method implemented in ComBat to remove batch effects. Batch effects were completely removed from the BL, B7, and K9 data and considerably removed from the B7 and B8 data. Data were integrated between the rgu34a chip which had a total of 8799 probe-sets and rae230a chip that had a total of 15923 probe-sets. After data integration, the rg\_exclu category contained 2356 probe-sets exclusive to the rgu34a array only. The all5\_com category included 6384 rgu34a unique probe-sets mapping to 5435 rae230a unique probe-sets that are common among all five studies. Finally, the rae\_exclu category contained 10,431 probesets exclusive to the rae230a array type.

#### Gene Identification and Functional Analysis

a) Aged vs. young (AY). In order to assess the effect of aging, a comparison was made between aged vs. young animals. After combining probe-sets from all three categories and after summarization we had ES estimates for 10,619 unique annotated genes. After filtering, there were 3235 genes left, of which 2245 genes were found significant with a p-value  $\leq 0.05$  (Table S1) and 1753 genes were found significant after BH multiple testing correction (pBH.ES  $\leq 0.05$ ). Among the 1753 genes, 874 genes have an I<sup>2</sup> (ratio of true heterogeneity to total variation) value of 0% while 1347 genes have an I<sup>2</sup> value under 40%. Table 2 shows the top 10 most up- and down-regulated genes in the aged animals compared to the young animals. The forest plots of two representative genes *C3* (complement component) (up-regulated) and *Tubb2b* (tubulin, beta 2B class IIb) (down-regulated) are presented in Figure 5A and B, respectively. DE analysis was performed on the data sets in parallel to the ES analysis. Using the 3235 genes from the ES analysis, the log fold-change (logFC) and p-values (pDE) for the corresponding differentially expressed genes were pulled out and BH adjusted similar to that of ES data. This resulted in a total of 1946 genes (pDE <0.05) and 1569 genes after BH adjustment (pBH.DE <0.05).

Functional and Pathway analysis were performed using the IPA software. For this analysis, we considered the significant genes based on unadjusted p-value (pES  $\leq 0.05$ ) of the random effect size, which resulted in a total of 2245 genes. These genes were used as input in the IPA of which 2240 were mapped to their corresponding object in the IPA Knowledge Base. The functional analysis identified the biological functions and/or diseases that were most significant (activation z-score value-cutoff of 1.980) to the mapped gene list. The IPA functional analysis predicts that comparatively more functions are decreased than increased in the aged animals. Table 3 shows a summary of the most significant functions, increased or decreased, as predicted by the IPA algorithm based on the expression levels of the genes in the dataset. The results show that the functions that are specifically decreased include cell viability of central nervous system cells, formation of cells, quantity and synthesis of inositol phosphate, and axonogenesis. Thus they affect the cell death and survival, cellular growth and proliferation, carbohydrate metabolism, molecular transport, small molecule biochemistry, cell morphology, and nervous system development and function in the aged animals. Major functions categories that see an increase are cellular movement, cellular development, and connective tissue development and function. The specific functions of the genes in this category include the migration of cells and differentiation of chondrocytes. We have generated biological knowledge based gene interaction networks for the AY significant genes. A representative network graph is presented in Figure 6, which shows the network interactions of some of the aging and learning genes. A summary of the functions for the top five most significant networks is given in Table 4. The most critical canonical pathways that are affected in the aged animals include Eif2 (eukaryotic translation initiation factor 2) signaling, antigen presentation, and Ox40 (tumor necrosis factor) signaling pathways (Table 5).

**b)** Aged-impaired vs. aged-unimpaired (IU). In order to assess the effect of ASLI, a comparison was made between the aged-impaired vs. aged-unimpaired (IU) rats where we included three sets of controls (e.g. cage controls, visual controls, and stress controls (no platform during memory test in the water maze)) in

Table 1. Age-associated spatial learning impairment (ASLI) datasets for rats.

Dataset ID	Reference	Affymetrix Array Type	Number of Assays (one animal/array)
BL	Blalock et al. 2003 [31]	RG_U34	29
В7	Burger et. al. 2007 [4]	RG_U34	79
R7	Rowe et. al. 2007 [32]	RAE230a	50
B8	Burger et. al. 2008 [15]	RAE230a	80
К9	Kadish et al. 2009 [33]	RAE230a	49



Figure 3. Boxplot of R7 dataset before (A) and after (B) RMA normalization. Each color represents a batch of arrays that were hybridized and processed at the same time. doi:10.1371/journal.pone.0069768.g003



**Figure 4. Hierarchical clustering of RMA normalized R7 data.** Each color represents a batch of arrays, which were hybridized and processed at the same time. Batch effects are evident even after normalization and before batch adjustment (A) as arrays are mostly clustered in batches. However, following Empirical Bayes adjustment arrays are clustered based on aged and young phenotypes irrespective of batches (B). Leaf labels: A, aged; Y, young; I, impaired; U, unimpaired; c, control. doi:10.1371/journal.pone.0069768.g004



**Figure 5. Forest plots of four representative significant genes.** For the selected probe-set for each gene the individual study specific standardized mean differences (SMD) and their 95% confidence intervals (CI) are plotted and shown on each row. The effect size results are shown at the bottom of each plot. *C3* is up-regulated (A) and *Tubb2b* is down-regulated in the aged rats. *Arc* is down-regulated (C) and *Marcks* is up-regulated (D) in the aged-impaired rats. doi:10.1371/journal.pone.0069768.g005

Table 2. Top ten most up- and down-regulated genes (based on ES) in the AY comparison.

Up-regulated genes							
Probe ID	Symbol	ES	z-value	p-value of z- value	pBH of z-value	LogFC of DE	pBH of DE
1398892_at	Npc2	3.988	3.16	0.002	0.009	0.474	0
X52477_at	C3*	3.812	3.716	0	0.002	0.730	0
X13044_g_at	Cd74*	3.389	3.148	0.002	0.009	0.916	0
M15562_g_at	HLA-DRA*	3.236	3.284	0.001	0.007	1.011	0
1368187_at	Gpnmb*	3.189	2.827	0.005	0.017	0.610	0
L03201_at	Ctss*	3.110	3.362	0.001	0.006	0.368	0
1373575_at	Fcer1g*	2.846	2.821	0.005	0.018	0.473	0
1370885_at	Ctsz	2.606	3.201	0.001	0.008	0.432	0
J03752_at	Mgst1*	2.544	3.229	0.001	0.024	0.362	0
1376652_at	C1qa*	2.519	3.709	0	0.007	0.488	0

Down-regulated genes

Probe ID	Symbol	ES	z-value	p-value of z-value	pBH of z-value	LogFC of DE	pBH of DE
1376319_at	Sema3c*	-3.674	-3.867	0	0.001	-0.588	0
X57281_at	Glra2	-2.589	-4.599	0	0	-0.528	0
1388821_at	Trib2	-2.029	-2.578	0.01	0.028	-0.253	0
1388750_at	Tfrc*	-1.853	-2.576	0.01	0.028	-0.203	0
L03294_at	Lpl*	-1.803	-5.42	0	0	-0.395	0
1374966_at	Dcx*	-1.783	-3.42	0	0.005	-0.262	0
1389533_at	Fbln2	-1.756	-2.332	0.02	0.04	-0.192	0
D45412_s_at	Ptpro*	-1.721	-2.499	0.013	0.032	-0.319	0
M58369_at	Pnlip*	-1.618	-3.26	0.001	0.007	-0.200	0
X03369_s_at	Tubb2b*	-1.607	-2.907	0.004	0.015	-0.174	0

Top genes identified by IPA are indicated by an asterisk (\*). Legends: ES, effect size; pBH, p-value with Benjamini and Hochberg correction; FC, fold change; DE, differentially expressed.



**Figure 6. Network 3 from aged vs. young (AY) comparison.** Major functions of this network are cellular assembly and organization, tissue development, and cell morphology. Each biological relationship (an edge) between two genes (nodes) is supported by at least one reference from the literature or curated information stored in the Ingenuity Knowledge Base. The intensity of the node color indicates the degree of up- (red) or down- (green) regulation observed in the AY comparison. The effect size and p-value for each gene is shown below the gene symbol. Edges are displayed with various labels that describe the nature of relationship between the genes (e.g. P for phosphorylation, PP for protein-protein binding, PD for protein-DNA binding, A for activation, E for expression, L for proteolysis, LO for localization, RB for regulation of binding). Any specific findings for a gene whether it is associated with aging (A), learning (L), and/or spatial learning (SL) is presented inside a rectangle beside that gene. doi:10.1371/journal.pone.0069768.g006

the aged-impaired group as was done in the B7 and B8 studies [4,15]. After combining probe-sets from all three categories and after summarization there were 10,412 unique annotated genes with ES estimates. After filtering out uninformative genes there were 1310 genes left, of which 787 genes were found significant with a p-value  $\leq 0.05$  (Table S2). Among the 787 genes, 59 were significant with adjusted pBH.ES  $\leq 0.05$  and 55 of these genes have an I<sup>2</sup> value of 0%. Table 6 shows the top 10 most up- and down-regulated genes in the aged-impaired as compared to the aged-unimpaired animals. Figure 5C and D show the forest plots of two representative genes Arc (activity-regulated cytoskeleton-associated protein) (down-regulated) and Marcks (myristoylated alanine-rich protein kinase C substrate) (up-regulated). DE analysis for the 1310 IU genes identified 460 significant genes (pDE

 $\leq$ 0.05), of which 92 were significant with pBH.DE  $\leq$ 0.05. However, among the 92 genes significant in the DE analysis, 44 were also present in the ES meta-analysis (pES  $\leq$ 0.05) category and 14 in the pBH.ES  $\leq$ 0.05 category.

A total of 738 genes with significant effect sizes (pES  $\leq 0.05$ ) were used as input for the IU functional analysis in the IPA. Though cell viability of hippocampal neurons and CNS cells, cell-to-cell signaling, and molecular transport were the top functions in the results, none were statistically significant. However, when we reanalyzed with an effect size data set that was generated comparing the expression level of the aged-impaired animals with that of the aged-unimpaired animals without any controls, four functions e.g. molecular transport, cellular development, cellular growth and proliferation, and connective tissue development and

Table 3. Significantly increased or decreased functions in the AY comparison.

Functions Annotation	p-value	Predicted activation state	Activation z-score	High-level functions category	Genes
Cell viability of central nervous system cells	2.22E-03 to 2.35E-02	Decreased	-2.757 to -2.000	Cell death and survival	<b>ApoE<sup>A,L,SL</sup></b> , Atf3, Bdnl <sup>L,SL</sup> , CdkSr1 <sup>L,SL</sup> , Cycs, Hspb1, Ide, Igf2, Ntf3 <sup>L</sup> , PlagI1, Prkcg <sup>L,SL</sup> , <b>Rela<sup>L</sup></b> , Serpini1, Sh3kbp1, <b>Slc11a2<sup>L</sup></b> , Vegfa <sup>L</sup> , Vip
Formation of cells	7.96E-03	Decreased	-2.376	Cellular growth and proliferation	Bdnf <sup>L,SL</sup> , Egr1 <sup>L</sup> , Fgf18, <i>Icam1</i> , Igf2, <i>Nppa, Pf4,</i> <i>S100b</i> <sup>L</sup> , Sdc2, Wt1
Quantity and synthesis of inositol phosphate	1.54E-02	Decreased	-2.186	Carbohydrate metabolism, molecular transport, small molecule biochemistry	Agtr1, Avp <sup>L</sup> , Cckbr, Gal <sup>L</sup> , Gnaq, Grp <sup>L</sup> , <i>Icam1</i> , Mas1, Pthlh, Rgs2, <b>Rgs3,</b> S1pr1, Trhr
Axonogenesis	6.96E-03	Decreased	- 1.980	Cell morphology, assembly and organization, nervous system development and function	<b>Actb,</b> Actr3, Agrn, Bdnf <sup>L,SL</sup> , Cck, <b>Cntn2<sup>⊥</sup>, Igf1r,</b> L1cam <sup>L,SL</sup> , <b>Mbp, Picalm,</b> Ppp2ca, Snap91, Stk11
Migration of cells	8.08E-04 to 4.96E-03	Increased	2.158	Cellular movement	Abcc1, Actr3, Agt, Aif1, Anxa2, Bcar1, Bdnf <sup>4,SL</sup> , C3, Cck, Ccl3l1/Ccl3l3 <sup>L,SL</sup> , Ccl5, Cd44, Cd82, Cdc42, Dnm2, Drd5 <sup>L,SL</sup> , Gucy1a3, Gucy1b3, Icam1, Nfkbia <sup>L</sup> , Ntf3 <sup>L</sup> , Pten <sup>L</sup> , Reln, Stat3, Scpep1, Tac1 <sup>L</sup> , Tgfa, Tgfa, Tgfb1, Tgfb2, Tpm1, Tubb2b, Vegfa <sup>L</sup> , and etc.
Differentiation of chondrocytes	1.54E-02	Increased	2.183	Cellular development, connective tissue development and function	Grn, Por, Rela <sup>L</sup> , Tgfb1, Thrb <sup>L</sup>

Genes in bold were up-regulated and not bold were down-regulated in this analysis. Genes annotated as aging, learning, and spatial learning in the IPA knowledge base are indicated by "<sup>Ar</sup>, "<sup>L</sup>", and "<sup>SL"</sup>, respectively.

doi:10.1371/journal.pone.0069768.t003

function were significantly decreased (results not shown). The specific functions of these genes in these categories include transport of molecules and proliferation of fibroblast cell lines. In addition, growth of neuritis was also decreased among others. Similar to AY, we have generated biological knowledge based gene interaction networks for the IU related genes. A summary of the functions for the top five most significant networks is given in Table 7. The canonical pathways that are most affected in the aged-impaired compared to the aged-unimpaired animals include Nurr77 (nuclear receptor subfamily) signaling in lymphocytes, nNOS (nitric oxide) signaling in neurons, and glutamate receptor signaling (Table 8).

#### Aging and Learning Related Genes

We searched the IPA Knowledge Base for genes that are annotated as aging related and genes that are annotated as learning related, particularly spatial learning. We found a total of 61 genes related to general aging, of which five, *Adraid* (all-trans retinoic acid-induced differentiation factor), *Aldoc* (aldolase C, fructose-bisphosphate), *Clu* (clusterin), *ApoE* (apolipoprotein E), and *Mapt* (microtubule-associated protein tau) (Figure 6) were present in our data set (p.ES  $\leq 0.05$ ). Further, there were 401 genes annotated as learning genes in the IPA Knowledge Base, of which 177 were categorized under spatial learning (SL). Among these learning genes 86 (30 of which were SL related) were present in our AY comparison (Table S3) and 48 (15 of which were SL related) were present in the IU comparison (Table S4) with p.ES  $\leq 0.05$ . They were considered as the ASLI associated genes. Among the 86 genes for AY and 48 genes for IU, 15 were found common in both comparisons.

### Discussion

# Effective Meta-analysis Necessitates Proper Data Preprocessing and Integration

Meta-analysis has emerged as an essential tool in modern genetic and genomic analysis [34]. It can uncover a significant effect from a combined analysis as integration of a broader and/or richer collection of data has the potential to generate results that have greater confidence, and place less reliance on a single dataset [24,34]. Although meta-analysis often includes large number of unrelated studies, we followed a more conservative approach in order to concentrate on microarray gene expression datasets that

Table 4. Major functions associated with the top five networks in the AY comparison.

Network ID	Top functions associated with the networks	IPA score	Total focus genes
1	Molecular transport, cell-to-cell signaling and interaction, nervous system development and function	25	35
2	Endocrine system disorders, gastrointestinal disease, metabolic disease	21	33
3	Cellular assembly and organization, tissue development, cell morphology	17	30
4	Cell-to-cell signaling and interaction, cell signaling, molecular transport	14	28
5	Drug metabolism, protein synthesis, cancer	14	28

Table 5. Top canonical pathways in the AY comparison.

Name	p-value	Ratio
EIF2 signaling pathway	2.36E-07	58/170 (0.341)
Antigen presentation pathway	6.01E-05	14/27 (0.519)
OX40 signaling pathway	1.91E-04	19/60 (0.317)
Chondroitin sulfate degradation pathway	4.96E-03	6/14 (0.429)
IL-17A signaling in gastric cells pathway	5.17E-03	10/24 (0.417)
Complement system pathway	1.69E-02	10/32 (0.312)

doi:10.1371/journal.pone.0069768.t005

focused on the hippocampus dependent ASLI as assessed by MWM test. We started the data preprocessing with raw expression data (CEL files), which gave us the opportunity to perform consistent quality assessment, preprocessing, and filtering of imperfect arrays and outlier values. It also allowed correction of batch effects and removal of any unexplained technical variations. Our results (Figure 3 and 4) confirmed the findings of recent studies [21,35] and demonstrated the necessity of removing batch effects from microarray data before integrating them in any analysis. Next, we performed the random effect size meta-analysis by keeping the individual studies separate and then only combining the probe-set specific effects. We also performed the traditional differential expression analysis in parallel to the ES analysis after merging all probe-set data into a single pool through the process of cross-study and cross-platform data normalization (Figure 2). Even though the DE analysis was able to detect significant DE level, the difference was smaller compared to the ES. Overall, the ES analysis seems to be a better approach than DE analysis, particularly when combining data from different studies and platforms. Nonetheless, the DE results helped us verify the ES outcomes and better screen the aging and ASLI associated genes.

It is important to point out that during the data integration process we worked at the probe-set level rather than at the gene level. This is essential when combining data from independent microarray results from different platforms. Therefore, we integrated all data first before doing any filtering, annotation, and summarization. In the final filtering process we removed genes with a p-value (of the ES z-value) >0.1. Our observation is that, a gene may have a higher ES but not necessarily a lower p-value (Table S1 and S2). This is due to either the heterogeneity among studies or the fact that some datasets are lacking the expression information for that particular probe set. Also the genes whose treatment effect sizes are either zero or close to zero have higher p-

Table 6. Top ten most up- and down-regulated genes (based on ES) in the IU comparison.

# Up-regulated genes

Probe ID	Symbol	ES	z-value	p-value of z-value	pBH of z- value	LogFC of DE	p-value of DE	pBH of DE
1369775_at	Nucks1	1.187	4.105	0	0	0.129	0.008	0.074
S74393_s_at	Рахб	0.881	3.944	0	0.016	0.073	0.014	0.086
M27905_at	Rpl21	0.884	3.965	0	0.016	0.093	0.020	0.1
1388783_at	Hmgb1*	1.124	3.921	0	0.016	0.095	0.055	0.155
U93692_at	Nup88	0.814	3.665	0	0.026	0.083	0.004	0.056
J01436cds_s_at	CYTB	0.827	3.726	0	0.026	0.052	0.119	0.232
1373952_at	Prkag2	1.023	3.610	0	0.033	0.089	0.013	0.084
U78090_s_at	Alg10	0.780	3.522	0	0.033	0.061	0.041	0.135
AB002111_at	Pex12	0.780	3.534	0	0.033	0.100	0.001	0.034
1389373_at	Smad1*	0.949	3.375	0	0.04	0.099	0.045	0.144

Down-regulated genes

Probe ID	Symbol	ES	z-value	p-value of z-value	pBH of z- value	LogFC of DE	p-value of DE	pBH of DE
1390518_at	Emid1	-1.259	-4.314	0	0	-0.063	0.049	0.148
rc_AA891838_at	Mrto4	-0.951	-4.224	0	0	-0.095	0.000	0.013
1389264_at	Ankrd54	-1.149	-3.983	0	0.016	-0.088	0.004	0.056
1369203_at	Wif1*	-0.980	-3.478	0	0.034	-0.056	0.023	0.107
U19866_at	Arc	-0.764	-3.46	0	0.034	-0.215	0.000	0.008
1376569_at	Klf2*	-0.960	-3.405	0	0.04	-0.182	0.000	0.013
rc_AA800613_at	Zfp36	-0.750	-3.396	0	0.04	-0.089	0.008	0.073
1398380_at	Vwa1	-0.94	-3.350	0	0.04	-0.098	0.002	0.038
1368451_at	Hrh3*	-0.940	-3.349	0	0.04	-0.094	0.005	0.063
S49760_g_at	Dgka	-0.711	-3.226	0.0013	0.051	-0.081	0.006	0.063

Top genes identified by IPA are indicated by an asterisk (\*). Legends: ES, effect size; pBH, p-value with Benjamini and Hochberg correction; FC, fold change; DE, differentially expressed.

Network ID	Top functions associated with the networks	IPA score	Total focus genes
1	Neurological disease, tissue morphology	29	27
2	Cellular growth and proliferation, cancer, cell death and survival	16	19
3	Cell-to-cell signaling and interaction, nervous system development and function, carbohydrate metabolism	14	18
4	Cell death and survival, cellular development, hematological system development and function	10	15
5	Cell death and survival, metabolic disease, cellular function and maintenance	8	13

Table 7. Major functions associated with the top five networks in the IU comparison.

doi:10.1371/journal.pone.0069768.t007

values. These genes were therefore filtered out. Our data integration method has prevented any loss of information and generated a number of differentially expressed genes even after multiple testing corrections (Table S1 and S2) particularly for the AY comparison. It is also important to mention that ES estimates of some of these genes e.g. C3 and Tubb2b (Figure 5, A and B) present some degree of heterogeneity. It is not unexpected in a meta-analysis as the heterogeneity may arise, as in this case, from differences in the details of MWM training, memory test and sample collection procedure, and other experimental variables pertaining to the individual studies. However, during the selection of the aging and ASLI related genes that had high heterogeneity, we made sure that the estimates of the ES are on the same direction.

In order to include more genes in the functional and pathway analysis using IPA we considered the unadjusted p-value ( $p \le 0.05$ ) of the random effect size for gene selection. Also, we analyzed data with lower number of genes following more stringent criteria such as using pES ≤0.005 (e.g. 888 genes) or pBH.ES ≤0.05 (e.g. 1753 genes) for AY comparison. It was satisfactory to note that the IPA analysis returned similar results. Also the expression levels (up- or down-regulation) identified in this meta-analysis for all or most of the genes in each function annotation category in the AY comparison correspond to the predicted activation state (decreased or increased) for that category as supported by the literature in the IPA Knowledge Base. Further, we were able to verify the results by literature review using the PubMed. The results support the fact that the genes and pathways identified in this analysis follow biological expectations. The genes identified (see below) are known to partake in aging and in learning impairments. This conclusion is also supported by follow up analysis including regulatory interaction networks based on known functions and interaction. The results obtained are discussed below in the context of aging and learning impairments associated with aging.

Table	8. Top	canonical	pathways	in	the IU	comparison.
IUNIC	<b>0.</b> 10p	curiorneur	putrivuys			companyon.

Name	p-value	Ratio
Nur77 signaling in T lymphocytes	6.13E-04	13/51 (0.255)
nNOS signaling in neurons	5.13E-03	12/46 (0.261)
Glutamate receptor signaling	5.68E-03	13/60 (0.217)
Calcium-induced T lymphocyte apoptosis	1.07E-02	12/57 (0.211)
Glutamate dependent acid resistance	1.48E-02	2/2 (1)

doi:10.1371/journal.pone.0069768.t008

# Effect of Differential Gene Expression on Aging and Learning

Our IPA analysis has revealed major functions and pathways that are affected in the aged and aged-impaired animals. The results show that aging is affected by the genes functioning in cell viability, axonogenesis, and inositol phosphase metabolism. Further, these genes contribute to the imbalance in many major function categories including molecular transport, cell to cell signaling and interaction, and nervous system function. Considering the effect of the most significant differentially expressed genes on cellular biology, these genes could be classified into three distinct but non-exclusive categories: general aging (GA) genes that are associated with aging related disorders and not associated with any learning impairments, general aging genes associated with syndromic learning impairments (GASI), and general aging genes associated with non-syndromic learning impairments (GANSI). Given the confounding effect of aging on learning impairments one may expect an overlap in the three groups of genes. Below we summarize some key findings about some of the genes from each of the above three categories. These genes presented significant up- or down regulation in the AY and IU comparison (Table 2 and 6) and some of them were also identified as contributing to significantly increased or decreased function in the aged animals (Table 3).

**GA or general aging genes.** A majority of the genes that fall into this category were up-regulated in the aged in comparison to the young rats in our analysis and many have been implicated in disease vulnerability at old age in humans and animals. These GA genes may affect a number of pathways including Eif2 signaling, antigen presentation, complement system, and Ox40 signaling pathways (Table 5). EIF2 signaling is activated (through the phosphorylation of eIF2 $\alpha$ ) in response to a wide array of cellular stresses to protects cells by reducing the general rate of protein synthesis while facilitating programs of stress-induced gene expression [36]. OX40 is a member of the tumor necrosis factor (TNF) receptor family and plays a key role in the survival and homeostasis of effector and memory T cells and T-cell-mediated inflammatory diseases [37].

The GA genes that are of special interest to this discussion are C3, Cd74 (CD74 molecule, major histocompatibility complex, class II invariant chain), Ctss (cathepsin S), Ctsz (cathepsins Z), Agt (angiotensinogen), Mbp (myelin basic protein), and Cck (Cholecystokinin). Specifically, C3, Cd74, and Agt expression level was increased (Table 3, migration of cells function) and they affect the endocrine system disorders, gastrointestinal disease, and metabolic disease functions. C3 (Table 2 and Figure 5A) plays a central role in the activation of complement system and is needed to restore tissue injury. However, inappropriate or excessive activation of the complement system can lead to cell death and tissue destruction, thus contributing to further injury and impaired wound healing

[38]. These consequences are clinically manifested in various disorders [39]. Cd74 (Table 2) participates in several key processes of the immune system including antigen presentation, B-cell differentiation, and inflammatory signalling. Overexpression of Cd74 has been reported in some inflammatory diseases and several forms of cancer (reviewed in [40]), and also known as an indicator of disease in some conditions. The longer form of CD74 also interacts with CTSS by direct binding [41], and both Ctss and Ctsz are also highly up-regulated in the aged rats (Table 2). Further, there is strong evidence implicating different AGT molecular variants as the cause of human essential hypertension and organ damage during aging (reviewed in [42]).

Expression of Mbp is known to decrease and Cck is known to affect the axonogenesis function in the aged animals (Table 3). Our analysis has revealed an increased expression of Mbp and a decreased expression of Cck. MBP is a major constituent of the myelin sheath of oligodendrocytes and has an important role in the pathophysiology of multiple sclerosis [43], which is a chronic inflammatory and neurodegenerative disease of the CNS of unknown cause. Cck is extensively expressed in the brain and a number of diverse changes to hippocampal Cck expression profile have been documented in various models of epilepsy [44]. Cck is also known to have a role in modulating the neuronal network of anxiety and panic disorder that involves other parts of the brain e.g. amygdale, hypothalamus [45]. Such results argue that the GA genes in general are associated with reduction in physiological and immunological efficiency leading to deterioration (senescence) with advancing age in the aged rats.

GASI or general aging genes associated with syndromic learning impairments. Deterioration of mental and physical state is very common with advancing age, which manifest in various syndromes. It is apparent that many syndromes associated with aging are also involved in memory loss and learning impairments. One such syndrome is the Alzheimer's disease (AD), which has been studied extensively. Among the identified GASI genes in this analysis that have been implicated in AD or late-onset Alzheimer disease (LOAD) include ApoE [46,47], Mapt [48], Igf1r (insulin-like growth factor 1 receptor) [49], Clu [50,51], Picalm (phosphatidylinositol binding clathrin assembly protein) [50,51], Cdk5r1 (cyclin-dependent kinase 5, regulatory subunit 1, p35) [52], and Ide (Insulin degrading enzyme) [53]. ApoE, Mapt, Igf1r, Clu, and Picalm were up-regulated, and Cdk5r1 and Ide were down-regulated in the aged animals compared to the young. These GASI genes may also lead to syndromic learning impairments by affecting various key neuronal functions. For example, ApoE, Cdk5r1 and Ide are known to decrease cell viability and *Picalm* and *Igf1r* are known to affect axonogenesis (Table 3).

Specifically, ApoE and Mapt have been annotated as the aging and learning genes in the IPA Knowledge Base (Table S3). ApoE gene is known as the strongest risk factor for age-related cognitive decline during normal ageing [54]. APOE isoforms differentially regulate A $\beta$  (amyloid  $\beta$ -peptide) aggregation and clearance in the brain, and have distinct functions in regulating brain lipid transport, glucose metabolism, neuronal signalling, neuroinflammation, and mitochondrial function (reviewed in [55]). Toxicity of A $\beta$  also depends on *Mapt* (Figure 6). Increase in MAPT levels may represent a very early sign of NFT (neurofibrillary tangle) formation and AD in humans [48]. Down-regulated Igf1r activity has been implicated with prolonged human lifespan [49]. When considering age-related neurodegeneration in AD, signalling through the Igf1r is disturbed in the AD patients' brain, and an increased level of Igf1r has been reported in the degenerating synapses in cerebral cortex within and surrounding  $A\beta$  plaques in people with AD compared to people of the same age without the disease (reviewed in [49]). Through the deregulated activity of Cdk5, Cdk5r1 is involved in the pathology of AD [52], synaptic plasticity, learning, and memory [56]. IDE is involved in the degradation of A $\beta$  and other bioactive peptides e.g. insulin and IGF-1 and IGF-2 in vitro (reviewed in [57]). PICALM plays a critical role in iron homeostasis and cell proliferation [58]. PICALM knockdown can result in reduced APP (amyloid precursor protein) internalization and A $\beta$  generation, while overexpression can cause increased APP internalization and amyloid plaque load [59]. Irregularities in the A $\beta$  clearance pathway are thought to initiate A $\beta$  and tau protein accumulation in specific brain regions and consequent toxic events that lead to synaptic dysfunction and neurodegeneration in AD. This is associated with the progressive destruction of synaptic circuits controlling memory and higher mental function.

Besides the above genes associated with AD, there is a number of GASI genes associated with other age-related disease syndromes and related memory impairment. For example, Cntn2 (Contactin-2), a learning gene, is up-regulated in the aged (Table S3), Hmgb1 (high mobility group box-1) is up-regulated in the aged-impaired (Table 6), and *Tubb2b* is down-regulated in the aged rats (Table 2). *Cntn2* plays a role in the formation of axon connections [60] and autoimmune responses to Cntn2 have been implicated in multiple sclerosis [61]. Studies show that cellular stress, trauma, and inflammatory condition can also result in the up-regulation of *Hmgb1* in the hippocampus in aged rats, which results in reduced cognitive function in a reversal learning version of the MWM test [62, 63]. Further, *Tubb2b* is a major component of microtubules cytoskeletal structures essential for cell motility and function and one of the top ten most down-regulated genes in the AY comparison (Table 2). A spectrum of neurological disorders (e.g. Polymicrogyria) characterized by abnormal neuronal migration, differentiation, organization, axon guidance, and maintenance has recently been associated with various mutations in Tubb2b [64,65]. In summary, a number of genes identified in the aged and agedimpaired animals are associated with a number of syndromes and fall in the category of GASI genes, which may contribute to the memory loss and learning impairments observed in the agedimpaired animals.

**GANSI or general aging related genes associated with non-syndromic learning impairments.** It is apparent that the majority of the differentially expressed genes in the aged or aged-impaired animals are known to facilitate learning and memory formation and are not implicated in any syndromes. They have been annotated as learning or spatial learning genes in the IPA Knowledge Base (Table S3 and S4). The canonical pathways that are most relevant to the GANSI genes functioning in the brain include nNos signaling pathway and glutamate receptor signaling pathway, which were identified most significant in the IU comparison (Table 8). nNos [66,67] and glutamate receptors (reviewed in [68]) play an important role in neurotransmission and are critical to LTP, memory formation and synaptic plasticity.

The genes that deserve particular attention in the GANSI category are the 59 genes identified in the IU comparison following BH correction on the effect sizes (pBH.ES  $\leq 0.05$ ). These genes were differentially expressed in the aged rats with spatial learning impairment as compared to those without spatial learning impairment. *Arc*, a learning gene, is one of the most interesting of these 59 genes and is among the top ten most down-regulated genes in the aged-impaired animals (Table 6). The immediate-early gene *Arc* (aka *Arg3*) (Figure 5C) expression is found vital for spatial memory consolidation and long-term synaptic plasticity in a variety of hippocampal-dependent and hippocampal-independent.

dent tasks, including spatial learning in the MWM (reviewed in [69,70]). Are is known for its tight experience-dependent regulation, dendritic mRNA transport, and local protein expression in activated synapses. For example, blocking Are expression either using Are knockout mice [71] or intra-hippocampal injections of Are antisense oligonucleotides [72] is known to impair or prevent LTP without affecting short-term memory performance.

When we considered the larger list of 787 differentially expressed genes in the IU comparison (BH uncorrected, pES  $\leq 0.05$ ), we also found 48 genes annotated as learning or spatial learning genes in the IPA Knowledge Base (Table S4). Some of the interesting learning genes among theses 48 genes include Camk2a (calcium/calmodulin-dependent protein kinase II alpha), Creb1 (cAMP responsive element binding protein 1), Crem (cAMP responsive element modulator), Egr1 (early growth response 1), Homer 1 (homer homolog 1) (Figure 6), Junb (jun B proto-oncogene) (Figure 6), Psen2 (presenilin 2), Slc11a2 (solute carrier family 11), and Marcks. Particularly, Marcks (pES = 0.004) (Figure 5D) is highly up-regulated in the aged-impaired animals. Timofeeva and colleagues recently reported that local infusions of MARCKS long peptide into the rat hippocampus resulted in a dramatic impairment of both working and reference memory in a dosedependent manner with robust impairment at higher doses [73], most likely through the inhibition of alpha7 nicotinic acetylcholine receptors [74]. Thus, our analysis has identified the two genes, Arc and Marcks, as prime candidates for further investigation for their role in ASLI.

Additional GANSI genes include Bdnf (brain-derived neurotrophic factor), Ntf3 (neurotrophin 3), Igf2, Serpini1 (neuroserpin), Gucy1a3 (guanylate cyclase 1, soluble, alpha 3), Gucy1b3 (guanylate cyclase 1, soluble, beta 3), Avp (arginine vasopressin), Gnaq (guanine nucleotide binding protein), Grp (gastrin releasing peptite), Pthlh (parathyroid hormone-like hormone), Trhr (thyrotropin-releasing hormone receptor), Agm (agrin), L1cam (Cell adhesion molecule L1), and *Ppp2ca* (protein phosphatase 2, catalytic subunit, alpha isozyme). These differentially expressed genes in the AY comparison play a critical role in the increase or decrease of several significant functions (Table 3) in the aged animals. Since a majority (73%) of the aged animals in the AY comparison were also impaired in the spatial learning task, it is not surprising that some of the aging genes may also contribute to the ASLI in these animals. Below we highlight some major functions of these GANSI genes.

For example, the genes *Bdnf*, *Ntf3*, *Igf2*, and *Serpini1* were downregulated in the aged animals and are known to decrease cell viability of CNS cells (Table 3). Neurotrophins such as *Bdnf* and *Ntf3* expression is strongly associated with synaptic function and plasticity. Specifically, *Bdnf* is known as a strong mediator for LTP (long term potentiation) in the hippocampus and play an essential role in memory formation in the adult brain (see review [75]). *Igf2* is a late response gene regulated by the CREB-C/EBP pathway and plays a critical role in memory consolidation and enhancement [76]. Furthermore, injections of recombinant IGF-II into the hippocampus after either training or memory retrieval significantly enhance memory retention and prevent forgetting. Neuroserpin e.g. *Serpini1* expression is involved in regulating the proteolytic balance associated with axonogenesis and synaptogenesis during development and synaptic plasticity in the adult [77,78].

Further, *Gucy1a3* and *Gucy1b3* are involved in the increase of cellular movement function (Table 3). They are soluble guanylate cyclises (sGC) and are part of the nitric oxide (NO)/sGC/cGMP dependent protein kinase (PKG) signaling pathway that plays a key role in memory processing [66,67]. Inhibition of sGC, of PKG or of cGMP-degrading phosphodiesterase has been found to

impair LTP [79]. Both GUCY1A3 and GUCY1B3 were found down-regulated in the aged animals, which may explain the ASLI in these animals.

The products of the genes Avp [80], Gnaq [81], Grp [82], Pthlh [83], and Trhr [84] maintain the quantity and synthesis of IP<sub>3</sub> (inositol 1,4,5-triphosphate) level in the cell (Table 3) and are down-regulated in the aged rats. These genes facilitate IP<sub>3</sub> production in the brain and some through the activation of phospolipase C (PLC) [80,81,82,83,84]. Some of them e.g. AVP [85] and GRPs (see review [82]) are specifically involved in regulating cognition and memory. IP<sub>3</sub> is an important second messenger in the neuron produced from phosphatidylinsitol biphosphate (PIP<sub>2</sub>) and cleaved by PLC. IP<sub>3</sub> binds to IP<sub>3</sub> receptors, which are gated Ca2+ channels that release calcium from the endoplasmic reticulum in to the cytosol [86].  $Ca^{2+}$  in turn controls many different signalling events within neuron, including neurotransmitter release and gene expression in the cell nucleus. At least two Ca<sup>2+</sup>-activated protein kinases e.g. Ca<sup>2+</sup>/calmodulindependent protein kinase (CaMKII) and protein kinase C (PKC) have been implicated in LTP induction. LTP is the underlying cellular molecular mechanism that correlates with learning and memory formation [87]. Thus, down regulation of the genes Avp, Gnaq, Grp, Pthlh, and Trhr can have a negative effect on the Inositol phospholipid-calcium-CamK-protein kinase C transduction pathway through decreased quantity and synthesis of  $IP_3$  in the aged animals and directly or indirectly contribute to age-associated nonsyndromic learning impairments such as ASLI.

A number of genes e.g. Agm, L1cam, Ppp2ca that are downregulated in the aged animals demonstrated spatial learning impairment. Lower expression of these genes is known to decrease axonogenesis (Table 3). They play a critical role in neurite outgrowth, synaptogenesis, and synaptic plasticity. For example, high level of Agm (Figure 6) expression was found in regions of the adult brain that show extensive synaptic plasticity. Recent studies demonstrated a substantial loss of excitatory synapses in the adult transgenic mice brain that lacked agrin expression. Further, they demonstrated inhibition of synaptogenesis by agrin antisense oligonucleotides or agrin siRNA in neuronal cell culture (reviewed in [88]). L1cam promotes the outgrowth of neurites and thereby contributes to formation of neuronal connections, learning, and memory [89,90] via activation of the mitogen-activated protein kinase (MAPK) pathway [91]. Ppp2ca (aka Pp2a) is involved in Ca2+-dependent dephosphorylation of SNAP-25 [92] and SNAP-25 phosphorylation plays an important role in neural plasticity and long-term potentiation in the hippocampus [93].

It is important to note that Fischer 344 strain of rats have a median life-span of 23–31 months in captivity [94,95]. Their normal age-related incidence of neoplasms and degenerative diseases is high, particularly, once the rats pass 24 months of age [94,95]. Also, the effect of aging and ASLI on brain gene expression is evident in the aged (24–26 months old) in comparison to the young (3–6 months old) rats as discussed above. Indeed, it is expected that studies on animals beyond 26 weeks of age may show involvement of additional genes in this phenomenon and the effects observed could be more pronounced at later stages of the rat's life-span.

In conclusion, we report that aged animals display a significant decrease in cell viability, axonogenesis, and inositol phosphate metabolism. They also show a significant increase in the migration of cells and differentiation of cells functions due to the altered gene expression. The regulatory interactions of the differentially expressed genes seems to affect molecular transport, cell to cell signaling and interaction, nervous system development and function, and cell death and survival. The genes that are known to be involved in the above functional changes and/or those present most significant expression changes in the aged or agedimpaired animals could be broadly classified into three major categories such as GA, GASI, and GANSI. The GA genes are mostly involved in inflicting various aging related senescence (e.g. stress, disorders, and inflammation conditions) and generally are not associated with any learning impairment. The GASI genes, on the other hand, are associated with age-related neurological disease syndromes e.g. Alzheimer's disease, which generally affect normal cognitive functioning and may result into syndromic memory impairments. The most important group of genes perhaps is the GANSI genes, most of which show down-regulation in the aged or aged-impaired rats and by themselves usually are not associated with any syndromes. These genes affect various signal transduction pathways and functions in the brain contributing to the disruption of proper learning and memory formation. We suggest that the GANSI genes should form the foundation for future studies in understanding age-associated memory impairments such as ASLI. These GASI and GANSI genes form a set of interesting candidates for future investigations as to how they interact with each other, how they are regulated, and what target genes they may affect in order to elucidate the mechanisms behind aging and age-associated spatial learning impairment.

#### References

- Landfield PW, Thibault O, Mazzanti ML, Porter NM, Kerr DS (1992) Mechanisms of neuronal death in brain aging and Alzheimer's disease: role of endocrine-mediated calcium dyshomeostasis. J Neurobiol 23: 1247–1260.
- Tanzi RE, Bertram L (2001) New frontiers in Alzheimer's disease genetics. Neuron 32: 181–184.
- Mattson MP, Magnus T (2006) Ageing and neuronal vulnerability. Nat Rev Neurosci 7: 278–294.
- Burger C, Lopez MC, Feller JA, Baker HV, Muzyczka N, et al. (2007) Changes in transcription within the CA1 field of the hippocampus are associated with age-related spatial learning impairments. Neurobiol Learn Mem 87: 21–41.
- Peleg S, Sananbenesi F, Zovoilis A, Burkhardt S, Bahari-Javan S, et al. (2010) Altered histone acetylation is associated with age-dependent memory impairment in mice. Science 328: 753–756.
- Glorioso C, Oh S, Douillard GG, Sibille E (2011) Brain molecular aging, promotion of neurological disease and modulation by Sirtuin5 longevity gene polymorphism. Neurobiol Dis 41: 279–290.
- Morris RG, Garrud P, Rawlins JN, O'Keefe J (1982) Place navigation impaired in rats with hippocampal lesions. Nature 297: 681–683.
- Burgess N (2002) The hippocampus, space, and viewpoints in episodic memory. Q.J Exp Psychol A 55: 1057–1080.
- Verbitsky M, Yonan AL, Malleret G, Kandel ER, Gilliam TC, et al. (2004) Altered hippocampal transcript profile accompanies an age-related spatial memory deficit in mice. Learn Mem 11: 253–260.
- Mesulam MM (1999) Neuroplasticity failure in Alzheimer's disease: bridging the gap between plaques and tangles. Neuron 24: 521–529.
- Small SA, Tsai WY, DeLaPaz R, Mayeux R, Stern Y (2002) Imaging hippocampal function across the human life span: is memory decline normal or not? Ann Neurol 51: 290–295.
- Pawlowski TL, Bellush LL, Wright AW, Walker JP, Colvin RA, et al. (2009) Hippocampal gene expression changes during age-related cognitive decline. Brain Res 1256: 101–110.
- Lu T, Pan Y, Kao SY, Li C, Kohane I, et al. (2004) Gene regulation and DNA damage in the ageing human brain. Nature 429: 883–891.
- Glorioso C, Sibille E (2011) Between destiny and disease: Genetics and molecular pathways of human central nervous system aging. Prog Neurobiol 93: 165–181.
- Burger C, Lopez MC, Baker HV, Mandel RJ, Muzyczka N (2008) Genomewide analysis of aging and learning-related genes in the hippocampal dentate gyrus. Neurobiol Learn Mem 89: 379–396.
- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) Introduction to Meta-Analysis: Wiley. 421 p.
- Li C, Wong W (2001) Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol 2: RESEARCH0032.
- Bolstad BM, Collin F, Brettschneider J, Simpson K, Cope L, et al. (2005) Quality assessment of affymetrix GeneChip data. In: Gentleman R, Irizarry RA, Carey VJ, Dudoit S, Huber W, editors. Bioinformatics and computational biology solutions using R and Bioconductor. First ed. New York: Springer. 33–48.

#### **Supporting Information**

 Table S1
 Details of all 2245 significant genes in the AY comparison.

(XLSX)

Table S2Details of all 787 significant genes in the IUcomparison.

(XLSX)

Table S3 List of 86 learning genes found in the AY significant genes. (XLSX)

Table S4 List of 48 learning genes found in the IU significant genes.

(XLSX)

**Checklist S1 Preferred reporting items for systematic reviews and meta-analyses (PRISMA) checklist.** (DOCX)

#### **Author Contributions**

Conceived and designed the experiments: RKU SMS. Performed the experiments: RKU. Analyzed the data: RKU. Contributed reagents/ materials/analysis tools: RKU SMS. Wrote the paper: RKU SMS.

- Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19: 185–193.
- Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy-analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307–315.
- Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8: 118–127.
- Chen C, Grennan K, Badner J, Zhang D, Gershon E, et al. (2011) Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. PLoS One 6: e17238.
- Sirbu A, Ruskin HJ, Crane M (2010) Cross-platform microarray data normalisation for regulatory network inference. PLoS One 5: e13822.
- Ramasamy A, Mondry A, Holmes CC, Altman DG (2008) Key issues in conducting a meta-analysis of gene expression microarray datasets. PLoS Med 5: e184.
- Moreau Y, Aerts S, De Moor B, De Strooper B, Dabrowski M (2003) Comparison and meta-analysis of microarray data: from the bench to the computer desk. Trends Genet 19: 570–577.
- Cochran W (1937) Problems arising in the analysis of a series of similar experiments. J R Stat Soc: 102–118.
- Fleiss JL (1993) The statistical basis of meta-analysis. Stat methods Med Res: 121–145.
- Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM (2002) Metaanalysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res 62: 4427–4433.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B 57: 299–314.
- Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article3.
- Blalock EM, Chen KC, Sharrow K, Herman JP, Porter NM, et al. (2003) Gene microarrays in hippocampal aging: statistical profiling identifies novel processes correlated with cognitive impairment. J Neurosci 23: 3807–3819.
- Rowe WB, Blalock EM, Chen KC, Kadish I, Wang D, et al. (2007) Hippocampal expression analyses reveal selective association of immediateearly, neuroenergetic, and myelinogenic pathways with cognitive impairment in aged rats. J Neurosci 27: 3098–3110.
- Kadish I, Thibault O, Blalock EM, Chen KC, Gant JC, et al. (2009) Hippocampal and cognitive aging across the lifespan: a bioenergetic shift precedes and increased cholesterol trafficking parallels memory impairment. J Neurosci 29: 1805–1816.
- Goldstein DR, Guerra R (2010) A brief introduction to meta-analysis, genetics and genomics. In: Goldstein DR, Guerra R, editors. Meta-analysis and combining information in genetics and genomics: CRC Press. 3–20.
- Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, et al. (2010) Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet 11: 733–739.

- 36. Donnelly N, Gorman AM, Gupta S, Samali A (2013) The eIF2alpha kinases: their structures and functions. Cell Mol Life Sci.
- 37. Ishii N, Takahashi T, Soroosh P, Sugamura K (2010) OX40-OX40 ligand interaction in T-cell-mediated immunity and immunopathology. Adv Immunol  $105 \cdot 63 - 98$
- 38. Cazander G, Jukema GN, Nibbering PH (2012) Complement activation and inhibition in wound healing. Clin Dev Immunol 2012: 534291.
- 39. Maier M, Peng Y, Jiang L, Seabrook TJ, Carroll MC, et al. (2008) Complement C3 deficiency leads to accelerated amyloid beta plaque deposition and neurodegeneration and modulation of the microglia/macrophage phenotype in amyloid precursor protein transgenic mice. J Neurosci 28: 6333-6341.
- 40. Borghese F, Clanchy FI (2011) CD74: an emerging opportunity as a therapeutic target in cancer and autoimmune disease. Expert Opin Ther Targets 15: 237-251.
- 41. Mihelic M, Dobersek A, Guncar G, Turk D (2008) Inhibitory fragment from the p41 form of invariant chain can regulate activity of cysteine cathepsins in antigen presentation. J Biol Chem 283: 14453-14460.
- 42. Arnold AC, Gallagher PE, Diz DI (2013) Brain renin-angiotensin system in the nexus of hypertension and aging. Hypertens Res 36: 5-13.
- 43. Moscarello MA, Mastronardi FG, Wood DD (2007) The role of citrullinated proteins suggests a novel mechanism in the pathogenesis of multiple sclerosis. Neurochem Res 32: 251-256.
- 44. Wyeth MS, Zhang N, Houser CR (2012) Increased cholecystokinin labeling in the hippocampus of a mouse model of epilepsy maps to spines and glutamatergic terminals. Neuroscience 202: 371-383.
- 45. Zwanzger P, Domschke K, Bradwejn J (2012) Neuronal network of panic disorder: the role of the neuropeptide cholecystokinin. Depress Anxiety 29: 762-774.
- 46. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, et al. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261: 921–923.
- 47. Bu G (2009) Apolipoprotein E and its receptors in Alzheimer's disease: pathways, pathogenesis and therapy. Nat Rev Neurosci 10: 333-344.
- 48. Maeda S, Sahara N, Saito Y, Murayama S, Ikai A, et al. (2006) Increased levels of granular tau oligomers: an early sign of brain aging and Alzheimer's disease. Neurosci Res 54: 197-201.
- 49. O'Neill C, Kiely AP, Coakley MF, Manning S, Long-Smith CM (2012) Insulin and IGF-1 signalling: longevity, protein homoeostasis and Alzheimer's disease. Biochem Soc Trans 40: 721-727
- 50. Ferrari R, Moreno JH, Minhajuddin AT, O'Bryant SE, Reisch JS, et al. (2012) Implication of common and disease specific variants in CLU, CR1, and PICALM. Neurobiol Aging 33: 1846 e1847-1818.
- 51. Chen LH, Kao PY, Fan YH, Ho DT, Chan CS, et al. (2012) Polymorphisms of CR1, CLU and PICALM confer susceptibility of Alzheimer's disease in a southern Chinese population. Neurobiol Aging 33: 210 e211-217.
- 52. Shukla V, Skuntz S, Pant HC (2012) Deregulated Cdk5 activity is involved in inducing Alzheimer's disease. Arch Med Res 43: 655-662.
- 53. Miller BC, Eckman EA, Sambamurti K, Dobbs N, Chow KM, et al. (2003) Amyloid-beta peptide levels in brain are inversely correlated with insulysin activity levels in vivo. Proc Natl Acad Sci U S A 100: 6221-6226.
- 54. Alzheimer's A (2012) 2012 Alzheimer's disease facts and figures. 8, 131-168 (2012). Alzheimers Dement 8: 131-168.
- 55. Liu CC, Kanekiyo T, Xu H, Bu G (2013) Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol.
- 56. Angelo M, Plattner F, Giese KP (2006) Cyclin-dependent kinase 5 in synaptic plasticity, learning and memory. J Neurochem 99: 353-370.
- 57. Nalivaeva NN, Beckett C, Belyaev ND, Turner AJ (2012) Are amyloiddegrading enzymes viable therapeutic targets in Alzheimer's disease? J Neurochem 120 Suppl 1: 167-185.
- 58. Scotland PB, Heath JL, Conway AE, Porter NB, Armstrong MB, et al. (2012) The PICALM protein plays a key role in iron homeostasis and cell proliferation. PLoS One 7: e44252.
- 59. Xiao Q, Gil SC, Yan P, Wang Y, Han S, et al. (2012) Role of phosphatidylinositol clathrin assembly lymphoid-myeloid leukemia (PICALM) in intracellular amyloid precursor protein (APP) processing and amyloid plaque pathogenesis. J Biol Chem 287: 21279-21289.
- 60. Lin JF, Pan HC, Ma LP, Shen YQ, Schachner M (2012) The cell neural adhesion molecule contactin-2 (TAG-1) is beneficial for functional recovery after spinal cord injury in adult zebrafish. PLoS One 7: e52376.
- 61. Derfuss T, Parikh K, Velhin S, Braun M, Mathey E, et al. (2009) Contactin-2/ TAG-1-directed autoimmunity is identified in multiple sclerosis patients and mediates gray matter pathology in animals. Proc Natl Acad Sci U S A 106: 8302-8307.
- 62. He HJ, Wang Y, Le Y, Duan KM, Yan XB, et al. (2012) Surgery upregulates high mobility group box-1 and disrupts the blood-brain barrier causing cognitive dysfunction in aged rats. CNS Neurosci Ther 18: 994-1002.
- 63. Klune JR, Dhupar R, Cardinal J, Billiar TR, Tsung A (2008) HMGB1: endogenous danger signaling. Mol Med 14: 476-484.
- 64. Romaniello R, Tonelli A, Arrigoni F, Baschirotto C, Triulzi F, et al. (2012) A novel mutation in the beta-tubulin gene TUBB2B associated with complex malformation of cortical development and deficits in axonal guidance. Dev Med Child Neurol 54: 765-769.

- 65. Cushion TD, Dobyns WB, Mullins JG, Stoodley N, Chung SK, et al. (2013) Overlapping cortical malformations and mutations in TUBB2B and TUBA1A. Brain 136: 536-548.
- 66. Bartus K, Pigott B, Garthwaite J (2013) Cellular targets of nitric oxide in the hippocampus. PLoS One 8: e57292.
- 67. Shen F, Li YJ, Shou XJ, Cui CL (2012) Role of the NO/sGC/PKG signaling pathway of hippocampal CA1 in morphine-induced reward memory. Neurobiol Learn Mem 98: 130-138.
- 68. Menard C, Quirion R (2012) Group 1 metabotropic glutamate receptor function and its regulation of learning and memory in the aging brain. Front Pharmacol 3: 182
- 69. Bramham CR, Alme MN, Bittins M, Kuipers SD, Nair RR, et al. (2010) The Arc of synaptic memory. Exp Brain Res 200: 125–140. Shepherd JD, Bear MF (2011) New views of Arc, a master regulator of synaptic
- 70 plasticity. Nat Neurosci 14: 279-284.
- 71. Plath N, Ohana O, Dammermann B, Errington ML, Schmitz D, et al. (2006) Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. Neuron 52: 437-444.
- 72. Guzowski JF, Lyford GL, Stevenson GD, Houston FP, McGaugh JL, et al. (2000) Inhibition of activity-dependent arc protein expression in the rat hippocampus impairs the maintenance of long-term potentiation and the consolidation of long-term memory. J Neurosci 20: 3993-4001.
- 73. Timofeeva OA, Eddins D, Yakel JL, Blackshear PJ, Levin ED (2010) Hippocampal infusions of MARCKS peptides impair memory of rats on the radial-arm maze. Brain Res 1308: 147-152.
- 74. Gay EA, Klein RC, Melton MA, Blackshear PJ, Yakel JL (2008) Inhibition of native and recombinant nicotinic acetylcholine receptors by the myristoylated alanine-rich C kinase substrate peptide. J Pharmacol Exp Ther 327: 884-890.
- 75. Park H, Poo MM (2013) Neurotrophin regulation of neural circuit development and function. Nat Rev Neurosci 14: 7-23.
- 76. Chen DY, Stern SA, Garcia-Osta A, Saunier-Rebori B, Pollonini G, et al. (2011) A critical role for IGF-II in memory consolidation and enhancement. Nature 469: 491-497.
- 77. Osterwalder T, Contartese J, Stoeckli ET, Kuhn TB, Sonderegger P (1996) Neuroserpin, an axonally secreted serine protease inhibitor. EMBO J 15: 2944-2953
- 78. Lee TW, Coates LC, Birch NP (2008) Neuroserpin regulates N-cadherinmediated cell adhesion independently of its activity as an inhibitor of tissue plasminogen activator. J Neurosci Res 86: 1243-1253.
- Monfort P, Munoz MD, Kosenko E, Llansola M, Sanchez-Perez A, et al. (2004) Sequential activation of soluble guanylate cyclase, protein kinase G and cGMPdegrading phosphodiesterase is necessary for proper induction of long-term potentiation in CA1 of hippocampus. Alterations in hyperammonemia. . Neurochem Int 45: 895–901.
- 80. Poulin P, Pittman QJ (1993) Arginine vasopressin-induced sensitization in brain: facilitated inositol phosphate production without changes in receptor number. J Neuroendocrinol 5: 23–31.
- 81. Montmayeur JP, Barr TP, Kam SA, Packer SJ, Strichartz GR (2011) ET-1 induced Elevation of intracellular calcium in clonal neuronal and embryonic kidney cells involves endogenous endothelin-A receptors linked to phospholipase C through Galpha(q/11). Pharmacol Res 64: 258-267.
- 82. Roesler R, Schwartsmann G (2012) Gastrin-releasing peptide receptors in the central nervous system: role in brain function and as a drug target. Front Endocrinol (Lausanne) 3: 159.
- Smogorzewski M, Islam A (1995) Parathyroid hormone stimulates the 83 generation of inositol 1,4,5-triphosphate in brain synaptosomes. Am J Kidney Dis 26: 814-817
- 84. Ramsdell JS, Tashjian AH, Jr. (1986) Thyrotropin-releasing hormone (TRH) elevation of inositol trisphosphate and cytosolic free calcium is dependent on receptor number. Evidence for multiple rapid interactions between TRH and its receptor. J Biol Chem 261: 5301-5306.
- 85. Ebstein RP, Knafo A, Mankuta D, Chew SH, Lai PS (2012) The contributions of oxytocin and vasopressin pathway genes to human behavior. Horm Behav 61: 359-379
- 86. Finch EA, Augustine GJ (1998) Local calcium signalling by inositol-1,4,5trisphosphate in Purkinje cell dendrites. Nature 396: 753-756.
- 87. Foster TC (2012) Dissecting the age-related decline on spatial learning and memory tasks in rodent models: N-methyl-D-aspartate receptors and voltagedependent Ca2+ channels in senescent synaptic plasticity. Prog Neurobiol 96: 283-303
- 88. Daniels MP (2012) The role of agrin in synaptic development, plasticity and signaling in the central nervous system. Neurochem Int 61: 848-853.
- Kenwrick S, Watkins A, De Angelis E (2000) Neural cell recognition molecule 89. L1: relating biological complexity to human disease mutations. Hum Mol Genet 9:879-886
- 90. Maness PF, Schachner M (2007) Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration. Nat Neurosci 10: 19-26.
- 91. Poplawski GH, Tranziska AK, Leshchyns'ka I, Meier ID, Streichert T, et al. (2012) L1CAM increases MAP2 expression via the MAPK pathway to promote neurite outgrowth. Mol Cell Neurosci 50: 169-178.
- 92. Iida Y, Yamamori S, Itakura M, Miyaoka H, Takahashi M (2013) Protein phosphatase 2A dephosphorylates SNAP-25 through two distinct mechanisms in mouse brain synaptosomes. Neurosci Res.

- Genoud S, Pralong W, Riederer BM, Eder L, Catsicas S, et al. (1999) Activitydependent phosphorylation of SNAP-25 in hippocampal organotypic cultures. J Neurochem 72: 1699–1706.
- Sass B, Rabstein LS, Madison R, Nims RM, Peters RL, et al. (1975) Incidence of spontaneous neoplasms in F344 rats throughout the natural life-span. J Natl Cancer Inst 54: 1449–1456.
- Coleman GL, Barthold W, Osbaldiston GW, Foster SJ, Jonas AM (1977) Pathological changes during aging in barrier-reared Fischer 344 male rats. J Gerontol 32: 258–278.