Article

Characterizing Gene and Protein Crosstalks in Subjects at Risk of Developing Alzheimer's Disease: A New Computational Approach

Kanchana Padmanabhan [1,2], Kelly Nudelman [3], Steve Harenberg [1], Gonzalo Bello [1], Dongwha Sohn [1,2], Katie Shpanskaya [4], Priyanka Tiwari Dikshit [1], Pallavi S. Yerramsetty [1], Rudolph E. Tanzi [5] (iD), Andrew J. Saykin [3], Jeffrey R. Petrella [6], P. Murali Doraiswamy [7] and Nagiza F. Samatova [1,2,*] for the Alzheimer's Disease Neuroimaging Initiative [†]

1   Department of Computer Science, North Carolina State University, Raleigh, NC 27695, USA; kanchanapadmanabhan@gmail.com (K.P.); sdharenb@ncsu.edu (S.H.); gabellol@ncsu.edu (G.B.); dsohn@ncsu.edu (D.S.); ptdikshi@ncsu.edu (P.T.D.); pallaviyerramsetty98@gmail.com (P.S.Y.)
2   Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
3   Indiana Alzheimer Disease Center and the Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN 46202, USA; kholohan@iupui.edu (K.N.); asaykin@iupui.edu (A.J.S.)
4   Department of Radiology, Stanford University School of Medicine, Stanford, CA 94025, USA; kss@stanford.edu
5   Genetics and Aging Research Unit and Department of Neurology, Massachusetts General Hospital and Harvard Medical School Stanford University School of Medicine, Stanford, CA 02129, USA; tanzi@helix.mgh.harvard.edu
6   Department of Radiology, Duke University Medical Center, Durham, NC 27710, USA; jeffrey.petrella@duke.edu
7   Neurocognitive Disorders Program, Department of Psychiatry and the Duke Institute for Brain Sciences, Duke University Health System, Durham, NC 27710, USA; murali.doraiswamy@duke.edu
*   Correspondence: samatova@csc.ncsu.edu; Tel.: +1-919-513-7575
†   Data used in preparation of this article were obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Abstract: Alzheimer's disease (AD) is a major public health threat; however, despite decades of research, the disease mechanisms are not completely understood, and there is a significant dearth of predictive biomarkers. The availability of systems biology approaches has opened new avenues for understanding disease mechanisms at a pathway level. However, to the best of our knowledge, no prior study has characterized the nature of pathway crosstalks in AD, or examined their utility as biomarkers for diagnosis or prognosis. In this paper, we build the first computational crosstalk model of AD incorporating genetics, antecedent knowledge, and biomarkers from a national study to create a generic pathway crosstalk reference map and to characterize the nature of genetic and protein pathway crosstalks in mild cognitive impairment (MCI) subjects. We perform initial studies of the utility of incorporating these crosstalks as biomarkers for assessing the risk of MCI progression to AD dementia. Our analysis identified Single Nucleotide Polymorphism-enriched pathways representing six of the seven Kyoto Encyclopedia of Genes and Genomes pathway categories. Integrating pathway crosstalks as a predictor improved the accuracy by 11.7% compared to standard clinical parameters and apolipoprotein E ε4 status alone. Our findings highlight the importance of moving beyond discrete biomarkers to studying interactions among complex biological pathways.

1. Introduction

It is common knowledge that the prognostics of diseases such as Alzheimer's disease (AD) is of national importance. AD alone affects about 10% of the population over 65 years old [1,2], and is among the leading causes of death in patients over 75 years of age in the U.S. [3]. There is evidence suggesting that the progression to AD dementia begins years before it is clinically determined and is preceded by a phase of mild cognitive impairment (MCI), during which AD-related treatments are likely to be more effective. Thus, it is important to discover the mechanisms underlying risk of AD and to develop accurate biomarkers that reflect the complexity of the disease at an individual level. Although a number of biomarkers are currently being evaluated for use to predict AD or study disease progression (e.g., tau, *p*-tau181P, *β*-amyloid1-42, apolipoprotein E *ε*4 (*APOE ε*4), and microRNAs) [4–7], none of these markers are yet fully validated or approved for predicting the risk of AD. Indeed, AD is no longer seen as a disease of single discrete lesions, but as a perturbation of altered cortical networks by pathological processes in interlinked pathways. Hence, the application of systems biology methods to the discovery and characterization of novel biomarkers [8–20] has taken on greater promise and urgency.

The cellular mechanisms underlying many neurological disorders are complex, with crosstalks between multiple molecular pathways likely contributing to disease initiation and progression. In living organisms, pathways are said to crosstalk if they are linked together to perform biological functions as a system. Crosstalks can also be defined as interactions between signal transduction pathways, and usually take the form of protein or transmembrane interactions. A number of potential crosstalks have been noted in vitro in AD, such as those between amyloid and tau pathways, oxidative phosphorylation, the p53 signaling pathway, and apoptosis [21–23]. Another example is the reported crosstalk among MAPK, insulin, and calcium signaling pathways [24]. There is also evidence of crosstalk among pathways involved in the regulation of glycolysis metabolism, pathways involved in the regulation of the actin cytoskeleton, and apoptosis [24]. The latter crosstalk is also associated with other neurodegenerative disorders, such as Huntington disease and amyotrophic lateral sclerosis [24]. Furthermore, the cellular signaling pathways in AD have been reported, such as *Wnt* signaling, 5′ adenosine monophosphate-activated protein kinase, mammalian target of rapamycin, Sirtuin 1, and peroxisome proliferator-activated receptor gamma co-activator 1-*α*, and possible crosstalk between these pathways has been discussed [25]. For a review of multiple interacting pathways in neurodegenerative disease, see [26]. In clinical AD research studies of diagnosis or prognosis, biomarkers are typically treated as discrete entities, in part because biological pathway crosstalks between genes or proteins have not yet been fully characterized at a systems biology level in AD.

From the computational methodology standpoint, the study of pathway crosstalks is still in its infancy. Existing methods predict crosstalks between known metabolic pathways using chemical protein interaction networks [24,27–29]. However, these computational methods do not take advantage of the different chemical evidence available, such as direct binding, the biochemical evidence, such as phosphorylation, and the functional evidence, such as transcriptional regulation. Moreover, the discovery, characterization, and utilization of pathway crosstalks as biomarkers for disease prognosis has not been investigated.

Here, we use clinical, cognitive, and genetic data from a national cohort study, the Alzheimer's Disease Neuroimaging Initiative (ADNI-1), along with a systematic computational methodology to discover and characterize biological pathway crosstalks in subjects with MCI. We further examine the utility of these novel biomarkers to discriminate stable MCI from those who progress to AD dementia. The first part of the methodology (Figure 1), focuses on utilizing several existing evidence, such as chemical interaction, genetic interaction, domain interaction, and transcription factors, to identify

potential pathway crosstalks. In the second part (Figure 2), Single Nucleotide Polymorphisms (SNPs) are used to find patient-specific pathway crosstalks as biomarkers. In the third part, we build and test initial prognostic models that use pathway crosstalks as biomarkers to predict patient progression from MCI progression to AD dementia (see Results). To the best of our knowledge, this is the first such systematic characterization of biological pathway crosstalk biomarkers associated with the risk of AD.
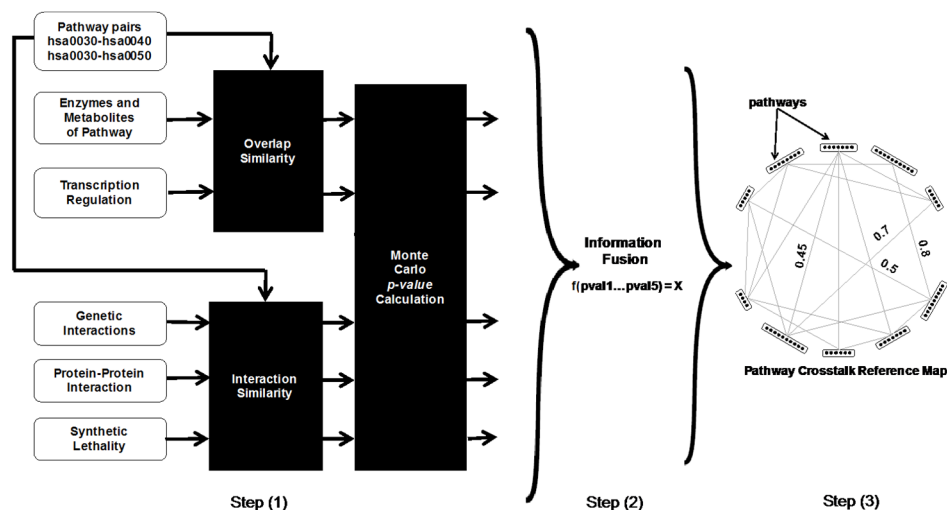

Figure 1. Identification of potential pathway crosstalks. The methodology has three steps: (**1**) quantifying crosstalk likelihood using multiple individual evidence to score each pathway pair, (**2**) obtaining a combined score using information fusion, and (**3**) building the crosstalk reference map.
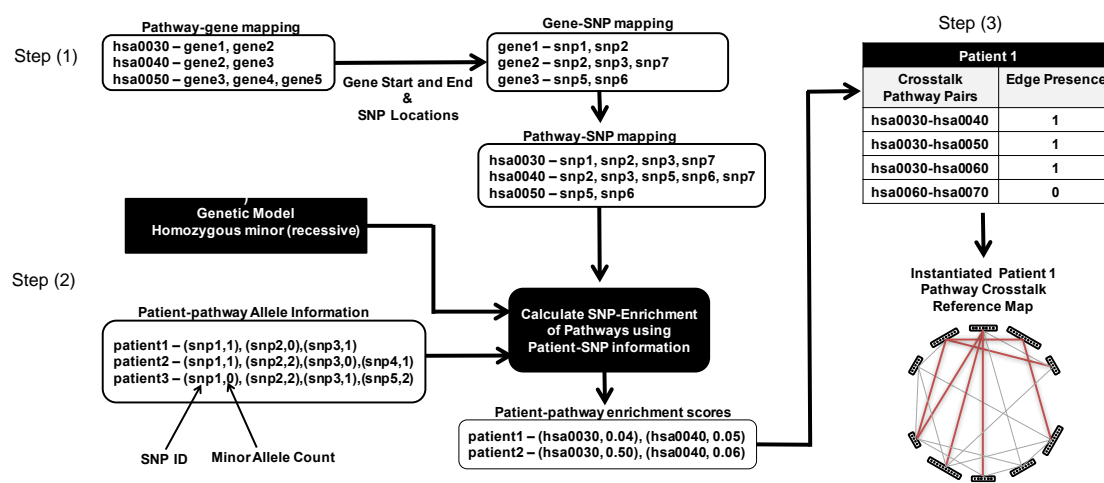

Figure 2. Identification of patient-specific pathway crosstalks. The methodology has three steps: (**1**) mapping the Single Nucleotide Polymorphisms (SNPs) to genes and in turn to pathways using the SNP and gene location information, (**2**) choosing a genetic model and calculating a patient-specific SNP enrichment score for each pathway using the patient's allele information, and (**3**) overlaying the pathway enrichment scores on the reference crosstalk map to build patient-specific pathway crosstalk maps.

2. Materials and Methods

Our methodology consists of the following steps: (A) identifying potential pathway crosstalks by using existing gene and protein data (Figure 1), (B) identifying patient-specific pathway crosstalks via SNP information (Figure 2), and (C) identifying significant pathway crosstalks as biomarkers for MCI progression to AD dementia progression prediction.

2.1. Identification of Potential Pathway Crosstalks

We quantify how likely it is that a pair of pathways will crosstalk based on biological datasets that provide evidence for possible crosstalks (including chemical interaction, genetic interaction, and transcription factors). To have a more robust pathway crosstalk map, we incorporate a wide array of evidence. The scores from each of these evidence are then combined to build one generic pathway crosstalk reference map analogous to the "Kyoto Encyclopedia of Genes and Genomes" (KEGG) pathway reference map.

The likelihood of a pathway pair crosstalking can be scored by utilizing one of two different methods. The first method is based on the presence of common elements, such as kinases and enzymes. The second method is based on the presence of interacting elements, such as chemically interacting proteins. In the following sections, we will discuss the different evidence used and their corresponding scoring methods.

2.1.1. Scoring Pathway Crosstalks Based on Common Elements

The pathway pairs were scored for how likely they are to crosstalk based on common elements from each of the following evidence:

- Shared enzymes and metabolites: The number of enzymes and metabolites shared by a pair of pathways is utilized as one of the evidence to identify potential pathway crosstalks. This is reasonable because a variation in the concentration of common enzymes or metabolites will affect both pathways.
- Phosphorylation: Phosphorylation, performed by protein kinases, is the addition of a phosphate group to a protein, which results in a change of the protein's function. Co-phosphorylated proteins in different pathways suggest potential pathway crosstalks.
- Transcriptional regulation: Genes with common transcription factors are likely coexpressed. Coexpressed genes in different pathways provide an avenue for the pathways to crosstalk. For each pathway pair, we find the group of transcription factors that have coexpressed genes in both pathways.

For each pair of pathways, P_i and P_j, we define the scoring function as Equation (1):

$$Overlap_{score}(P_i, P_j) = \frac{|Y(P_i) \cap Y(P_j)|}{min(|Y(P_i)|, |Y(P_j)|)}, \tag{1}$$

where $Y(P_i)$ is the set of proteins (enzymes, metabolites, transcription factors, kinases) associated with pathway P_i.

2.1.2. Scoring Pathway Crosstalks Based on Interacting Elements

The pathway pairs were scored for how likely they are to crosstalk based on interacting elements from each of the following evidence:

- Chemical interactions: Protein interactions have previously been used to identify pathway crosstalks [24,30]. Chemical interaction between proteins belonging to different pathways provides a mechanism for pathways to crosstalk.
- Genetic interactions: The use of genetic interactions for identifying pathway crosstalks stems from the concept of "between-pathway" interactions. This essentially states that if there is a genetic interaction between pathways, one pathway covers for the defects in the other pathway.
- Protein domain: Protein function is closely related to fundamental units of protein structure called "domains". In the domain interaction network, a pair of proteins has an edge if they are associated with the same set of protein domains. These edges are taken into consideration to assess for potential pathway crosstalks because of the common domains.

- Synthetically lethal gene pairs: Gene pairs whose simultaneous low- or non-expression can cause the organism to die are called synthetically lethal pairs [31,32]. The presence of synthetically lethal pairs of genes across two pathways is a possible sign of pathway crosstalks.

For each pair of pathways, P_i and P_j, we define the scoring function as Equation (2):

$$Interaction_{score}\left(P_i, P_j\right) = \frac{N_{inter}\left(P_i, P_j\right)}{|Y(P_i)| * |Y(P_j)|}, \tag{2}$$

where $N_{inter}\left(P_i, P_j\right)$ is the number of interactions (genetic, chemical, domain, synthetically lethal) that exist among the proteins associated with pathway P_i and the proteins associated with pathway P_j.

2.1.3. Significance Estimation of Pathway Crosstalk Scores

Estimating *p*-values using Monte Carlo methods [33] is a robust technique for statistical significance assessments. This technique was utilized to assess the significance of the scores obtained for the pathway crosstalks using different evidence, as follows:

1. For each pair of pathways, a score for how likely they are to crosstalk is calculated based on each evidence.
2. Each pathway is randomized by replacing all proteins in that pathway with randomly selected proteins from the set of all proteins in the organism. This pathway randomization step is repeated $W = 1000$ times, i.e., we obtain W sets of pathways with randomized proteins.
3. The evidence-specific scores for each pathway pair are recalculated W times using each set of pathways with randomized proteins.
4. An evidence-specific *p*-value is estimated for each pathway pair as R/W, where R is the number of randomized versions of that pathway pair that produce an evidence-specific score greater than or equal to the score obtained for the original pathway pair.

2.1.4. Combining the Scores for Each Pathway Crosstalk

For each pathway pair, we combine the evidence-specific *p*-values obtained using Monte Carlo methods. This gives a combined estimation for crosstalk likelihood between the pathway pair. To combine the *p*-values, we use the QFAST information fusion methodology proposed by Bailey and Gribskov [34], which is based on a theorem by Feller [35]. The QFAST methodology uses the product of the individual *p*-values as a test statistic to calculate the combined *p*-value; using the product of *p*-values as a test statistic has been shown to be a desirable method for information fusion [34]. One issue to consider is that some pathway pairs may not be scored by some of the evidence due to missing data. For those cases, we assign a *p*-value of 1 to denote that the particular evidence offers no information about those pathways crosstalking. The QFAST formula to calculate the combined *p*-value is Equation (3):

$$\left(\prod_{i=1}^{n} p_i\right) \sum_{i=0}^{n-1} \frac{-\ln(\prod_{i=1}^{n} p_i)}{i!}, \tag{3}$$

where P_i is the *p*-value obtained for evidence i, and n is the number of evidence.

A generic pathway crosstalk reference map is then built as a network, where the nodes represent pathways and the edges represent a statistically significant combined *p*-value for crosstalk likelihood between a pathway pair (at a significance level of $\alpha = 0.01$).

2.2. Identification of Patient-Specific Pathway Crosstalks

To determine which of the pathway crosstalks in the generic reference map may be utilized as a biomarker for MCI progression to AD dementia progression, we identify patient-specific pathway crosstalks. For this purpose, we make use of SNP data. SNPs are variations in the deoxyribonucleic

acid (DNA) sequence at particular locations, which can influence phenotypes such as proneness to disease or reaction to drugs. Initiatives such as the ADNI collect patient-specific SNP information. We utilize this information to identify patient-specific pathway crosstalks via the following four steps (Figure 2):

1. Obtain a mapping of SNPs to pathways using genetic information.
2. Identify the list of SNPs that are present in a patient.
3. Use the mapping obtained in Step 1 and the patient-specific SNP list in Step 2 to obtain the pathways that are "SNP-enriched" in the patient.
4. Use the "SNP-enriched" pathways from Step 3 to obtain patient-specific pathway crosstalks.

2.2.1. Obtain a Mapping of SNPs to Pathways

Every SNP is assigned a chromosome number and a location on the genome, which can be used to map SNPs to genes, and, in turn, SNPs to pathways. Starting with a list of all genes that map to at least one pathway, we assign an SNP to a gene if it is present within 10 kilo base pairs (kbp) distance upstream or downstream of that gene. This method has been previously used by Silver et al. [36,37]. Note that since SNPs are mapped to all genes within a range of 10 kbp, the same SNP may be mapped to more than one gene. The set of SNPs assigned to a pathway is the union of all SNPs assigned to the genes of that pathway.

2.2.2. Identify Patient-Specific SNPs That Are Present

For each patient, we identify a list of SNPs that are present based on the homozygous minor (recessive) genetic model. This genetic model requires a minor allele count of 2 for an SNP to be considered present, i.e., the minor allele is inherited from both parents.

2.2.3. Identify Patient-Specific SNP-Enriched Pathways

Given the set of SNPs assigned to a pathway, $\text{SNP}_{pathway}$, the set of SNPs that are present in a patient, $\text{SNP}_{patient}$, and the set of SNPs of interest, $\text{SNP}_{interest}$, we define an enrichment score for this pathway and patient as Equation (4):

$$Enrichment(patient, pathway) = \frac{\left| \text{SNP}_{patient} \cap \text{SNP}_{pathway} \cap \text{SNP}_{interest} \right|}{\left| \text{SNP}_{interest} \right|}, \tag{4}$$

where $\text{SNP}_{interest}$ is the set of all SNPs found on the human genome or a set of relevant SNPs from the scientific literature.

A *p*-value for the enrichment score is calculated using Monte Carlo methods, as discussed previously. The "SNP-enriched" pathways for each patient are then defined as the pathways with a statistically significant *p*-value for that patient (at a significance level of $\alpha = 0.05$).

2.2.4. Identify Patient-Specific Pathway Crosstalks

Given the SNP-enriched pathways for each patient, we build patient-specific pathway crosstalk maps from the generic pathway crosstalk reference map, analogous to building organism-specific pathway maps from the KEGG pathway reference map. A pathway crosstalk, i.e., an edge in the patient-specific reference map, is present if both pathways are SNP-enriched for that patient.

2.3. Identification of Biased Pathway Crosstalk

The pathways and patient-specific pathway crosstalks that are biased towards MCI progressive patients or MCI non-progressive patients (at a significance level of $\alpha = 0.01$) are incorporated as features into the model to predict MCI progression to AD dementia progression. The bias of an active pathway crosstalk towards MCI progressive patients is quantified using the hypergeometric test (Equation (5)):

$$\phi(n, x, v, w) = \sum_{i=w}^{x} \frac{\binom{x}{i}\binom{n-x}{v-i}}{\binom{n}{v}},$$ (5)

where

- Population: n is the total number of patients.
- Success in population: x is the total number of MCI progressive patients and y is the number of MCI non-progressive patients.
- Sample: v is the total number of patients (both MCI progressive and MCI non-progressive patients) a pathway crosstalk is enriched in.
- Success in sample: w is the number of MCI progressive patients and z is the number of MCI non-progressive patients the pathway crosstalk is enriched in.

Similarly, the bias of an active pathway crosstalk towards MCI non-progressive patients can be calculated via $\phi(n, y, v, z)$.

2.4. Datasets

In this study, we utilize cellular subsystems that model biological pathways. Henceforth, we will refer to a cellular subsystem as a pathway. To create a potential pathway crosstalk reference map, we used cellular pathway data from the KEGG database [38–40]. We obtained evidence for human chemical interaction, genetic interaction, and synthetic lethal gene pairs from BioGRID [41], domain interaction from GeneMania [42], transcription factors from the FANTOM database [43,44], and protein phosphorylation [45]. We obtained SNPs associated with genes that were manually curated to be associated with AD from the Comparative Toxicogenomics Database [46], and we obtained a compilation of genes from the literature that have been identified as likely risk factors of AD from SNPedia [47]. This information was utilized as our biologically meaningful knowledge priors. Some of the genes associated with Alzheimer's that were used in this study can be found in Table 1.

Table 1. Some of the genes associated with Alzheimer's disease (AD) that were used in this study.

Gene	Evidence
APP amyloid beta (A4) precursor protein	Mutations in this gene have been implicated in autosomal dominant AD and cerebroarterial amyloidosis (NCBI Entrez Gene)
IL-1β	Four new genetic studies underscore the relevance of IL-1 to Alzheimer's pathogenesis, showing that homozygosity of a specific polymorphism in the IL-1α gene at least triples Alzheimer's risk, especially for an earlier age of onset and in combination with homozygosity for another polymorphism in the IL-1β gene [48]
SOD2	A polymorphism in *SOD2* is associated with development of AD [49]
NOS3	*NOS3* may be a new genetic risk factor of late onset AD [50]

The data used in the preparation of this manuscript were obtained from the ADNI [51] database. The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of the ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

For our predictive study, we utilized the dataset from an earlier study by Shaffer et al. [52] based on ADNI-1. That particular study identified 97 MCI patients and predicted progression to AD dementia based on their clinical parameters, MRI results, PET scans, cerebrospinal fluid (CSF)

markers (tau, *p*-tau181P, and *β*-amyloid1-42), the *APOE ε*4 genotype, and results from at least one follow-up clinical examination. Out of the 97 patients from the earlier study, only 91 patients have corresponding SNP data in the ADNI database. Hence, for the current study, we only utilized these 91 patients. However, this reduction in the number of patients did not considerably affect the ratio of MCI progressive patients to MCI non-progressive patients. The original study had 43 MCI progressive patients and 54 MCI non-progressive patients, and the reduced dataset has 41 MCI progressive patients and 50 MCI non-progressive patients. Thus, there is still sufficient representation of the two classes of patients.

3. Results/Discussion

3.1. Sample Characteristics

The mean age for all 91 MCI patients was 74.96 ± 7.32 years (mean ± standard deviation). The male-to-female ratio was 2.37, and 96.7% of subjects were white. A total of 36.26% of subjects had a family history of AD, and 54.94% had a positive finding for the *APOE ε*4 genotype. The mean follow-up duration for all of the subjects was 31.6 ± 10.6 months. Of these, 41 progressed to AD during follow-up (MCI progressive patients) and 50 did not (MCI non-progressive patients), with MCI progressive patients tending to have longer follow-up times by about 4.5 months. Statistically, MCI progressive patients did not differ from MCI non-progressive patients in mean age, sex ratio, education, race, ethnicity, family history of AD, or *APOE ε*4 prevalence. See Table 2 for details.

Table 2. Baseline Characteristics of mild cognitive impairment (MCI) Study Sample.

Subjects with MCI (*n* = 91)	MCI Progressive Patients (*n* = 41)	MCI Non-Progressive Patients (*n* = 50)	*p*-Value [1]
Age (years)	75.17 ± 7.30	74.78 ± 7.44	0.8011
Male-to-female-ratio [2,3]	2.42 (29/12)	2.33 (35/15)	0.9394
Family history of AD [2,3]	36.39 (15/41)	36.00 (18/50)	0.9539
*APOE ε*4 carriers, % [2,3]	60.98 (25/41)	50.00 (25/50)	0.4036
Average follow-up time (months)	34.10 ± 9.70	29.64 ± 10.94	0.0426

[1] Data in parentheses are number of participants. [2] *p*-values obtained using χ^2-tests. [3] Use of χ^2- or *t*-test to compare difference between MCI progressive patients and MCI non-progressive patients. Unless otherwise indicated, the data is written as mean ± standard deviation and *p*-values were calculated using a *t*-test.

3.2. SNP-Enriched Pathways and Associated Crosstalks

Our analysis identified SNP-enriched pathways that represent six of the seven KEGG pathway categories, including Cellular Processes, Metabolism, Environmental Information Processing, Genetic Information Processing, Human Diseases, and Organismal Systems. This broad array of pathway categories represents the complex nature of AD pathogenesis, which has been attributed to many different biological mechanisms, ranging from amyloid toxicity to metabolic dysfunction to immune dysregulation. Figure 3 depicts the distribution of SNP-enriched pathways amongst the six KEGG categories. The majority of enriched pathways are classified under Human Diseases (31%). This supports the well-established relationships between AD and multiple other cardiovascular, autoimmune, and neurodegenerative diseases. For instance, diabetes, obesity, and heart diseases are well-established risk factors of AD, so much so that AD has been referred to as type 3 diabetes. As such, finding SNP-enriched pathways for cardiovascular, endocrine, and metabolic diseases in individuals with MCI is anticipated [53].

Similarly, the enrichment of metabolic pathways, organismal systems including nervous and immune system pathways, and common signaling pathways of the environmental information processing category is also expected and well-supported in the literature [54–68]. Interestingly, several genetic information processing pathways, including cell cycle regulation and DNA replication and repair, were found to be enriched. Evidence for the roles of these pathways in AD has only

recently begun to surface [69–71]. Our findings of the SNP-enrichment of these pathways among MCI individuals may provide support for further investigations into such pathways.

SNP-enriched pathway crosstalks were discovered between six KEGG categories, with the greatest number of crosstalks occurring between Human Diseases and Organismal systems. It is difficult to stipulate the significance of these findings. However, given that the etiology of many diseases, including AD, is complex and likely involves the failure/dysregulation of many pathways that are involved in the normal functioning of multiple organ systems, such significant crosstalk between these two categories among MCI individuals is not unexpected. The ageing process itself may facilitate a greater number of crosstalks in many pathways, since aging is associated with degeneration in many tissues and raises the risk for other chronic diseases besides dementia.
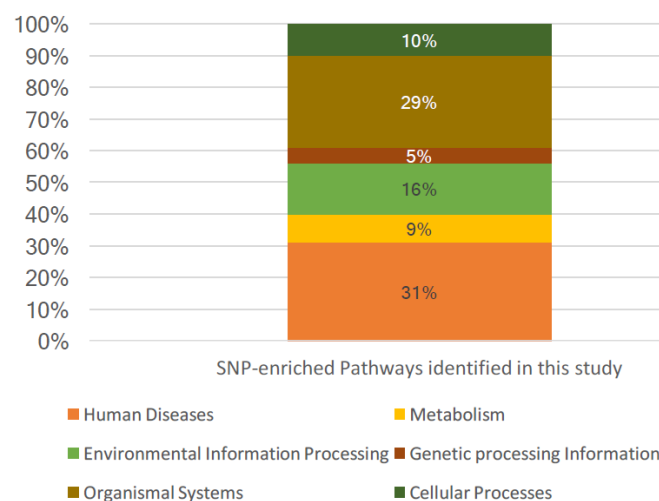
Figure 3. The distribution of the types of SNP-enriched pathways identified in this study and a comparison to the pathway distribution of the Kyoto Encyclopedia of Genes and Genomes (KEGG). NOTE: Although there are seven KEGG pathway categories, here we only show the six KEGG pathway categories that included identified SNP-enriched pathways in this study.

To investigate the genetic load in regards to AD, we further examined enriched pathway crosstalks specifically relating to the KEGG AD pathway. We identified 97 AD-related crosstalks and grouped the participating pathways by KEGG category (Figure 4).

Figure 4. Pathways found to have significant crosstalk with the AD pathway and corresponding KEGG categories (shown in colored blocks). Specific KEGG pathway types are listed below each category with the number of occurrences in parentheses. NOTE: Although there are seven KEGG pathway categories, here we only show the six KEGG pathway categories that included identified SNP-enriched pathways in this study.

In line with the overall findings of crosstalk enrichment, the AD-specific pathway crosstalks primarily fell between the categories Human Diseases and Organismal Systems, supporting the importance of the pathways within these categories in AD genetic load. In contrast, pathways of Metabolism and Genetic Information Processing had very few crosstalks, suggesting that genetic load in these processes is not as important to the disease process, at least in this particular cohort. Similar findings were seen in the analysis of all pathway crosstalks. Focusing in on the AD pathway, we observe significant crosstalk in between all pathway categories supporting the complex etiology of this disease.

3.3. SNP-Enriched Features with Baseline Clinical Parameters

We predicted MCI progression to AD dementia progression using a support vector machine (SVM) with a linear kernel function with baseline clinical parameters (age, education, and Alzheimer's disease assessment scale-cognitive subscale (ADAS-Cog)), significant pathways, or significant pathway crosstalks as predictors. The results for 100 iterations of 10-fold cross-validation are shown in Table 3. The model built with the clinical parameters only produced an accuracy of $59.19 \pm 2.46\%$ with $83.64 \pm 0.29\%$ of training data points as support vectors. The model built with significant pathways alone produced an accuracy of $56.78 \pm 3.5\%$ with $68.36 \pm 3.5\%$ support vectors. Typically, we expect a random guessing model to yield an accuracy of 50%; thus, both models only perform moderately above a random model.

Table 3. Performance of support vector machine (SVM) models with baseline clinical parameters.

Metrics	Baseline Clinical: Age, Education, ADAS-Cog	Significant Pathways (Only)	Significant Pathway Crosstalks (Only)	Baseline Clinical + Significant Pathways	Baseline Clinical + Significant Pathway Crosstalks
Accuracy in %	59.19 ± 2.46	56.78 ± 3.5	60.97 ± 3.24	64.57 ± 3.56	70.9 ± 3.3
Support Vectors in %	83.64 ± 0.29	68.36 ± 2.1	50.83 ± 4.77	63.3 ± 1.15	54.29 ± 0.56
True Positives	30.78 ± 1.7	31.21 ± 4.7	40.9 ± 3.1	33.64 ± 2.4	37.93 ± 2.16
False Negatives	19.22 ± 1.7	18.79 ± 4.7	9.06 ± 3.13	16.36 ± 2.4	12.07 ± 2.16
False Positives	17.9 ± 1.6	20.51 ± 3.09	26.46 ± 4.17	15.9 ± 2.03	14.41 ± 1.6
True Negatives	23.08 ± 1.55	20.49 ± 3.09	14.54 ± 4.17	25.11 ± 2.03	26.59 ± 1.6
Sensitivity	0.62 ± 0.03	0.62 ± 0.09	0.82 ± 0.06	0.67 ± 0.05	0.76 ± 0.04
Specificity	0.56 ± 0.04	0.51 ± 0.07	0.35 ± 0.11	0.61 ± 0.05	0.65 ± 0.04
Precision	0.62 ± 0.03	0.60 ± 0.09	0.61 ± 0.03	0.68 ± 0.03	0.74 ± 0.03

ADAS-Cog, Alzheimer's disease assessment scale-cognitive subscale.

A high percentage of support vectors indicate that an SVM model is overfitted and unlikely to generalize well. Thus, if we have two models that produce the same accuracy, then we pick the model that has the lower percentage of support vectors. Sixty-eight percent (68%) or more of the training data points were used as support vectors and this indicates highly overfitted models, which is shown by the poor cross-validation accuracy.

Incorporating both the baseline clinical parameters and significant pathways as predictors produced a model with an accuracy of $64.57 \pm 3.56\%$ with $63.3 \pm 1.15\%$ support vectors. This combined model demonstrated a 5.38% increase in accuracy compared to the baseline clinical parameters model and a 7.79% increase in accuracy compared to the model using significant pathways alone. Additionally, the reduced support vector percentage of this combined model indicates a better generalizability than the baseline clinical parameters model (20.34% decrease in support vectors) and the significant pathways model (5.04% decrease in support vectors).

With our novel approach of using significant pathway crosstalks to predict AD progression, our model provides an accuracy of 60.97 ± 3.24, which is higher than using baseline clinical parameters or significant pathways alone. Furthermore, this crosstalks model has the lower support vector percentage of $50.83 \pm 4.77\%$, and thus the greatest generalizability of all of the models.

The enhancement of the significant pathway crosstalks model with the inclusion of baseline clinical parameters produced a model that has the greatest accuracy of 70.9 ± 3.3% with a moderate support vectors percentage of 54.29 ± 0.56%. These initial results support the utility of using pathway crosstalks as significant predictors in the progression from MCI progression to AD dementia and warrant replication in larger samples followed for longer periods.

3.4. Comparison of Model Performances from Shaffer et al. (2013) with Our Model Performance including SNP-Enriched Features

We compared models built using the clinical parameters and the SNP-enriched features (significant pathways or significant pathway crosstalks) to a logistic regression model with only clinical parameters by Shaffer et al. [52] (Table 4). We also noticed that the average accuracy of the logistic regression model slightly increased (from 58.7% to 59.10 ± 1.71%) when we repeatedly created random 10-folds instead of using the 10 original folds from Shaffer et al. [52]. It decreased (to 57.04 ± 2%) when we removed the six patients that did not have corresponding SNP data in the ADNI database. Our method, when incorporating either significant pathways or significant pathway crosstalks, had a higher average accuracy on 100 randomly generated 10-folds than the method by Shaffer et al. [52]. Impressively, the combination of the baseline clinical parameters, *APOE ε*4, and significant pathway crosstalks in our logistic regression model yielded an accuracy of 72.1 ± 2.66. Also, a similar accuracy was obtained using a linear kernel SVM built on the SNP-enriched features. This indicates that the pathways and pathway crosstalks indeed lead to a better rate of prediction from MCI progression to AD dementia progression.

Table 4. Performance of Shaffer et al. [52] model with clinical parameters with 97 patients in comparison to our model with 97 and 91 patients.

Model	Logistic Regression				SVM with Linear Kernel	
No. Patient	97 Patients		91 Patients		91 Patients	
Metrics	original 10-fold cross-validation: Baseline Clinical + *APOE ε*4 Shaffer et al. paper [52]	100 Iterations of Random 10-fold cross-validation: Baseline Clinical + *APOE ε*4	100 Iterations of Random 10-fold cross-validation: Baseline Clinical + *APOE ε*4	100 Iterations of Random 10-fold cross-validation: Baseline Clinical + *APOE ε*4 + significant pathway crosstalks	100 Iterations of Random 10-fold cross-validation: Baseline Clinical + *APOE ε*4 + significant pathway	100 Iterations of Random 10-fold cross-validation: Baseline Clinical + *APOE ε*4 + significant pathway crosstalks
Accuracy in %	58.7	59.10 ± 1.71	57.04 ± 2	72.1 ± 2.66	63.56 ± 3.4	69.53 ± 2.9
Support Vectors in %	N/A	N/A	N/A	N/A	63.29 ± 1.16	53.65 ± 0.6
True Positives	17	39.98 ± 1.21	16.48 ± 1.16	39.22 ± 2.08	33.74 ± 2.4	37.6 ± 1.9
False Negatives	26	14.02 ± 1.21	24.43 ± 1.3	10.78 ± 2.08	16.3 ± 2.35	12.44 ± 1.9
False Positives	14	25.65 ± 1.23	14.67 ± 1.58	14.61 ± 1.3	16.91 ± 2	15.29 ± 1.5
True Negatives	40	17.35 ± 1.23	35.52 ± 1.53	26.39 ± 1.3	24.09 ± 2	25.71 ± 1.5
Sensitivity	0.40	0.74 ± 0.02	0.41 ± 0.05	0.78 ± 0.04	0.68 ± 0.05	0.75 ± 0.04
Specificity	0.74	0.40 ± 0.03	0.72 ± 0.03	0.64 ± 0.03	0.59 ± 0.05	0.63 ± 0.04
Precision	0.46	0.61 ± 0.02	0.54 ± 0.06	0.75 ± 0.03	0.68 ± 0.04	0.73 ± 0.03

3.5. Randomized SNP-Enriched Features

To demonstrate that the pathway crosstalks found in this study have true predictive power and the results are not a random occurrence, we generated 25 random samples of pathway crosstalks with no prior association to Alzheimer's and performed 100 iterations of 10-fold cross-validation for each of these 25 samples. The results are shown in Table 5.

Table 5. Performance of models with randomized pathway cross-talk features.

Metrics	Baseline Clinical + Randomized Significant Pathway Crosstalks	Baseline Clinical + Significant Pathway Crosstalks
Accuracy in %	59.27 ± 3.66	70.9 ± 3.3
Support Vectors in %	83.47 ± 1.84	54.29 ± 0.56
True Positives	30.86 ± 1.98	37.93 ± 2.16
False Negatives	19.14 ± 1.97	12.07 ± 2.16
False Positives	17.95 ± 1.59	14.41 ± 1.6
True Negatives	23.05 ± 1.56	26.59 ± 1.6
Sensitivity	0.62 ± 0.97	0.76 ± 0.04
Specificity	0.56 ± 1.45	0.65 ± 0.04
Precision	0.63 ± 0.02	0.74 ± 0.03

The model with the baseline clinical parameters and randomized significant pathway crosstalks gave an accuracy of 59.27 ± 3.66 with 83.47 ± 1.84 support vectors. This model yields 12% less accuracy and a 29.1% increase in support vectors, in comparison to the original model that uses baseline parameters and significant pathway crosstalks (instead of randomized significant pathway crosstalks). As expected, our randomly generated pathway crosstalks shows worse performance than significant pathway crosstalks. The model accuracy is still moderately above a random guessing model, likely due to the presence of the clinical parameters. A similar trend was seen when investigating models with baseline clinical parameters and all AD biomarkers to determine the effects of randomized pathways.

In this work, we focus on the development of a novel computational methodology for the discovery of pathway crosstalks to be used as biomarkers for the prognosis of AD. To demonstrate the efficacy of our methodology, we compared it with methods and results from prior studies in this area, which used ADNI-1 data. Although there is more recent data available, ADNI-1 data was used so that we could benchmark our methodology against these prior studies. In future work, we will continue our characterization efforts by incorporating the newer ADNI datasets as well as increasing the sensitivity of the proposed methodology through the use of the additive genetic model for the identification of patient-specific SNPs. There are also some limitations to our study. The ADNI is not a population-based study; it is essentially a biomarker cohort at research sites and our sample size was relatively small: we relied on a sample that was previously studied, since our initial goal was to examine the additive value of crosstalk biomarkers. We also did not incorporate other biomarkers such as tau, *p*-tau181P, *β*-amyloid1-42, *APOE ε*4, and microRNAs at this time, since our main focus was on methodological development for discovering and characterizing pathway crosstalks. However, the ADNI results have formed the basis for many current clinical prevention drug trials, and hence the ADNI is a highly relevant dataset. Moreover, its careful selection criteria and the way it makes available rich biomarker and genetic data and longitudinal cognitive data are enormous strengths. Indeed, the study of pathway crosstalks may yield novel insights into how AD pathological (e.g., beta-amyloid, tau) and neuronal loss (e.g., apoptosis, atrophy) mechanisms interact, and our methods lay the foundation for such future work.

The generic pathway crosstalk reference map was built using several different datasets, and hence the question arises as to whether all datasets should be treated equally. For simplicity, in this study, we treated all datasets equivalently. However, a modification to our information fusion method would allow us to introduce parameters to weigh evidence differently based on expert knowledge or trustworthiness. In the future, we would like to perform additional experiments to see the effects of these parameters on disease AD prognosis. This is non-trivial, as we would first need to define a weighting scheme and then develop additional methods to gauge the weights for different evidence.

4. Conclusions

AD is a major public health challenge, and there remain substantial gaps in our knowledge of its biology and treatment targets. Fully characterizing AD at a systems biology level is a priority

for these reasons. In this work, we demonstrate a new methodology to build a pathway crosstalk reference map using the combined power of several gene and protein knowledge antecedents, and use this to make AD-specific discovery pathway crosstalks by enrichment with patient-specific SNP information. Our pilot data documents the promise of utilizing those SNP-enriched pathway crosstalks to identify potential AD-linked mechanisms at a systems level. More specifically, we demonstrate a three-step methodology to build a generic pathway crosstalk reference map by combining several protein/gene evidence. We then used the identified pathway crosstalks from this map as potential AD biomarkers by enriching them with patient-specific SNP information. In an initial sample of at risk subjects, we found that utilizing SNP-enriched pathway crosstalks as additional features significantly improved the prediction accuracy of MCI progression to AD dementia progression.

In addition, we verified some previously identified pathways and identified some new pathway crosstalks that warrant further study. Furthermore, we built the prediction model including the identified pathways and crosstalks, and compared our model's outputs with a previous study. These prediction model comparison analyses show that the identified pathways and crosstalks can be used as significant biomarkers of MCI progression to AD dementia progression prediction with other clinical information. Additional analysis would be required to understand the biological mechanisms that explain the association of these pathways to AD.

In summary, this is the first report to our knowledge that characterizes biological crosstalk pathways in subjects at risk of AD using gene and protein knowledge antecedents and studies their potential utility as prognostic biomarkers. Further application of this methodology to the full ADNI-1 and ADNI-2 cohort as well as to other population studies is warranted, and may yield further insights into disease mechanisms as well as novel targets for biomarker development and drug discovery.

Author Contributions: K.P. and N.F.S. conceived and designed the computational and genetic crosstalk experiments with advice from R.E.T., A.J.S., J.R.P., and P.M.D.; P.M.D. oversaw biomarker and clinical data collection at Duke and A.J.S. headed the genetic core of A.D.N.I. K.P. performed the experiments and analyzed the data with K.N, S.H., G.B., and P.T.D; K.P. wrote the paper and D.S., K.S., and P.S.Y. did the manuscript drafting. All authors assisted with interpretation and editing.

References

1. Brookmeyer, R.; Evans, D.A.; Hebert, L.; Langa, K.M.; Heeringa, S.G.; Plassman, B.L.; Kukull, W.A. National estimates of the prevalence of Alzheimer's disease in the United States. *Alzheimers Dement.* **2011**, *7*, 61–73. [CrossRef] [PubMed]
2. Alzheimer's Association. 2017 Alzheimer's Disease Facts and Figures. *Alzheimers Dement.* **2017**, *13*, 325–373.
3. Heron, M. Deaths: Leading causes for 2010. *Natl. Vital Stat. Rep.* **2013**, *62*, 1–96. [PubMed]
4. Saykin, A.J.; Shen, L.; Yao, X.; Kim, S.; Nho, K.; Risacher, S.L.; Ramanan, V.K.; Foroud, T.M.; Faber, K.M.; Sarwar, N.; et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. *Alzheimers Dement.* **2015**, *11*, 792–814. [CrossRef] [PubMed]
5. Sheinerman, K.S.; Tsivinsky, V.G.; Abdullah, L.; Crawford, F.; Umansky, S.R. Plasma microRNA biomarkers for detection of mild cognitive impairment: Biomarker validation study. *Aging Albany N.Y.* **2013**, *5*, 925–938. [CrossRef] [PubMed]
6. Galimberti, D.; Villa, C.; Fenoglio, C.; Serpente, M.; Ghezzi, L.; Cioffi, S.M.; Arighi, A.; Fumagalli, G.; Scarpini, E. Circulating miRNAs as potential biomarkers in Alzheimer's disease. *Alzheimers Dis.* **2014**, *42*, 1261–1267. [CrossRef]
7. Femminella, G.D.; Ferrara, N.; Rengo, G. The emerging role of microRNAs in Alzheimer's disease. *Front. Physiol.* **2015**, *6*. [CrossRef] [PubMed]
8. Chen, Z.; Padmanabhan, K.; Rocha, A.M.; Shpanskaya, Y.; Mihelcic, J.R.; Scott, K.; Samatova, N.F. SPICE: Discovery of phenotype-determining component interplays. *BMC Syst. Biol.* **2012**, *6*, 40. [CrossRef] [PubMed]
9. Goh, C.S.; Gianoulis, T.A.; Liu, Y.; Li, J.; Paccanaro, A.; Lussier, Y.A.; Gerstein, M. Integration of curated databases to identify genotype-phenotype associations. *BMC Genom.* **2006**, *7*, 257. [CrossRef] [PubMed]
10. Gonzalez, O.; Zimmer, R. Assigning functional linkages to proteins using phylogenetic profiles and continuous phenotypes. *Bioinformatics* **2008**, *24*, 1257–1263. [CrossRef] [PubMed]
11. Hendrix, W.; Rocha, A.M.; Elmore, M.T.; Trien, J.; Samatova, N.F. Discovery of Enriched Biological Motifs Using Knowledge Priors with Application to Biohydrogen Production. In Proceedings of the BIOCOMP, Las Vegas, NV, USA, 12–15 July 2010; pp. 17–23.
12. Hendrix, W.; Rocha, A.M.; Padmanabhan, K.; Choudhary, A.; Scott, K.; Mihelcic, J.R.; Samatova, N.F. DENSE: Efficient and prior knowledge-driven discovery of phenotype associated protein functional modules. *BMC Syst. Biol.* **2011**, *5*, 172. [CrossRef] [PubMed]
13. Jim, K.; Parmar, K.; Singh, M.; Tavazoie, S. A cross-genomic approach for systematic mapping of phenotypic traits to genes. *Genom. Res.* **2004**, *14*, 109–115. [CrossRef] [PubMed]
14. Schmidt, M.C.; Rocha, A.M.; Padmanabhan, K.; Chen, Z.; Scott, K.; Mihelcic, J.R.; Samatova, N.F. Efficient α,β-motif finder for identification of phenotype-related functional modules. *BMC Bioinform.* **2011**, *12*, 440. [CrossRef] [PubMed]
15. Schmidt, M.C.; Rocha, A.M.; Padmanabhan, K.; Shpanskaya, Y.; Banfield, J.; Scott, K.; Mihelcic, J.R.; Samatova, N.F. NIBBS-Search for fast and accurate prediction of phenotype-biased metabolic systems. *PLoS Comput. Biol.* **2012**, *8*, e1002490. [CrossRef] [PubMed]
16. Slonim, N.; Elemento, O.; Tavazoie, S. Ab initio genotype-phenotype association reveals intrinsic modularity in genetic networks. *Mol. Syst. Biol.* **2006**, *2*. [CrossRef] [PubMed]
17. Korbel, J.O.; Doerks, T.; Jensen, L.J.; Perez-Iratxeta, C.; Kaczanowski, S.; Hooper, S.D.; Andrade, M.A.; Bork, P. Systematic association of genes to phenotypes by genome and literature mining. *PLoS Biol.* **2005**, *3*, e134. [CrossRef] [PubMed]
18. Levesque, M.; Shasha, D.; Kim, W.; Surette, M.G.; Benfey, P.N. Trait-to-Gene: A computational method for predicting the function of uncharacterized genes. *Curr. Biol.* **2003**, *13*, 129–133. [CrossRef]
19. Tamura, M.; D'haeseleer, P. Microbial genotype-phenotype mapping by class association rule mining. *Bioinformatics* **2008**, *24*, 1523–1529. [CrossRef] [PubMed]
20. Zalesky, A.; Fornito, A.; Bullmore, E.T. Network-based statistic: Identifying differences in brain networks. *Neuroimage* **2010**, *53*, 1197–1207. [CrossRef] [PubMed]
21. Ballatore, C.; Lee, V.M.; Trojanowski, J.Q. Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nat. Rev. Neurosci.* **2007**, *8*, 663–672. [CrossRef] [PubMed]

22. Lanni, C.; Uberti, D.; Racchi, M.; Govoni, S.; Memo, M. Unfolded p53: A potential biomarker for Alzheimer's disease. *J. Alzheimers Dis.* **2007**, *12*, 93–99. [CrossRef] [PubMed]

23. Selkoe, D.J. Alzheimer's disease: Genes, proteins, and therapy. *Physiol. Rev.* **2001**, *81*, 741–766. [CrossRef] [PubMed]

24. Li, Y.; Agarwal, P.; Rajagopalan, D. A global pathway crosstalk network. *Bioinformatics* **2008**, *24*, 1442–1447. [CrossRef] [PubMed]

25. Godoy Zeballos, J.A. Signaling pathway cross talk in Alzheimer's disease. *Cell Commun. Signal.* **2014**, *12*, 23. [CrossRef] [PubMed]

26. Ramanan, V.K.; Saykin, A.J. Pathways to neurodegeneration: Mechanistic insights from GWAS in Alzheimer's disease, Parkinson's disease, and related disorders. *Am. J. Neurodegener. Dis.* **2013**, *2*, 145–175. [PubMed]

27. Myers, C.L.; Robson, D.; Wible, A.; Hibbs, M.A.; Chiriac, C.; Theesfeld, C.L.; Dolinski, K.; Troyanskaya, O.G. Discovery of biological networks from diverse functional genomic data. *Genome Biol.* **2005**, *6*, R114. [CrossRef] [PubMed]

28. Xu, Y.; Hu, W.; Chang, Z.; Duanmu, H.; Zhang, S.; Li, Z.; Yu, L.; Li, X. Prediction of human protein-protein interaction by a mixed Bayesian model and its application to exploring underlying cancer-related pathway crosstalk. *J. R. Soc. Interface* **2011**, *8*, 555–567. [CrossRef] [PubMed]

29. Liu, Z.P.; Wang, Y.; Zhang, X.S.; Chen, L. Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer's disease brains. *BMC Syst. Biol.* **2010**, *4*, S11. [CrossRef] [PubMed]

30. Li, Y. Pathway crosstalk network. In *Systems Biology for Signaling Networks*; Choi, S., Ed.; Springer: New York, NY, USA, 2010; pp. 491–504.

31. Hartman, J.L.; Garvik, B.; Hartwell, L. Principles for the buffering of genetic variation. *Science* **2001**, *291*, 1001–1004. [CrossRef] [PubMed]

32. Suthers, P.F.; Zomorrodi, A.; Maranas, C.D. Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Mol. Syst. Biol.* **2009**, *5*, 301. [CrossRef] [PubMed]

33. North, B.V.; Curtis, D.; Sham, P.C. A note on the calculation of empirical *p*-values from Monte Carlo procedures. *Am. J. Hum. Genet.* **2002**, *71*, 439–441. [CrossRef] [PubMed]

34. Bailey, T.L.; Gribskov, M. Combining evidence using p-values: Application to sequence homology searches. *Bioinformatics* **1998**, *14*, 48–54. [CrossRef] [PubMed]

35. Feller, W. *An Introduction to Probability Theory and Its Applications*; Wiley: New York, NY, USA, 1957.

36. Silver, M.; Janousova, E.; Hua, X.; Thompson, P.M.; Montana, G. Identification of gene pathways implicated in Alzheimer's disease using longitudinal imaging phenotypes with sparse regression. *Neuroimage* **2012**, *63*, 1681–1694. [CrossRef] [PubMed]

37. Silver, M.; Montana, G. Fast identification of biological pathways associated with a quantitative trait using group lasso with overlaps. *Stat. Appl. Genet. Mol. Biol.* **2012**, *11*. [CrossRef] [PubMed]

38. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef] [PubMed]

39. Kanehisa, M.; Goto, S.; Furumichi, M.; Tanabe, M.; Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **2010**, *38*, 355–360. [CrossRef] [PubMed]

40. Kanehisa, M.; Goto, S.; Hattori, M.; Aoki-Kinoshita, K.F.; Itoh, M.; Kawashima, S.; Katayama, T.; Araki, M.; Hirakawa, M. From genomics to chemical genomics: New developments in KEGG. *Nucleic Acids Res.* **2006**, *34*, 354–357. [CrossRef] [PubMed]

41. Stark, C.; Breitkreutz, B.J.; Reguly, T.; Boucher, L.; Breitkreutz, A.; Tyers, M. BioGRID: A general repository for interaction datasets. *Nucleic Acids Res.* **2005**, *34*, 535–539. [CrossRef] [PubMed]

42. Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **2010**, *38*, W214–W220. [CrossRef] [PubMed]

43. Kawaji, H.; Severin, J.; Lizio, M.; Forrest, A.R.; van Nimwegen, E.; Rehli, M.; Schroder, K.; Irvine, K.; Suzuki, H.; Carninci, P.; et al. Update of the FANTOM web resource: From mammalian transcriptional landscape to its dynamic regulation. *Nucleic Acids Res.* **2010**, *39*, 856–860. [CrossRef] [PubMed]

44. Kawaji, H.; Severin, J.; Lizio, M.; Waterhouse, A.; Katayama, S.; Irvine, K.M.; Hume, D.A.; Forrest, A.R.R.; Suzuki, H.; Carninci, P.; Hayashizaki, Y.; Daub, C.O. The FANTOM web resource: From mammalian transcriptional landscape to its dynamic regulation. *Genome Biol.* **2009**, *10*, R40. [CrossRef] [PubMed]

45. Li, T.; Du, P.; Xu, N. Identifying human kinase-specific protein phosphorylation sites by integrating heterogeneous information from various sources. *PLoS ONE* **2010**, *5*, e15411. [CrossRef] [PubMed]

46. Davis, A.P.; Murphy, C.G.; Johnson, R.; Lay, J.M.; Lennon-Hopkins, K.; Saraceni-Richards, C.; Sciaky, D.; King, B.L.; Rosenstein, M.C.; Wiegers, T.C.; et al. The comparative toxicogenomics database: Update 2013. *Nucleic Acids Res.* **2013**, *41*, 1104–1114. [CrossRef] [PubMed]

47. Cariaso, M.; Lennon, G. SNPedia: A wiki supporting personal genome annotation, interpretation and analysis. *Nucleic Acids Res.* **2012**, *40*, 1308–1312. [CrossRef] [PubMed]

48. Mrak, R.E.; Griffin, S. Interleukin-1, neuroinflammation, and Alzheimer's disease. *Neurobiol. Aging* **2001**, *22*, 903–908. [CrossRef]

49. Wiener, H.W.; Perry, R.T.; Chen, Z.; Harrell, L.E.; Go, R.C. A polymorphism in *SOD2* is associated with development of Alzheimer's disease. *Genes Brain Behav.* **2007**, *6*, 770–775. [CrossRef] [PubMed]

50. Dahiyat, M.; Cumming, A.; Harrington, C.; Wischik, C.; Xuereb, J.; Corrigan, F.; Breen, G.; Shaw, D.; St Clair, D. Association between Alzheimer's disease and the *NOS3* gene. *Ann. Neurol.* **1999**, *46*, 664–667. [CrossRef]

51. Weiner, M.W.; Veitch, D.P.; Aisen, P.S.; Beckett, L.A.; Cairns, N.J.; Green, R.C.; Harvey, D.; Jack, C.R.; Jagust, W.; Liu, E.; et al. The Alzheimer's disease neuroimaging initiative: A review of papers published since its inception. *Alzheimers Dement.* **2012**, *8*, 1–68. [CrossRef] [PubMed]

52. Shaffer, J.L.; Petrella, J.R.; Sheldon, F.C.; Choudhury, K.R.; Calhoun, V.D.; Coleman, R.E.; Doraiswamy, P.M. Predicting cognitive decline in subjects at risk for Alzheimer disease by using combined cerebrospinal fluid, MR imaging, and PET biomarkers. *Radiology* **2013**, *266*, 583–591. [CrossRef] [PubMed]

53. Vignini, A.; Giulietti, A.; Nanetti, L.; Raffaelli, F.; Giusti, L.; Mazzanti, L.; Provinciali, L. Alzheimer's disease and diabetes: New insights and unifying therapies. *Curr. Diabetes Rev.* **2013**, *9*, 218–227. [CrossRef] [PubMed]

54. Ramanan, V.K.; Kim, S.; Holohan, K.; Shen, L.; Nho, K.; Risacher, S.L.; Foroud, T.M.; Mukherjee, S.; Crane, P.K.; Aisen, P.S.; et al. Genome-wide pathway analysis of memory impairment in the Alzheimer's disease neuroimaging initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. *Brain Imaging Behav.* **2012**, *6*, 634–648. [CrossRef] [PubMed]

55. Roe, C.M.; Behrens, M.I.; Xiong, C.; Miller, J.P.; Morris, J.C. Alzheimer disease and cancer. *Neurology* **2005**, *64*, 895–898. [CrossRef] [PubMed]

56. Sardi, F.; Fassina, L.; Venturini, L.; Inguscio, M.; Guerriero, F.; Rolfo, E.; Ricevuti, G. Alzheimer's disease, autoimmunity and inflammation. The good, the bad and the ugly. *Autoimmun. Rev.* **2011**, *11*, 149–153. [CrossRef] [PubMed]

57. Honjo, K.; van Reekum, R.; Verhoeff, N.P. Alzheimer's disease and infection: Do infectious agents contribute to progression of Alzheimer's disease? *Alzheimers Dement.* **2009**, *5*, 348–360. [CrossRef] [PubMed]

58. Jung, B.K.; Pyo, K.H.; Shin, K.Y.; Hwang, Y.S.; Lim, H.; Lee, S.J.; Moon, J.H.; Lee, S.H.; Suh, Y.H.; Chai, J.Y.; et al. Toxoplasma gondii infection in the brain inhibits neuronal degeneration and learning and memory impairments in a murine model of Alzheimer's disease. *PLoS ONE* **2012**, *7*, e33312. [CrossRef] [PubMed]

59. Al-Mansoori, K.M.; Hasan, M.Y.; Al-Hayani, A.; El-Agnaf, O.M. The role of α-synuclein in neurodegenerative diseases: From molecular pathways in disease to therapeutic approaches. *Curr. Alzheimer Res.* **2013**, *10*, 559–568. [CrossRef] [PubMed]

60. Mandrekar-Colucci, S.; Karlo, J.C.; Landreth, G.E. Mechanisms underlying the rapid peroxisome proliferator-activated receptor-γ-mediated amyloid clearance and reversal of cognitive deficits in a murine model of Alzheimer's disease. *J. Neurosci.* **2012**, *32*, 10117–10128. [CrossRef] [PubMed]

61. Berntorp, K.; Frid, A.; Alm, R.; Fredrikson, G.N.; Sjöberg, K.; Ohlsson, B. Antibodies against gonadotropin-releasing hormone (GnRH) in patients with diabetes mellitus is associated with lower body weight and autonomic neuropathy. *BMC Res. Notes* **2013**, *6*, 329. [CrossRef] [PubMed]

62. Giunta, B.; Fernandez, F.; Nikolic, W.V.; Obregon, D.; Rrapo, E.; Town, T.; Tan, J. Inflammaging as a prodrome to Alzheimer's disease. *J. Neuroinflamm.* **2008**, *5*, 51. [CrossRef] [PubMed]

63. Liu, Y.H.; Zeng, F.; Wang, Y.R.; Zhou, H.D.; Giunta, B.; Tan, J.; Wang, Y.J. Immunity and Alzheimer's disease: Immunological perspectives on the development of novel therapies. *Drug Discov. Today* **2013**, *18*, 1212–1220. [CrossRef] [PubMed]

64. Freo, U.; Pizzolato, G.; Dam, M.; Ori, C.; Battistin, L. A short review of cognitive and functional neuroimaging studies of cholinergic drugs: Implications for therapeutic potentials. *J. Neural Transm.* **2002**, *109*, 857–870. [CrossRef] [PubMed]

65.　Tamburri, A.; Dudilot, A.; Licea, S.; Bourgeois, C.; Boehm, J. NMDA-Receptor activation but not ion flux is required for amyloid-beta induced synaptic depression. *PLoS ONE* **2013**, *8*, e65350. [CrossRef] [PubMed]

66.　Chen, Q.S.; Wei, W.Z.; Shimahara, T.; Xie, C.W. Alzheimer amyloid beta-peptide inhibits the late phase of long-term potentiation through calcineurin-dependent mechanisms in the hippocampal dentate gyrus. *Neurobiol. Learn. Mem.* **2002**, *77*, 354–371. [CrossRef] [PubMed]

67.　Ghebranious, N.; Mukesh, B.; Giampietro, P.F.; Glurich, I.; Mickel, S.F.; Waring, S.C.; Mc-Carty, C.A. A pilot study of gene/gene and gene/environment interactions in Alzheimer disease. *Clin. Med. Res.* **2011**, *9*, 17–25. [CrossRef] [PubMed]

68.　Mosch, B.; Morawski, M.; Mittag, A.; Lenz, D.; Tarnok, A.; Arendt, T. Aneuploidy and DNA replication in the normal human brain and Alzheimer's disease. *J. Neurosci.* **2007**, *27*, 6859–6867. [CrossRef] [PubMed]

69.　Yurov, Y.B.; Vorsanova, S.G.; Iourov, I.Y. The DNA replication stress hypothesis of Alzheimer's disease. *Sci. World J.* **2011**, *11*, 2602–2612. [CrossRef] [PubMed]

70.　Yang, Y.; Geldmacher, D.S.; Herrup, K. DNA replication precedes neuronal cell death in Alzheimer's disease. *J. Neurosci.* **2001**, *21*, 2661–2668. [PubMed]

71.　Bradley, W.G.; Polinsky, R.J.; Pendlebury, W.W.; Jones, S.K.; Nee, L.E.; Bartlett, J.D.; Hartshorn, J.N.; Tandan, R.; Sweet, L.; Magin, G.K. DNA repair deficiency for alkylation damage in cells from Alzheimer's disease patients. *Prog. Clin. Biol. Res.* **1989**, *317*, 715–732. [PubMed]