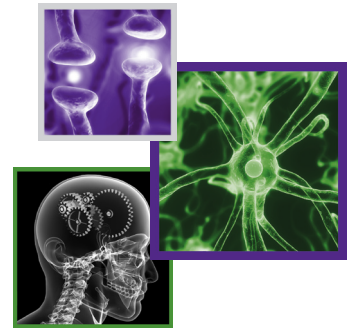
INTERVIEW

Neurodegenerative Disease Management



Identifying therapeutic targets for Alzheimer's disease with big data

Nilüfer Ertekin-Taner is a Professor of Neuroscience and Neurology at the Mayo Clinic, Jacksonville, FL, USA: A neurogeneticist and board-certified behavioral neurologist, she received her medical degree from Hacettepe University Medical School in Ankara, Turkey and her doctorate degree in Molecular Neuroscience from Mayo Graduate School. She completed her residency training in the Department of Neurology at Mayo Clinic in Rochester (MN, USA) and fellowship in Behavioral Neurology at Mayo Clinic in Jacksonville. Her laboratory aims to discover and characterize genetic factors underlying the complex genetics of Alzheimer's disease (AD) and related neurodegenerative conditions. Her earlier work contributed to the establishment of the endophenotype approach in genetic studies of AD and pioneered the use of amyloid β peptide levels as an endophenotype in AD genetic research. Her laboratory currently uses biological traits such as gene expression levels and cognitive scores and leverages combined genome, transcriptome and epigenetic data to uncover genetic risk factors for neurodegenerative conditions. She is the principal investigator of numerous NIH and foundation grants. She leads multiple collaborative projects aimed at gene and pathway discoveries in AD and other neurodegenerative diseases, as a part of the NIH initiatives Accelerating Medicines Partnership AD (AMP-AD) and Molecular Mechanisms of the Vascular Etiology of AD (M^2OVE-AD) consortia. She is the Principal Investigator of the Florida Consortium for African–American AD Studies (FCA^3DS). As a neurologist in the Memory Disorders Clinic at Mayo Clinic in Jacksonville, Florida, she continues to evaluate and manage patients with AD and other dementias and aspires to improve the quality of care that we provide to our patients and their families.

First draft submitted: 7 February 2017; Accepted for publication: 9 February 2017; Published online: 22 May 2017

Q Your current research focuses on identifying therapeutic targets & biomarkers for Alzheimer's disease with big data; can you explain what this research involves?

Big data research in health sciences is the analysis of very large datasets to identify specific patterns and associations pertinent to human biology and disease. To explain how we and others apply this approach to Alzheimer's disease (AD), we first need to briefly describe the complexity and magnitude of the health problem posed by AD.

AD is the most common cause of dementia in the elderly and a deadly epidemic

KEYWORDS
• Alzheimer's disease • big data • biomarker • genome • network analysis • prevention • therapeutic targets • transcriptome

Mayo Clinic, Departments of Neurology and Neuroscience, 4500 San Pablo Road, Jacksonville, FL 32224, USA; taner.nilufer@mayo.edu

Future Medicine part of fsg

that affects nearly 44 million people worldwide. Currently, there is no cure for this condition. AD is characterized by two pathologic hallmarks in the brain, namely senile plaques and neurofibrillary tangles, composed of amyloid β peptide and tau protein respectively; and a typical clinical presentation of memory loss. Decades of research focused on the genetics, neuropathology, biomarkers and model systems of AD identified perturbations in a multitude of biological pathways including but not limited to proteostasis, lipid metabolism, inflammation and vascular and oxidative mechanisms.

Big data research aims to identify the biological pathways and key molecules in those pathways that are perturbed in AD through the generation and analysis of large-scale –omics (e.g., transcriptome, epigenetic and proteome) and clinical data. The underlying hypothesis is that genetic and environmental risk factors for AD influence changes in gene and/or protein expression, which in turn leads to perturbations in biological pathways, ultimately culminating in clinical AD.

Gene and protein expression levels measured in tissue samples from large cohorts of AD and control patients are subjected to computational analyses. These analyses enable identification of networks of genes that have concerted changes in AD patients. The networks that correlate with AD and AD-related outcomes, such as pathology, imaging or cognition, can lead to nominations of pathways and molecules that are critical for the initiation and propagation of AD pathophysiology. The nominated molecules will serve as candidate drug targets and/or biomarkers.

It should be stated that the big data approach taken in isolation cannot distinguish whether the perturbed networks are a cause or consequence of the disease. The roles of the nominated molecules in various aspects of the disease cascade need to be investigated through additional studies including validations in model systems. Nevertheless, big data research is a powerful approach in generating testable hypotheses regarding novel drug and biomarker targets in AD.

Q Could you also provide an insight into the goal & successes achieved so far through this research?

The first examples of big data research in AD are the genome-wide association studies (GWAS) conducted in AD and control cohorts of tens of thousands of individuals. These studies led to the identification of 20 novel genetic loci where genetic variants are associated with risk for AD. Although these findings constitute a major milestone in AD research, in order to translate this success to therapeutic and biomarker discoveries, we need to identify the actual risk gene and functional variants at these AD susceptibility loci, and importantly, uncover their mechanism of action. In addition, the collective risk factors identified from the AD risk GWAS do not account for the complete genetic risk component of AD, suggesting the presence of additional factors.

To bridge this knowledge gap, alternative approaches are utilized including use of endophenotypes such as gene expression, neuropathology and cognitive measures as AD-related phenotypes in genetic research. These alternative approaches have been instrumental in the characterization of the AD GWAS risk loci for their effects on these endophenotypes, thus providing testable hypothesis for the mechanism of action of the genes and variants at these loci. Furthermore, they led to the prioritization of additional risk variants for downstream validation analyses.

More recently, we [1] and others are generating and analyzing large-scale multiomics data from well-characterized cohorts of AD and control individuals as part of initiatives including NIH Accelerating Medicines Partnership AD and Molecular Mechanisms of the Vascular Etiology of AD consortia. These initiatives embrace an open-science approach with significant collaborations, transparent and reproducible research and wide sharing of data with the scientific community through the Accelerating Medicines Partnership AD Knowledge Portal [2]. Using these data, we and others are generating network models of AD and utilizing these networks to nominate candidate drug targets and biomarkers. These candidates will be validated in model systems. The ultimate goal is to translate these candidates to viable therapeutic and biomarker discoveries through the partnership of multiple academic and pharmaceutical industry teams.

Q One of the ultimate goals of your lab is to commence drug therapy in the presymptomatic stage for AD, how do you think this can be achieved successfully?

Longitudinal neuroimaging, cerebrospinal fluid and cognitive biomarker studies in AD have revealed that the pathophysiologic cascades in this disease began decades before the onset of symptoms. Given this and the failure of clinical trials to date which may at least in part be due to

intervening too late in the disease process, there are now prevention trials targeting asymptomatic individuals at risk for developing AD. Current prevention trials are being conducted in carriers of deterministic early onset familial AD mutations, carriers of two copies of the *APOE e4* risk allele or those who have biomarkers for AD. These trials will be instrumental in testing the effects of the trial drugs, primarily aimed at the amyloid cascade, in preventing or delaying disease onset.

I envision the next generation of prevention trials to include additional at-risk groups, to utilize trial drugs targeting various pathophysiologic cascades, to leverage a novel set of biomarkers for patient selection and stratification and to individualize trial drugs to selective set of patients based on their specific risk profile. It is clear that AD is a heterogeneous disease with substantial variability from patient to patient with respect to their risk factors and their downstream pathophysiologic consequences. To achieve success in generating and testing therapies for AD, we have to acknowledge and embrace this heterogeneity. The expectation is that the big data research will identify additional biomarkers that will be informative about the stage of disease as well as specific pathways that are perturbed in individual patients. Drugs specifically targeting not one but multiple perturbed pathways will be utilized in the appropriate group of patients. The target engagement of the drug is then followed via pathway-specific biomarkers in addition to key clinical outcomes. While to some this may sound like science fiction, these therapeutic concepts are currently being applied in other areas of healthcare for complex conditions, such as cancer. Given the advances in our understanding of AD pathophysiology, we have every reason to believe that such next-generation preventative trials will be launched for this condition in the near future.

Q Considering that AD is highly heterogeneous, how do you feel the future of personalized medicine for the disease will progress?

This question is partly answered above. As a clinician, it is evident to me that each patient has a 'different Alzheimer's disease' with respect to the clinical picture of this condition, despite the common aspects. The heterogeneity is also evident in studies of biomarkers, neuropathology and genetics. Therapeutic success will likely require targeting common pathways such as the amyloid cascade in the majority of the patients, while also being able to provide individualized therapies aimed at specific pathways that are preferentially perturbed in subsets of patients.

Q What, in your opinion, are some of the major hurdles that have prevented the full translation of genomics into personalized healthcare for neurodegenerative diseases?

This is in part due to the fact that the genomic revolution is a relatively recent entity. The first well-powered GWAS in AD was published in 2009. Since then, 20 novel genetic loci for AD risk have been discovered. Research is underway to identify the specific genetic variants at these loci and their mechanism of action. This knowledge is necessary to translate these discoveries to personalized genetic risk stratification paradigms, novel therapies or biomarkers. Another difficulty is the heterogeneity of AD, which we have addressed above. One common challenging aspect of neurodegenerative diseases is the fact that they have a very long 'incubation period', where pathophysiologic changes occur decades before clinical symptoms become apparent. This makes it difficult to determine whether the transcriptome, epigenetic and proteome changes detected in patients are a cause or consequence of the disease, which has implications for drug discovery efforts that leverage such –omics datasets. Furthermore, success of clinical trials may depend on targeting patients at much earlier stages of their disease than has been the case. This requires accurate identification and risk stratification of patients, which increases cost and duration of clinical trials. Despite these hurdles, there is now a significant momentum and concerted efforts by the NIH and other institutions to bring together teams with complementary expertise in genomics, drug development and clinical trials with the common goal of finding a cure for AD and other neurodegenerative diseases. My hope and expectation is that the translation of genomic research to personalized healthcare will hit its stride within the next decade.

Q What has been your most significant academic achievement to date?

To quote my favorite poet, Nazim Hikmet, *"The most beautiful days we haven't seen yet. And the most beautiful words I wanted to tell you, I haven't said yet."* For me the best part of what I do every day comes from the privilege in working with the colleagues that I have, from generating and analyzing these amazingly complex datasets that we are just beginning to understand, and from

being part of this singular quest to find a cure for AD. I consider myself fortunate to be able to do these, while continuing to see my patients. I would also like to think that the most significant achievement is the one we have not reached yet.

Q Your lab recently won a grant to study & identify vascular risk factors in aging & dementia; do you believe that targeting these factors could be used to prevent dementia?

We know from epidemiological studies that vascular risk factors such as diabetes, hypertension and high cholesterol are also risk factors for AD and other dementias. Many patients with AD neuropathology also have concurrent vascular pathologies. The most substantial genetic risk factor for late-onset AD, APOE-ε4, is also associated with vascular disease outcomes. Hence, there is already significant evidence for a role of vascular risk factors in the pathophysiology of AD and other dementias, however, the precise molecular mechanisms underlying this vascular contribution to neurodegeneration is unknown. Together with my colleague, Dr Guojun Bu at Mayo Clinic in Jacksonville, FL, we have been funded to identify novel genes and pathways that influence vascular risk in aging, AD and other dementias. We will also specifically explore the role of APOE- and sex-dependent mechanisms. This grant is one of five funded as part of the NIH Molecular Mechanisms of the Vascular Etiology of AD consortium [3]. I believe that the collective outcomes from these and similar research studies can identify novel molecular targets by which vascular pathways influence AD, thus leading to new therapeutic and biomarker candidates. In addition, epidemiologic studies, which focus on control of vascular risk factors and which promote cardiovascular health, will continue to inform the public on the role of these measures in prevention of dementia.

Q How do you see the field of neurogenetics developing over the next few years?

Genomic revolution enabled large-scale genotyping and sequencing of the human genome, thereby generating repositories of human DNA variation, sequence, transcriptome and epigenetic data. These data in isolation have limited use for transformative healthcare research. The next big step is to link these data to multiple layers of clinically relevant phenotypes to identify patterns and associations that uncover the genes, pathways and mechanisms that underlie various aspects of human neurologic diseases. Another missing link in neurogenetics research is the interaction of environmental factors with genetics, a highly challenging area of research. Deeply phenotyped longitudinal cohorts with well-documented environmental exposures coupled with genomic data can be instrumental to fill this knowledge gap. Given cost, lack of sample availability or both, neurogenetics research has largely been focused on 'static' –omics data, such as DNA sequence or single time-point transcriptome or epigenetic data. Just as protein levels are utilized as biomarkers in longitudinal studies, transcripts and epigenetic factors, which are more immediate outcomes of risk factors, can also be exploited as such. Once these –omics measures are established as biomarkers of disease, their utility in clinical trials as described above can be explored.

Q Do you have a particular message to our readership or the general research/clinical community?

This is an exceptional time in neurogenetics research due to the convergence of technological advancement, availability of well-characterized study cohorts and model systems and multidisciplinary initiatives. Despite the complexities and challenges in neurodegenerative research, there are also many opportunities. I would hope to see especially junior investigators and those who are new to the field of neurogenetics get involved in and take advantage of the vast amount of data being generated. Some of the most exciting findings in science can arise from seemingly unlikely partnerships. We can expect these partnerships to lead to discoveries that will translate into clinical benefits for our patients.

Disclaimer

The opinions expressed in this interview are those of the interviewee and do not necessarily reflect the views of Future Medicine Ltd.

Financial & competing interests disclosure

N Ertekin-Taner is funded by NIH U01 AG046139, R01 NS080820, RF1 AG051504 and R21 AG048101. She served as an unpaid consultant for Cytox. N Ertekin-Taner has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1 Allen M, Carrasquillo MM, Funk C *et al.* Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. *Sci. Data* 3, 160089 (2016).

2 AMP-AD Knowledge Portal. www.synapse.org/#!Synapse:syn2580853

3 NIH M2OVE-AD. http://sagebase.org/research-projects