The human genome is composed of viral DNA: Viral homologues of the translated products cause Alzheimer's disease and other disorders via autoimmune and other mechanisms.

C.J. Carter

PolygenicPathways, Flat 4, 20 Upper Maze hill, Saint Leonard's on Sea. East Sussex, TN38 OLG, UK

Chris_car@yahoo.com

The human genome is composed of millions of fragmented contiguous viral DNA sequences, dating from the dawn of evolution and reflecting retroviral insertions over millions of years of coexistence. Herpes and other viral insertion points correspond to the locations of over 120 Alzheimer's disease susceptibility genes and to linkage hotspots. The greater the number of pathogen matches, the more important the gene. These DNA sequences are translated into short contiguous 5-12 amino acid stretches (vatches), identical in viruses and man, and in other pathogens implicated in Alzheimer's disease (Borrelia, Chlamydia, Helicobacter, C.Neoformans, P.Gingivalis). C.Neoformans, which has been associated with a rare but curable form of dementia, expresses the most number of hits to Alzheimer's disease proteins. Vatches are often immunogenic and antibodies to viral proteins may knock down their human counterparts or activate immune responses killing the cells containing their human homologues. This is supported by the presence of the complement membrane attack complex in Alzheimer's disease neurones and by the ability of *tau* antigens (homologous to pathogen proteins) to promote the formation of neurofibrillary tangles and Alzheimer's disease pathology in mice. Vatches may act as dummy ligands or decoy receptors and interfere with the interactome of their human counterparts. Alzheimer's disease is thus a "pathogenetic" disorder caused by pathogens but dependent on the genes that create these matching sequences. This scenario is relevant to many other, and perhaps most human disorders, given the massive genomic extent of viral coverage. The vatches in the human proteome, dictated by polymorphisms and mutations, may predict, from birth, the spectrum of pathogens that match our proteins and which pathogenetic disease we are likely to develop. These may all be preventable by vaccination, pathogen detection and elimination and curable by immunosuppressant approaches, perhaps with a unique, safe, and effective immunosuppressant panacea.

Introduction

Alzheimer's disease is a devastating degenerative disorder characterised by extensive neuronal loss particularly of the cholinergic system⁶⁸, severance of the afferent and efferent hippocampal connections⁷³, loss of corticocortical glutamatergic association fibres⁷⁸ and massive cerebral shrinkage⁷². It is characterised pathologically by the presence of amyloid containing plaques⁵⁵ and neurofibrillary tangles³³ containing the hyperphosphorylated microtubule protein, *tau*⁴³. These features, including beta-amyloid deposition, tau phosphorylation^{35, 59, 75}, entorhinal and hippocampal cell loss and cerebral shrinkage⁴ can all be induced by Herpes simplex viral infection in mice. Several hundred genes have been implicated in the disease⁸ (see

<http://www.polygenicpathways.co.uk/alzpolys.html>

, and a number of environmental risk factors including viral infection (Herpes viruses HSV-1^{36,37} and the cytomegalovirus), and bacterial and parasitic infection (C.Pneumoniae^{17,49}, Borrelia Burgdorferi⁵³ and Helicobacter pylori) have also been shown to promote risk³¹. Indeed eradication of H.Pylori has been reported to improve cognitive performance⁴⁴. The AIDS virus HIV-1 can also provoke dementia with the classical pathology of Alzheimer's disease²⁴. Tooth loss, caused by periodontitis and the oral pathogen P.Gingivalis (*inter alia*), has also been linked to Alzheimer's disease⁴¹. Two case reports, for patients diagnosed with Alzheimer's disease, and in care as such for over three years, have shown that the misdiagnosed dementia was in fact a Cryptococcal infection caused by C.Neoformans. Both patients were treated with antifungal agents and returned to normal life^{2,30}. Previous work has shown that beta-amyloid and the mutant APP peptides are homologous to a number of viruses, bacteria, phages and allergens. Many of these homologous regions are immunogenic and it has been suggested that both familial and late-onset Alzheimer's disease are autoimmune disorders caused by viral mimicry of and antibody generation to beta-amyloid and mutant APP peptides¹³. This survey extended this observation to other pathogens and genes implicated in Alzheimer's disease. During the course of the survey, it became evident that viral DNA incorporation into the human genome is universal, suggesting a general application of the processes exemplified by Alzheimer's disease to most other diseases and syndromes.

Methods

The genomes of Herpes simplex (HSV-1), the human immunodeficiency virus (HIV-1) the cytomegalovirus, Rhinovirus 14, Chlamydia pneumoniae, Borrelia Burgdorferi, Helicobacter Pylori, Porphyromonas gingivalis and of Cryptococcus Neoformans were screened against the human proteome (translated nucleotide vs. protein: Blastx) and for HSV-1 *vice versa*. The raw results of these and other comparisons are stocked in an online database at PolygenicPathways <http://www.polygenicpathways.co.uk/blasts.htm> for visualisation and a summary available in the supplementary data file. The results in the subsequent figures are a summary of the hit tables for each BLAST. Alzheimer's disease susceptibility genes (any that have been reported as risk factors) are referenced at the PolygenicPathways database (<http://www.polygenicpathways.co.uk/alzpolys.htm>). The initial survey exceeded the cut off point for the BLAST results (>20,000 alignments) and only these results are shown³. Alignment with human chromosomes was realised via the BLAST-like alignment tool (BLAT)³⁹ server at Ensembl <http://www.ensembl.org/Multi/blastview>

Immunogenicity (B Cell epitope) was predicted using the BepiPred server⁴⁶
<http://www.cbs.dtu.dk/services/BepiPred>

It was not possible to BLAT whole viral genomes in either the UCSC or Ensembl applications, to map the homologies by human chromosome, as the input sequences are deemed too large. This could provide a detailed whole genome human map of the contiguous regions for any virus or vatch, which according to the results below, is likely to concord with human disease linkage data. With the computing power available, a concerted bioinformatic attack on these homologous regions should be able to rapidly identify all viral/human homologues and their associated diseases, both known and unsuspected.

Results

The type of gapped consensus homology between viral and human proteins is illustrated in Fig 1, together with an example of the insertion of a single viral match (“vatch”) into different regions of the human genome. These insertion sites correspond to both intra and intergenic regions (junk DNA), which appears to code for protein segments awaiting assembly and the addition of start and stop codons. Pictorial representations of the BLAST results comparing the HSV-1, HIV-1 and rhinovirus genomes with the human genome are shown in Fig 2, and numerous other results posted at <http://www.polygenicpathways.co.uk/blasts.htm>. The data in these website files contain almost half a million alignments from diverse viruses and other pathogens. To briefly summarise the overall data, the DNA of all viruses tested is inserted into the human genome. In the case of the HIV-1 and XMRV retroviruses, the insertion appears to be relatively recent, as the DNA is less fragmented, and the number of protein hits lower than that of other viruses. However the contiguous amino acid stretches are longer. For older insertions, for example HSV-1, the DNA fragmentation is greater, the contiguous amino acid stretches shorter, but the number of hits far greater. In addition, the repetitive patterns on the BLAST pictograms show that the viral DNA is reincorporated into multiple protein isoforms and into multiple members of the same gene families, for example the APP or APP binding protein families. These data show that the viral DNA is broken up by repeated chromosome shuffling over millions of years, but that contiguous stretches are retained. The effect is translated into short contiguous stretches of 5-12 amino acids, which are probably responsible for most, if not all, human diseases.

In relation to Alzheimer’s disease, the areas of viral insertion within the human genome (rhinovirus, HSV1, cytomegalovirus, HIV-1) correspond to genetic linkage hotspots in Alzheimer’s disease (157 genes in the linkage hotspot on chromosome 10 alone)⁴⁸ and to over 120 Alzheimer’s disease susceptibility genes. These genes included all the major players (APP and beta-amyloid, Presenilins 1 and 2, BACE1, clusterin, complement receptor 1, PICALM, glycogen synthase kinase-3 beta and tau) This viral/human overlap is effectively a viral linkage map for Alzheimer’s disease and others where the HSV-1 virus has been implicated.

These viruses mostly target the same genes (supplementary data) although there is a slight difference in their overall “pathogram” profile (Fig 3) and the regions within the genes are distinct. For example, the rhinovirus and HIV-1 are not homologous to complement receptor 1 or its homologue CR1L, nor to the APP binding protein APPBA2, fractalkine the phosphodiesterase ENTPD2 or teratocarcinoma-derived growth factor 1 (Fig 3).

All other genomes tested (*Borrelia*, *C. Pneumoniae*, *Helicobacter*, *C. Neoformans*) were homologous to over 120 Alzheimer's disease susceptibility gene products and to 157 genes within Alzheimer's disease linkage hotspots, most notably chromosome 10. The most relevant genes in Alzheimer's disease (with the exception of APOE4) were homologous to a greater number of pathogen proteins as shown in Figs 4 and 5. Heading the list were complement receptor 1, APP and 5 proteins related to mitochondrial respiration (NADH dehydrogenases, cytochrome B, ATP synthase, and the foetal Alzheimer's antigen. Tau, (MAPT), Presenilins 1 and 2, BACE1 and 2, clusterin and choline acetyltransferase were among the top 35 hits. and PICALM, APOE4 and APOE among the top 70. Mitochondrial energy related proteins figured prominently among the top scorers likely explaining the reduction in cerebral glucose consumption prior to symptomatology in late-onset Alzheimer's disease¹⁹. Conversely, *C. Neoformans* was the pathogen scoring the greatest number of hits in relation to Alzheimer's disease proteins, by an order of magnitude. Each of the 14 *C. Neoformans* chromosomes expresses proteins returning over 20,000 alignments, the cut-off point for the BLAST server (See website Tables).

APOE4 is homologous to these and other viral, bacterial and allergen vatches.

Surprisingly, APOE4, the most potent genetic risk factor¹⁸, was not at the head of the list of susceptibility genes in relation to its homology with pathogen proteins. It was therefore compared with viral, bacterial, yeast and allergen genomes to further elaborate its role. The APOE4 isoform is homologous to over 200 phages and viruses including salmonella, streptococcus, lactococcus, shigella and Enterobacteria phages as well as to herpes viruses (HSV 1,3,4 and 6) and to the measles, hepatitis and influenza viruses and HIV-1. APOE4 is also homologous to over 9,000 bacterial and 11,000 yeast proteins, 2574 from *Candida* species. Such extensive homology might thus explain the key role of APOE4 in this and other diseases^{6,9,45} (See website tables).

HIV-1/AIDS

The HSV-1 virus also expresses proteins homologous to our own, but fewer than observed for other viruses. However the homologous regions are much longer, suggesting a more recent event that has not undergone as much chromosomal shuffling (or that the virus has evolved independently to create matching sequences). For example no vatches longer than 8 amino acids were observed for HSV-1, queried versus Alzheimer's disease, while several of 10 and one of 11 amino acids long were observed in relation to the AIDS virus. As already reported, the AIDS virus expresses immunogenic proteins that are homologous to many proteins of the human immune defence network suggesting that AIDS too is an autoimmune disease¹². Its selective targeting of the immune system may explain many of the features of this disease, where rather counter intuitively, immunosuppressants may be of benefit.

Immunogenic proteins

In mice, immunisation with the tau peptide induces tauopathy, neurofibrillary tangles, gliosis and axonal loss, essentially reproducing the tau related pathology of

Alzheimer's disease⁶⁵ illustrating the mechanism through which immunogenic viral matches induce the pathological features of Alzheimer's disease (see discussion). Beta-amyloid autoantibodies derived from APP transgenic mice also potentiate the toxicity of the beta-amyloid peptide⁵⁴. A clinical trial with beta-amyloid antibodies also had to be halted, sadly resulting in the death of two patients from encephalitis²⁶.

Many of the pathogen vatches will be immunogenic and antibodies raised to infection are also likely to target their human counterparts, producing effects similar to those described above. The foetal Alzheimer antigen (Alz50) which labels tau-containing filaments⁵⁶, and is homologous to these pathogens (Fig 2) is an evident example. There are too many proteins to analyse individually, and beta-amyloid is used as an example, primarily because of the presence of beta-amyloid autoantibodies in the sera of the ageing population⁶⁹ and in Alzheimer's disease^{38,58}. As seen in Fig 6, the beta-amyloid peptide is studded with vatches from a large number of phages and viruses, many within the immunogenic regions of the peptide. Herpes simplex⁶⁰ and 68 viruses and phages¹³ are homologous to a VGGVV c-terminal sequence that has been used as an epitope to label beta-amyloid in the Alzheimer's disease brain⁶⁶. The peptide also contains matches to influenza, C. Tetani, C. Diphtheria and the poliovirus. Vaccination against these pathogens has been shown to reduce the risk of developing Alzheimer's disease⁷⁴.

While certain beta-amyloid autoantibodies may trigger a hazardous immune response, potentially destroying beta-amyloid containing cells, others possess catalytic properties and are less likely to provoke an immune response^{58,71}. These antibodies target the beta-amyloid peptide within the regions depicted in Fig 6. Interestingly three cancer-inducing viruses target these regions (Epstein-Barr, Hepatitis B and the papillomavirus). Antibodies raised to these viruses may catalytically dispose of the beta-amyloid peptide, perhaps explaining the inverse association between cancer and Alzheimer's disease^{63,64}. As a vaccine to the papillomavirus already exists²⁷, it may perhaps have a role to play in Alzheimer's disease prevention, provided that the epitopes are restricted to this catalytic region.

The autoimmune aspect is also applicable to familial Alzheimer's disease as four APP₇₁₇ and the Swedish mutants convert the surrounding peptides to ones homologous with proteins from the common cold virus, to the ubiquitous HHV-6 virus and to common commensal bacteria, including E.Coli and E.Faecalis¹³.

On the absence of viral DNA.

A retrospective analysis has shown that the presence of IgM antibodies to Herpes simples, a marker of active infection, predicted the subsequent development of Alzheimer's disease⁴⁷. However many studies have failed to demonstrate the presence of viral or pathogen DNA in Alzheimer's disease sera (a problem common to many viral-related diseases) e.g.^{61,62}. If however, the pathogen has already invoked an antibody response, that also targets its human homologues, the immune response will become self-sustaining as antibodies continue to encounter the human proteins, even if the virus has been successfully eliminated. Indeed, the more successful the immune response against the pathogen, the greater the risk of autoimmunity, a

disastrous pyrrhic victory. The absence of viral DNA therefore need not necessarily mitigate against the involvement of any particular pathogen.

Discussion.

The human genome is composed of millions of fragmented contiguous viral DNA sequences from diverse viruses. For older incorporations, the DNA is more fragmented, the amino acid matches (vatches) shorter, but the number of proteins involved much greater. For relatively recent insertions, for example the XMRV virus, the DNA is less fragmented, the vatches longer, and the number of hits fewer.

Almost a century ago J.B.S.Haldane and Francois D'Herelle suggested that viruses and phages are responsible for the origins of higher forms of life^{20,29}, a supposition supported by the present survey. Retroviruses also play a role in human evolution⁴⁰. At each retroviral insertion, our genomes change dramatically with the addition of several viral-derived genes, and if passed on through our germ lines, have effectively created a novel being. Over millions of years, evolution may thus proceed in jumps as dozens of new viral genes are added at each insertion. Some of these conferring disadvantage, for example the AIDS virus, will be weeded out by natural selection, while AIDS virus resistant genomes⁵⁷ will be maintained. Others may confer advantage. Those with little selection pressure will also be maintained, as is the case in Alzheimer's disease, where the genes have already been ceded to children and grandchildren before the onset of the disease. Viruses may thus be responsible not only for our existence but also for our progression through the evolutionary ladder. The design of new forms of life as already achieved²⁸ may therefore be fraught with hidden danger, or perhaps herald a new step in our evolution.

They are also responsible for many human diseases. Homologous viral DNA is translated into proteins containing short contiguous amino acid stretches (vatches) typically 5-12 amino acids long. These are identical to the vatches expressed by viruses, and other pathogens, and it is these that are the likely agents of destruction in many diseases. Their effects are reviewed in relation to Alzheimer's disease below.

Their relevance is applicable to many human diseases¹²⁻¹⁴, as shown in the website BLAST files. For example, the HIV-1 vatches target over 50 components of the immune defence network, while the XMRV virus implicated in chronic fatigue and prostate cancer⁵⁰ targets proteins related to fatigue (mitochondrial respiration genes) and prostate cancer. The Epstein-Barr virus implicated in multiple sclerosis⁵ targets multiple sclerosis autoantigens (and identifies others) and the coronavirus implicated in Parkinson's disease²⁵ targets Parkins and other proteins relevant to this disorder (dopamine receptors, transporters, free radical, glutathione and quinone-related proteins). The Herpes virus, influenza, rubella and T.Gondii, implicated in schizophrenia^{10,11,77} target DISC1, neuregulin dopamine, glutamate and oligodendrocyte related proteins implicated in this disease. Many viruses target proteins related to many diseases, For example the influenza virus targets proteins relevant to Alzheimer's disease, Bipolar disorder and schizophrenia. The Polyglutamine repeat protein responsible for Huntington's disease, and Spinocerebellar ataxias⁷ is also expressed by numerous phages, viruses, bacteria, fungi and yeast. The procedure can also be used to identify novel viral culprits in disease. For example BLASTS of multiple sclerosis autoantigens against viral proteomes revealed extensive homology with influenza, polyoma viruses, papillomaviruses, hepatitis B and C and shigella, vibrio and streptococcus phages, as

well as to the already implicated Epstein-Barr virus. A BLAST of a lung cancer oncogene NDUFC2 also revealed extensive homology with a number of viral species.

In relation to Alzheimer's disease, the viral chromosomal insertion points correspond to the locations of over 120 Alzheimer's disease susceptibility genes and to linkage hotspots. The pathogens implicated in Alzheimer's disease also express proteins homologous to these viral proteins. The importance of Alzheimer's disease susceptibility genes correlates with the extent of their matches to viral and other pathogen proteins. Conversely, the important pathogens will be those whose proteins target the greatest number of susceptibility gene products. The matches to C.Neoformans proteins were by far the greatest, by an order of magnitude. As there have been two case reports of complete recovery from dementia following eradication of this pathogen^{2,30}, this suggests that regular screening and elimination of fungal infection in the ageing population could have a dramatic effect on the incidence, severity and progression of Alzheimer's disease.

Many of the viral vatches in Alzheimer's disease are predicted to be immunogenic¹³. The dangers of such are perfectly illustrated by the report that immunisation in mice with the microtubule protein *tau* provokes the formation of neurofibrillary tangles and axonal injury, as seen in Alzheimer's disease⁶⁵. Antibodies directed to immunogenic viral vatches may thus also target their human homologues resulting in immune-related knockdown of the human protein, effectively creating, on a massive scale, the elements of behaviour and pathology seen in knockout transgenic mice^{1,15,21,42,51,67}. Vatches in beta-amyloid also correspond to those in numerous viruses, phages bacteria and allergens, and are highly immunogenic suggesting that the beta-amyloid autoantibodies seen in the ageing population may tag amyloid containing neurones for destruction. This is supported by the presence of numerous immune and inflammatory-related proteins in Alzheimer's disease amyloid plaques^{52,70}, by evidence of activated microglia in the areas surrounding these plaques²³ and by the presence of the complement membrane attack complex in neurofibrillary tangles and neurones³⁴ ..

Many vatches are homologous to peptide ligands including BDNF, fractalkine and interleukin 1-alpha, *inter alia*, and might therefore act as dummy ligands, either blocking, stimulating or otherwise using the cognate receptors of their human counterparts, as exemplified by the use of the chemokine CCR5 receptor by the AIDS virus²². Others are homologous to receptors including the muscarinic and nicotinic receptors, CHRM3 and CHRNA4 and multiple lipoprotein receptors, and may thus act as decoys for their ligands effectively acting as antagonists. They are also homologous to proteins within the same signalling network, for example APP, and the Fe65 class of APP binding proteins as well as BACE and gamma secretase components and are likely to interfere with interactomes of their human counterparts. Viral vatches are also homologous to ion channels, growth factors, cytokines, transcription factors, enzymes, immune components, mitochondrial components and others and antibodies to these proteins are capable of creating havoc in a wide variety of biological systems.

Nevertheless, Alzheimer's disease is a disease of old-age, and despite the many pathogens homologous to these key proteins, sufferers have evidently escaped the potential hazards until late in life. The generation of antibodies is a random process dependent upon the antigens encountered, and the immune armoury set up at birth, and it is possible that summative effects, for example multiple infections, are necessary to trigger the final end game. In this respect, the risk of developing Alzheimer's disease is related to the number of pregnancies¹⁶, and therefore to multiple exposure to childhood infections. The severity of Alzheimer's disease is also

attenuated in nuns ³², who are less exposed to sexually transmitted viral and other pathogen diseases. With repeated exposure, the antibody bank and diversity will accumulate and increase the risks of autoimmunity. It is possible that there is one final trigger creating saturation of the antibody response and it is tempting to suggest that this is related to C.Neoformans, whose homology to Alzheimer's disease related proteins was an order of magnitude higher than that of any other pathogen, representing a superantigen that may well play a critical role in Alzheimer's disease.

The results discussed above suggest that Alzheimer's disease is a "pathogenetic" disorder caused by pathogens in a gene-dependent manner. Pathogen detection and elimination are realistic therapeutic options, as already suggested ^{75,76}. In the long-term, vaccination against these viruses might be expected to reduce the incidence of Alzheimer's disease, as already shown for vaccination against influenza, diphtheria, tetanus and polio ⁷⁴. The autoimmune component also suggests that immunosuppression, anti-antibody antibodies or immunadsorption may be of benefit.

Our genomes are composed of millions of fragmented viral DNA segments, reflecting our evolution from these organisms and from many retroviral insertions over millions of years. While responsible for both our origin and evolution, viruses have left behind a deadly legacy of viral-derived human protein fragments that exactly match those in the current virome. In these pathogenetic diseases, our genomes, mutations and polymorphisms dictate which vatches we possess, and from birth, select the viruses we match and the diseases we might develop. This applies to viruses, bacteria, fungi, parasites and allergens, and indeed to all pathogenic species between viruses and Man in the tree of life. This inevitability is by no means pessimistic. *Given the phylogeny of the thousands of microbial genomes so far sequenced, it is inconceivable that other viruses and their descendants do not follow the same pattern, suggesting that most diseases are pathogenetic, with viruses and other pathogens causing their mischief via the same mechanisms, including vatch-triggered autoimmunity. This infers that most are therefore preventable and curable, perhaps, in relation to autoimmunity, by the same single, safe and effective immunosuppressant; in other words, the panacea.*

Fig 1: Examples of the types of gapped consensus sequence homologies observed between viral and human proteomes and a representation of the insertion of a single translated HSV-1 octapeptide into different regions of the human genome.

Fig 2: Example of the viral matches to the human proteome (HIV-1, HSV-1 and rhinovirus 14) of the human matches in the Herpes simplex (HSV-1) proteome. Note the repetitive patterns in the human proteome with respect to specific queries. These represent protein isoforms or protein families, for example the APP or APP binding protein family.

Fig 3 A pathogram of Alzheimer's disease showing selected examples of the similarities and differences between viral targeting of the human proteome. The table below indicates the number of viral proteins matching the human target for each virus (HSV-1, HIV-1 and rhinovirus 14). Note that the rhinovirus or the AIDS virus are not homologous to complement receptor 1, the APP binding protein APPBA2, fractalkine, CX3CL1, the phosphodiesterase ENTPD2 or the teratoma derived growth factor TDGF1.

Fig 4 and 5

Human proteins ranked in relation to the number of pathogen protein matches from HSV-1, HIV-1, cytomegalovirus, C.Pneumoniae, rhinovirus 14, Borrelia Burgdorferi, Helicobacter pylori, P.Gingivalis, and the 14 chromosomes of C.Neoformans. Gene products highly relevant to Alzheimer's disease are highlighted in red. 4) Complement receptor 1 (CR1) amyloid precursors protein APP, foetal Alzheimer antigen (Falz), tau (MAPT), clusterin (CLU), presenilins (PSEN1, PSEN2) beta-secretase (BACE1, 2) 5) Apolipoprotein E and E4 (APOE) and phosphatidylinositol-binding clathrin assembly protein (PICALM). See supplementary data and website for other abbreviations

Fig 6

Viral protein matches to different regions of the beta-amyloid peptide in relation to the predicted B Cell Epitope antigenicity. Predicted epitopes are marked with an asterisk and the Y-axis is an index of antigenicity. The VGGVV epitope has been used to label beta-amyloid and is marked + as are other short epitopes used to label beta-amyloid (QKLV, FFAE, IIGL) Other short epitopes used to label beta-amyloid include MGGVV, VGGVV, MVGGVV, VGGVV and GGVVIA). The viruses and bacteria in black boxes represent those where vaccination was reported to reduce the incidence of Alzheimer's disease. These alignments are to viral proteins rather than to epitopes within the vaccine. The arrows represent the beta-amyloid cleavage sites of the catalytic beta-amyloid autoantibodies isolated from Alzheimer's disease sera. H*↑Q is the major site of catalysis. Note that cancer and plant viruses overlap this region. Burk= Burkholderia Phage; Camp Phage = Campylobacter Phage; HumRespSyn = Human respiratory syncytial Virus; HepC = Hepatitis C; HIV = Human Immunodeficiency Virus; HSV-1 = Herpes simplex (Human Herpesvirus 1);

Mycobact = Mycobacteria phage; Strep = Streptococcus phage; C.Tet = Clostridium Tetani, C.Dip = Corynebacterium diphtheriae (See website for alignments)

Reference List

- ¹ A. Aguzzi, *et al.*, "Transgenic and knock-out mice: models of neurological disease," *Brain Pathol.* **4**(1), 3 (1994).
- ² T. A. Ala, R. C. Doss, and C. J. Sullivan, "Reversible dementia: a case of cryptococcal meningitis masquerading as Alzheimer's disease," *J. Alzheimers. Dis.* **6**(5), 503 (2004).
- ³ S. F. Altschul, *et al.*, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Res.* **25**(17), 3389 (1997).
- ⁴ A. G. Armien, *et al.*, "Chronic Cortical and Subcortical Pathology with Associated Neurological Deficits Ensuing Experimental Herpes Encephalitis," *Brain Pathol.* (2009).
- ⁵ B. A. Bagert, "Epstein-Barr virus in multiple sclerosis," *Curr. Neurol. Neurosci. Rep.* **9**(5), 405 (2009).
- ⁶ G. Bartzokis, *et al.*, "Apolipoprotein E genotype and age-related myelin breakdown in healthy individuals: implications for cognitive decline and dementia," *Arch. Gen. Psychiatry.* **63**(1), 63 (2006).
- ⁷ P. O. Bauer and N. Nukina, "The pathogenic mechanisms of polyglutamine diseases and current therapeutic strategies," *J. Neurochem.* **110**(6), 1737 (2009).
- ⁸ L. Bertram, *et al.*, "Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database," *Nat. Genet.* **39**(1), 17 (2007).
- ⁹ P. S. Bhattacharjee, *et al.*, "Effect of human apolipoprotein E genotype on the pathogenesis of experimental ocular HSV-1," *Exp. Eye Res.* **87**(2), 122 (2008).
- ¹⁰ S. L. Buka, *et al.*, "Maternal exposure to herpes simplex virus and risk of psychosis among adult offspring," *Biol Psychiatry.* **63**(8), 809 (2008).
- ¹¹ C. J. Carter, "Schizophrenia susceptibility genes directly implicated in the life cycles of pathogens: cytomegalovirus, influenza, herpes simplex, rubella, and *Toxoplasma gondii*," *Schizophr. Bull.* **35**(6), 1163 (2009).
- ¹² C. J. Carter, "Extensive Viral Mimicry of Human Proteins in AIDS, Autoimmune Disorders, Late-Onset and Familial Alzheimer's Disease and Other Genetic Diseases," in 2010), pp.1-15.
- ¹³ C. J. Carter, "Familial and Late-Onset Alzheimer's Disease: Autoimmune Disorders Triggered by Viral, Microbial and Allergen Mimics of Beta-Amyloid and APP Mutants ?," in 2010).

- 14 C. J. Carter, "Proteins of the XMRV Retrovirus Implicated in Chronic Fatigue Syndrome and Prostate Cancer Are Homologous to Human Proteins Relevant to Both Diseases.," in 2010).
- 15 A. M. Cataldo, *et al.*, "Presenilin mutations in familial Alzheimer disease and transgenic mouse models accelerate neuronal lysosomal pathology," *J. Neuropathol. Exp. Neurol.* **63**(8), 821 (2004).
- 16 M. Colucci, *et al.*, "The number of pregnancies is a risk factor for Alzheimer's disease," *Eur. J. Neurol.* **13**(12), 1374 (2006).
- 17 C. Contini, *et al.*, "Chlamydomonas pneumoniae Infection and Its Role in Neurological Disorders," *Interdiscip. Perspect. Infect. Dis.* **2010**, 273573 (2010).
- 18 E. H. Corder, *et al.*, "Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families," *Science* **261**(5123), 921 (1993).
- 19 N. R. Cutler, "Cerebral metabolism as measured with positron emission tomography (PET) and [18F] 2-deoxy-D-glucose: healthy aging, Alzheimer's disease and Down syndrome," *Prog. Neuropsychopharmacol. Biol Psychiatry.* **10**(3-5), 309 (1986).
- 20 F D'Herelle, *the bacteriophage; its role in immunity* (Williams and Wilkins, Baltimore, 1922).
- 21 S. E. Dodson, *et al.*, "Loss of LR11/SORLA enhances early pathology in a mouse model of amyloidosis: evidence for a proximal role in Alzheimer's disease," *J. Neurosci.* **28**(48), 12877 (2008).
- 22 T. Dragic, *et al.*, "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5," *Nature.* **381**(6584), 667 (1996).
- 23 P. Eikelenboom, *et al.*, "Neuroinflammation - an early event in both the history and pathogenesis of Alzheimer's disease," *Neurodegener. Dis.* **7**(1-3), 38 (2010).
- 24 M. M. Esiri, S. C. Biddolph, and C. S. Morris, "Prevalence of Alzheimer plaques in AIDS," *J. Neurol. Neurosurg. Psychiatry.* **65**(1), 29 (1998).
- 25 E. Fazzini, J. Fleming, and S. Fahn, "Cerebrospinal fluid antibodies to coronavirus in patients with Parkinson's disease," *Mov Disord.* **7**(2), 153 (1992).
- 26 I. Ferrer, *et al.*, "Neuropathology and pathogenesis of encephalitis following amyloid-beta immunization in Alzheimer's disease," *Brain Pathol.* **14**(1), 11 (2004).
- 27 S. M. Garland and J. S. Smith, "Human papillomavirus vaccines: current status and future prospects," *Drugs.* **70**(9), 1079 (2010).
- 28 D. G. Gibson, *et al.*, "Creation of a bacterial cell controlled by a chemically synthesized genome," *Science.* **329**(5987), 52 (2010).

- 29 J. B. S. Haldane, "The Origin of Life," in 148 ed. 1928), pp.3-10.
- 30 M. Hoffmann, *et al.*, "Cryptococcal meningitis misdiagnosed as Alzheimer's disease: complete neurological and cognitive recovery with treatment," *J. Alzheimers. Dis.* **16**(3), 517 (2009).
- 31 K. Honjo, R. van Reekum, and N. P. Verhoeff, "Alzheimer's disease and infection: do infectious agents contribute to progression of Alzheimer's disease?," *Alzheimers. Dement.* **5**(4), 348 (2009).
- 32 D. Iacono, *et al.*, "The Nun study: clinically silent AD, neuronal hypertrophy, and linguistic skills in early life," *Neurology.* **73**(9), 665 (2009).
- 33 H. Ishino and S. Otsuki, "Distribution of Alzheimer's neurofibrillary tangles in the basal ganglia and brain stem of progressive supranuclear palsy and Alzheimer's disease," *Folia Psychiatr. Neurol. Jpn.* **29**(2), 179 (1975).
- 34 S. Itagaki, *et al.*, "Ultrastructural localization of complement membrane attack complex (MAC)-like immunoreactivity in brains of patients with Alzheimer's disease," *Brain Res.* **645**(1-2), 78 (1994).
- 35 R. F. Itzhaki, S. L. Cosby, and M. A. Wozniak, "Herpes simplex virus type 1 and Alzheimer's disease: the autophagy connection," *J. Neurovirol.* **14**(1), 1 (2008).
- 36 R. F. Itzhaki and M. A. Wozniak, "Herpes simplex virus type 1 in Alzheimer's disease: the enemy within," *J. Alzheimers. Dis.* **13**(4), 393 (2008).
- 37 R. F. Itzhaki, *et al.*, "Infiltration of the brain by pathogens causes Alzheimer's disease," *Neurobiol. Aging* **25**(5), 619 (2004).
- 38 A. Kellner, *et al.*, "Autoantibodies against beta-amyloid are common in Alzheimer's disease and help control plaque burden," *Ann. Neurol.* **65**(1), 24 (2009).
- 39 W. J. Kent, "BLAT--the BLAST-like alignment tool," *Genome Res.* **12**(4), 656 (2002).
- 40 K. Khodosevich, Y. Lebedev, and E. Sverdlov, "Endogenous retroviruses and human evolution," *Comp Funct. Genomics.* **3**(6), 494 (2002).
- 41 J. M. Kim, *et al.*, "Dental health, nutritional status and recent-onset dementia in a Korean community population," *Int. J. Geriatr. Psychiatry.* **22**(9), 850 (2007).
- 42 R. Koldamova, M. Staufienbiel, and I. Lefterov, "Lack of ABCA1 considerably decreases brain ApoE level and increases amyloid deposition in APP23 mice," *J. Biol Chem.* **280**(52), 43224 (2005).
- 43 K. S. Kosik, "Tau protein and neurodegeneration," *Mol. Neurobiol.* **4**(3-4), 171 (1990).

- 44 J. Kountouras, *et al.*, "Eradication of *Helicobacter pylori* may be beneficial in the management of Alzheimer's disease," *J. Neurol.* **256**(5), 758 (2009).
- 45 I. Kuhlmann, *et al.*, "Apolipoprotein E genotype and hepatitis C, HIV and herpes simplex disease risk: a literature review," *Lipids Health Dis.* **9**, 8 (2010).
- 46 J. E. Larsen, O. Lund, and M. Nielsen, "Improved method for predicting linear B-cell epitopes," *Immunome. Res.* **2**, 2 (2006).
- 47 L. Letenneur, *et al.*, "Seropositivity to herpes simplex virus antibodies and risk of Alzheimer's disease: a population-based cohort study," *PLoS. One.* **3**(11), e3637 (2008).
- 48 X. Liang, *et al.*, "Genomic convergence to identify candidate genes for Alzheimer disease on chromosome 10," *Hum. Mutat.* **30**(3), 463 (2009).
- 49 C. S. Little, *et al.*, "Chlamydia pneumoniae induces Alzheimer-like amyloid plaques in brains of BALB/c mice," *Neurobiol. Aging.* **25**(4), 419 (2004).
- 50 V. C. Lombardi, *et al.*, "Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome," *Science.* **326**(5952), 585 (2009).
- 51 M. Maier, *et al.*, "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**(25), 6333 (2008).
- 52 P. L. McGeer, *et al.*, "Pathological proteins in senile plaques," *Tohoku J. Exp. Med.* **174**(3), 269 (1994).
- 53 J. Miklossy, *et al.*, "Borrelia burgdorferi persists in the brain in chronic lyme neuroborreliosis and may be associated with Alzheimer disease," *J. Alzheimers. Dis.* **6**(6), 639 (2004).
- 54 A. Nath, *et al.*, "Autoantibodies to amyloid beta-peptide (Abeta) are increased in Alzheimer's disease patients and Abeta antibodies can enhance Abeta neurotoxicity: implications for disease pathogenesis and vaccine development," *Neuromolecular. Med.* **3**(1), 29 (2003).
- 55 T. Nikaïdo, *et al.*, "Isolation and preliminary characterization of Alzheimer plaques from presenile and senile dementia," *Trans. Am. Neurol. Assoc.* **95**, 47 (1970).
- 56 N. Nukina, "The reinterpretation of the immunochemical study of Alzheimer neurofibrillary tangles," *Ann. Med.* **21**(2), 117 (1989).
- 57 S. J. O'Brien and J. P. Moore, "The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS," *Immunol. Rev.* **177**, 99 (2000).

- 58 S. Paul, S. Planque, and Y. Nishiyama, "Immunological Origin and Functional Properties of Catalytic Autoantibodies to Amyloid beta Peptide," *J. Clin. Immunol.* (2010).
- 59 R. Piacentini, *et al.*, "HSV-1 promotes Ca(2+)-mediated APP phosphorylation and Abeta accumulation in rat cortical neurons," *Neurobiol. Aging.* (2010).
- 60 R. B. Pyles, "The association of herpes simplex virus and Alzheimer's disease: a potential synthesis of genetic and environmental factors," *Herpes.* **8**(3), 64 (2001).
- 61 E. B. Renvoize, I. O. Awad, and M. H. Hambling, "A sero-epidemiological study of conventional infectious agents in Alzheimer's disease," *Age Ageing.* **16**(5), 311 (1987).
- 62 E. B. Renvoize and M. H. Hambling, "Cytomegalovirus infection and Alzheimer's disease," *Age Ageing.* **13**(4), 205 (1984).
- 63 C. M. Roe, *et al.*, "Alzheimer disease and cancer," *Neurology.* **64**(5), 895 (2005).
- 64 C. M. Roe, *et al.*, "Cancer linked to Alzheimer disease but not vascular dementia," *Neurology.* **74**(2), 106 (2010).
- 65 H. Rosenmann, *et al.*, "Tauopathy-like abnormalities and neurologic deficits in mice immunized with neuronal tau protein," *Arch. Neurol.* **63**(10), 1459 (2006).
- 66 C. Schwab, *et al.*, "Extracellular neurofibrillary tangles are immunopositive for the 40 carboxy-terminal sequence of beta-amyloid protein," *J. Neuropathol. Exp. Neurol.* **57**(12), 1131 (1998).
- 67 G. R. Seabrook and T. W. Rosahl, "Transgenic animals relevant to Alzheimer's disease," *Neuropharmacology.* **38**(1), 1 (1999).
- 68 N. R. Sims, *et al.*, "Glucose metabolism and acetylcholine synthesis in relation to neuronal activity in Alzheimer's disease," *Lancet.* **1**(8164), 333 (1980).
- 69 J. H. Sohn, *et al.*, "Identification of autoantibody against beta-amyloid peptide in the serum of elderly," *Front Biosci.* **14**, 3879 (2009).
- 70 S. D. Styren, W. H. Civin, and J. Rogers, "Molecular, cellular, and pathologic characterization of HLA-DR immunoreactivity in normal elderly and Alzheimer's disease brain," *Exp. Neurol.* **110**(1), 93 (1990).
- 71 H. Taguchi, *et al.*, "Exceptional amyloid beta peptide hydrolyzing activity of nonphysiological immunoglobulin variable domain scaffolds," *J. Biol Chem.* **283**(52), 36724 (2008).
- 72 P. M. Thompson, *et al.*, "Dynamics of gray matter loss in Alzheimer's disease," *J. Neurosci.* **23**(3), 994 (2003).

- ⁷³ G. W. Van Hoesen and B. T. Hyman, "Hippocampal formation: anatomy and the patterns of pathology in Alzheimer's disease," *Prog. Brain Res.* **83**, 445 (1990).
- ⁷⁴ R. Verreault, *et al.*, "Past exposure to vaccines and subsequent risk of Alzheimer's disease," *CMAJ*. **165**(11), 1495 (2001).
- ⁷⁵ M. A. Wozniak, A. L. Frost, and R. F. Itzhaki, "Alzheimer's disease-specific tau phosphorylation is induced by herpes simplex virus type 1," *J. Alzheimers. Dis.* **16** (2), 341 (2009).
- ⁷⁶ M. A. Wozniak, A. P. Mee, and R. F. Itzhaki, "Herpes simplex virus type 1 DNA is located within Alzheimer's disease amyloid plaques," *J. Pathol.* **217**(1), 131 (2009).
- ⁷⁷ R. H. Yolken and E. F. Torrey, "Are some cases of psychosis caused by microbial agents? A review of the evidence," *Mol. Psychiatry*. **13**(5), 470 (2008).
- ⁷⁸ A. B. Young, "Cortical amino acidergic pathways in Alzheimer's disease," *J. Neural Transm. Suppl.* **24**, 147 (1987).

Fig 1: Examples of the gapped contiguous viral sequences (Sbjct) within the human proteome (Query: Translated viral genome) and the insertion of a single HSV-1 vatch (Translated protein) into different chromosomal regions (From the BLAT Ensembl server).

```
Query 136083 QLLSAVDYIHRQGIHRDIKTENIFINTPE----DICLGDFGAA-----CFV 136211
              QL      YIH  GIHRDIK  +N  +P+      C  DFG  A      C
Sbjct 164    QLFRLAYIHSFGICHRIKPKQNLLLD-PDTAVLKLC--DFGSAKQLVRGEPNVSYIC-- 218

Query 139575 RPRTSP      CFTR  YPPLMGAGSGLRGWDGT  LTPYKASLEGKGQSISR*RG 139724
              RPRT P      C  R  Y  P      T  LTP  LE
Sbjct 306    RPRTPEAIALC  SRLLEYTP      TARLTP  LE----- 333

Query 12093  RKRPERHGRHFDRLDRGHHHLHHCHNHLRV 12004
              RKRPR+      L  GH  LHHC  LR+
Sbjct 75     RKRPPRQ      KPL  GHSLHHCSRI
```

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Beta-Amyloid 1-42