

RESEARCH ARTICLE

# DES-Amyloidoses “Amyloidoses through the looking-glass”: A knowledgebase developed for exploring and linking information related to human amyloid-related diseases

Vladan P. Bajic<sup>1†\*</sup>, Adil Salhi<sup>2</sup>, Katja Lakota<sup>3</sup>, Aleksandar Radovanovic<sup>2</sup>, Rozaimi Razali<sup>2</sup>, Lada Zivkovic<sup>3</sup>, Biljana Spremo-Potparevic<sup>4</sup>, Mahmut Uludag<sup>2</sup>, Faroug Tifratene<sup>2</sup>, Olaa Motwalli<sup>5</sup>, Benoit Marchand<sup>6</sup>, Vladimir B. Bajic<sup>2</sup>, Takashi Gojobori<sup>2,7</sup>, Esmā R. Isenovic<sup>1</sup>, Magbubah Essack<sup>2†\*</sup>

**1** Institute of Nuclear Sciences “VINCA”, Laboratory for Radiobiology and Molecular Genetics, University of Belgrade, Belgrade, Republic of Serbia, **2** Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology (KAUST), Thuwal, Kingdom of Saudi Arabia, **3** Department of Physiology, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia, **4** Department of Pathobiology, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia, **5** Saudi Electronic University (SEU), College of Computing and Informatics, Madinah, Kingdom of Saudi Arabia, **6** New York University, Abu Dhabi, UAE, **7** Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Kingdom of Saudi Arabia

† VPB and ME share first authorship on this work.

\* [magbubah.essack@kaust.edu.sa](mailto:magbubah.essack@kaust.edu.sa) (ME); [vladanbajic@vinca.rs](mailto:vladanbajic@vinca.rs) (VPB)

## Abstract

More than 30 types of amyloids are linked to close to 50 diseases in humans, the most prominent being Alzheimer’s disease (AD). AD is brain-related local amyloidosis, while another amyloidosis, such as AA amyloidosis, tends to be more systemic. Therefore, we need to know more about the biological entities’ influencing these amyloidosis processes. However, there is currently no support system developed specifically to handle this extraordinarily complex and demanding task. To acquire a systematic view of amyloidosis and how this may be relevant to the brain and other organs, we needed a means to explore “amyloid network systems” that may underly processes that leads to an amyloid-related disease. In this regard, we developed the DES-Amyloidoses knowledgebase (KB) to obtain fast and relevant information regarding the biological network related to amyloid proteins/peptides and amyloid-related diseases. This KB contains information obtained through text and data mining of available scientific literature and other public repositories. The information compiled into the DES-Amyloidoses system based on 19 topic-specific dictionaries resulted in 796,409 associations between terms from these dictionaries. Users can explore this information through various options, including enriched concepts, enriched pairs, and semantic similarity. We show the usefulness of the KB using an example focused on inflammasome-amyloid associations. To our knowledge, this is the only KB dedicated to human amyloid-related diseases derived primarily through literature text mining and complemented by data mining that provides a novel way of exploring information relevant to amyloidoses.



## OPEN ACCESS

**Citation:** Bajic VP, Salhi A, Lakota K, Radovanovic A, Razali R, Zivkovic L, et al. (2022) DES-Amyloidoses “Amyloidoses through the looking-glass”: A knowledgebase developed for exploring and linking information related to human amyloid-related diseases. PLoS ONE 17(7): e0271737. <https://doi.org/10.1371/journal.pone.0271737>

**Editor:** Maria Grazia Ortole, Marche Polytechnic University, ITALY

**Received:** August 17, 2021

**Accepted:** July 6, 2022

**Published:** July 25, 2022

**Copyright:** © 2022 Bajic et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data sources used for creating the concepts index are detailed in the manuscript, and are accessible to anyone for download and to use for the mining of PubMed. The authors also confirm additional available data (<https://www.cbrc.kaust.edu.sa/des-amyloidosis/>) is third party data and not owned by them.

**Funding:** This work has been supported by the Ministry of Education, Science and Technological Development, Republic of Serbia, and by the

KAUST grant OSR#4129 (to EI and VBB), which also supported SZ and VPB. VBB has been supported by the KAUST Base Research Fund (BAS/1/1606-01-01), while ME has been supported by KAUST Office of Sponsored Research (OSR) grant no. FCC/1/1976-20-01. TG has been supported by the King Abdullah University of Science and Technology (KAUST) Base Research Fund (BAS/1/1059-01-01). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

## 1. Introduction

Amyloid refers to aberrant extracellular proteins that clump together, forming fibrils [1, 2]. The formation of these fibrillar assemblies causes the native secondary structure of proteins to change its shape to predominantly cross- $\beta$ -sheet secondary structures essential for fiber formation [3–5]. Unlike normal fibrous proteins that provide structural support in cells, these amyloid types (protein aggregates) are associated with the pathology of almost 50 disorders with different symptoms collectively referred to as amyloidoses [6–10]. Rambaran *et al.* [11] and Sipe *et al.* [12] provide inventories of amyloids associated with these human diseases, and these works show that the number of known amyloids increased from 20 in 2008 to 31 in 2016. Amyloids display diversity in structure [13], aggregation time, and site of action [14]. Thus, today's amyloidosis represents all diseases with misfolded aggregated proteins or peptides as the common denominator.

Amyloidoses have been sub-classified into prion diseases and non-prion diseases; prions (misfolded proteins) become a self-perpetuating infectious agent in prion disease, i.e., they are transmissible, whereas misfolded proteins in non-prion diseases are non-transmissible [15]. Thus far, all amyloids do not have demonstrated infectivity as prions. However, a preformed fibril can seed amyloid formation. That is, in cell culture, seeding aggregation has been demonstrated for non-prion disease amyloids such as Tau [16–18],  $\alpha$ -synuclein [19–21], amyloid-beta (A $\beta$ ) [22–24], huntingtin [25, 26], superoxide dismutase 1 (SOD1) [27, 28] or TDP-43 [29, 30]. Moreover, Tau [16, 20, 31–34],  $\alpha$ -synuclein [19, 35, 36], and A $\beta$  [23, 37] further exhibit trans-cellular propagation and the ability to induce progressive pathology *in vivo*. Our understanding of these amyloids' roles is further complicated by experimental results showing that amyloids can interact to aggregate into hybrid amyloid fibrils (a process called cross-seeding) [38]. Reports of such co-localized amyloids include 1/ A $\beta$  and human islet amyloid polypeptide (hIAPP) in the brain and beta-pancreatic cells [39, 40], 2/ A $\beta$  and SAA in AD plaques [41], and 3/ A $\beta$  and phosphorylated tau (p-tau) in synaptic terminals of AD brains [42].

To expand our insights into amyloids mechanisms of action and roles, several methods/tools have been developed to predict the propensity of proteins to aggregate, examples being Tango [43], ZipperDB [44, 45], Pasta [46], NetCSSP [47], FoldAmyloid [48], AmyloidMutant [49, 50], AmylPred2 [51]. When developing such tools/methods, the key task is to find approaches for discovering sequence segments responsible for self-aggregation and protein destabilization [44, 52–54]. Other databases facilitate in-depth investigation of these amyloids due to their link to several disease states. Representative databases include AMYPdb, (Amyloid Protein Database) [55], CPAD (Curated Protein Aggregation Database) [56], and ALBase (Amyloid Light Chain Database) [57] [<http://albase.bumc.bu.edu/aldb/>]. Specifically, the AmyPDB database houses amyloid protein data from 33 amyloid families, including 1,705 proteins, complemented with bibliographic references and 3D structures. While CPAD provides a collection of more than 2300 experimentally observed aggregation rates for known amyloids that can be accessed based on various classifications. This data is further linked to Uniprot [58], Protein Data Bank [59], PubMed [60], GAP [61], TANGO [62] and WALTZ [63]. On the other hand, ALBase contains 4364 amyloid nucleotide sequences, of which 808 encode monoclonal proteins that form fibrillar deposits in patients with light chain, amyloidosis, including 295 control light chain sequences from healthy subjects used to analyze the amyloid sequences. Overall, the existing amyloid-related databases' scope is restrictive and does not allow for the comprehensive exploration of literature information and all biomedical terms/concepts related to amyloids.

Here we present DES-Amyloidoses (<https://www.cbrc.kaust.edu.sa/des-amyloidosis/>), a text and data mining-based KB developed to facilitate more efficient exploration of information

contained in the literature and link concepts related to human amyloidosis. The development strategy is similar to the one we presented in [70]. Specifically, for text indexing, we used pre-compiled biomedical terms/phrases (referred to as concepts) organized into thematic dictionaries (e.g., diseases, pathways, miRNAs, lncRNAs, and so forth). The KB searches for the dictionary terms in titles, abstracts, and full-length articles retrieved from PubMed and PubMed Central [37]. We incorporated this data into the DES framework designed to provide users with the statistically enriched concepts in the topic-specific literature (in this case, about the amyloids and associated diseases), statistically “enriched pairs” of concepts or concepts that co-occur in text, as well as semantic similarity, for further exploration. We also provide an AD-related example to show how DES-Amyloidoses can assist researchers in the amyloid domain.

## 2. The DES-Amyloidoses exploration system

Efficient exploration of information related to amyloids is challenging as published information is substantial. For example, a PubMed Central query (31 December 2019) using “amyloid” retrieves more than 125,000 articles, of which 50% were published in the last five years. The challenge associated with exploring this voluminous amyloid-related information is further augmented when searching for links between different amyloids or between amyloids and other relevant biomedical concepts. However, the literature exploration process of complex biomedical topics of this nature has been made easy by developing several topic-specific KBs [64–75]. These KBs provide users with topic-specific enriched concepts and pre-computed statistically enriched associations between the enriched concepts. The text mining in these KBs are based on the titles and abstracts of publicly available PubMed records [32]. However, we can access significantly more information in full length articles [33]. Thus, the more recently developed topic-specific KBs use text mining based on both titles and abstracts (PubMed [60]) and full length articles (the subset of open-access articles from PubMed Central [37]). Unfortunately, to our knowledge, an amyloidoses-related KB of this type does not exist.

Thus, we developed DES-Amyloidoses as a topic-specific KB using an upgraded version of the Dragon Exploration System (DES), DES v3.0, on 17 February 2020. This version of DES allows users to explore scientific concepts through literature-derived topic-specific enriched concepts and pairs of concepts similar to the older versions, and it enables users to explore these concepts through semantic similarity. The underlying systems, and concept enrichment process used in the current version of DES has been described in [74].

### 2.1 The DES-Amyloidoses literature corpus

To create the DES-Amyloidoses literature corpus, on the 17 February 2020, we used the following query: ((*Amyloid* AND (“*Serum Amyloid A*” OR *SAA* OR “*amyloid A*” OR *AA* OR *apoSAA* OR *ApoE* OR “*light chain amyloid*” OR *AL* OR “*amyloid enhancing factor*” OR *AEF* OR “*Serum amyloid P component*” OR *SAP* OR *glycosaminoglycans* OR “*Heparin sulfate proteoglycan*” OR “*islet amyloid polypeptide*” OR *IAPP* OR “*Immunoglobulin light chain*” OR “*Immunoglobulin heavy chain*” OR *AH* OR *Transytherthetin* OR *ATTR* OR “ $\beta$ 2-microglobulin” OR *A $\beta$ 2M* OR “*Apolipoprotein AI*” OR “*AApo AI*” OR *AApoAII* OR *AApoAIII* OR *Gelsolin* OR *Agel* OR *Lysosyme* OR *Alys* OR *Fibrinogen $\alpha$*  OR *Afib* OR “*Cystatin C*” OR *ACys* OR “*ABriPP variants*” OR *ABri* OR “ $\alpha$ -Synuclein” OR *A $\alpha$ Syn* OR *Tau* OR *ATau* OR “*Prion protein*” OR *APrP* OR “*Atrial natriuretic factor*” OR *AANF* OR *Prolactin* OR “*A Pro*” OR *Insulin* OR *AIns* OR “*Galectin 7*” OR *AGal7* OR *Lactoferrin* OR *ALac* OR “*Semenogelin 1*” OR *Asem1* OR *Enfurvitide* OR *AEnf*)) AND (*human* OR *humans* OR “*homo sapiens*”), to retrieve topic-specific articles from our local repository (MongoDB) of PubMed, and PubMed Central articles. This query retrieved 31,821 articles.

## 2.2 Dictionaries incorporated into DES-Amyloidoses

We imported 18 dictionaries from the pre-existing DES v2.0 vocabularies. To ensure completeness, we further compiled the topic-related dictionary, “Amyloids (Human and Mouse)” (see Table 1). The terms compiled in the 19 thematic dictionaries were normalized to a single internal identifier in the KB, where possible, to allow for more efficient mining of relevant terms in the text and enable linking terms to external sources. Additionally, term redundancies within the same dictionary are unified into one term. The text mining of most dictionary terms is generally straightforward. However, gene names are frequently interchangeably used with their protein product names/symbols in biomedical text. Thus, for example, in the “Human Genes and Proteins” dictionary, we combine EntrezGene [76] nomenclature (gene names/symbols) with UniProt [77] nomenclature (protein names/symbols). Also, concepts in all the dictionaries are normalized where possible, i.e., names/synonyms, and symbols referring to the same concept are retrieved by a single entity.

Initial indexing is performed, that is, concepts in these dictionaries are mined in the prepared literature corpus and color-coded to reflect the dictionary from which it was retrieved. In this manner, we can identify 1/ promiscuous terms through their high frequencies due to their use as ordinary English words, and 2/ terms in newly imported dictionaries not found in the prepared literature corpus, which we exclude as part of the dictionary cleaning process. Then, re-indexing is performed based on clean dictionary data.

**Table 1. DES-Amyloidoses dictionaries, terms per dictionary, and terms enriched in the literature corpus.**

Dictionary	# Enriched concepts	# Enriched pairs of concepts	# Enriched pairs that contain amyloids	Status
<b>Chemicals/Compounds</b>				
Amyloids (Human and Mouse) [in-house compiled]	298	19,218	1,084	<b>newly compiled</b>
Chemical Entities of Biological Interest (ChEBI) [78]	6,996	233,926	1,769	pre-existing in DES
Lipids (Lipid Maps) [79, 80]	580	19,678	61	pre-existing in DES
Metabolites (MetaboLights) [81]	1,636	54,755	240	pre-existing in DES
Toxins (T3DB) [82]	1,046	47,615	314	pre-existing in DES
<b>Functional Annotation</b>				
Biological Process (GO) [83]	2,327	44,829	523	pre-existing in DES
Cellular Component (GO) [83]	645	17,404	201	pre-existing in DES
Molecular Function (GO) [83]	820	14,245	157	pre-existing in DES
Pathways (KEGG [84], Reactome [85], UniPathway [86], PANTHER [87])	723	18,773	144	pre-existing in DES
<b>Diseases</b>				
DOID Ontology (Bioportal) Human Disease Ontology [88]	1,517	33,970	312	pre-existing in DES
HP Ontology (Bioportal) Human Phenotype Ontology [89]	1,592	38,701	312	pre-existing in DES
SIDER (Drug Indications and Side Effects) [90]	1,354	31,828	276	pre-existing in DES
<b>Drugs</b>				
Drugs (DrugBank) [91]	1,887	76,994	531	pre-existing in DES
<b>Anatomies</b>				
Human Anatomy [in-house compiled]	1,325	76,315	1,001	pre-existing in DES
<b>Human</b>				
Human Genes & Proteins (EntrezGene) [92]	13,346	756,763	11,085	pre-existing in DES
Human Long Non-Coding RNAs [93]	99	2,232	20	pre-existing in DES
Human microRNAs [93, 94]	741	26,002	61	pre-existing in DES
Human Transcription Factors [95]	1,010	63,353	647	pre-existing in DES
Mutations (tmVar) [96]	5,144	40,204	480	pre-existing in DES

<https://doi.org/10.1371/journal.pone.0271737.t001>

[Table 1](#) lists the dictionaries used and provide, 1/ the number of enriched concepts in the literature corpus per dictionary, 2/ the number of enriched concept pairs in the literature corpus per dictionary, and 3/ the number of enriched pairs that include an amyloid per dictionary. Out of all concepts in the 19 dictionaries, 43,086 were found to be statistically enriched. Based on the statistically enriched terms in the corpus, the system identified 796,409 enriched pairs of concepts in the literature corpus. Embedding the network of concept pairs enabled semantic similarity computation between the KB concepts.

**2.2.1 Enriched concepts.** The frequency at which a concept appears in the full literature set is expected to be similar to its frequency in that literature's random subset. Thus, in DES, a concept is defined as enriched when overrepresented in the topic-specific corpus, in this case, the DES-Amyloidoses corpus, compared to the complete set of PubMed and PubMed Central articles in DES MongoDB database. We calculated the false discovery rate (FDR)  $< 0.05$  (*P-value*) based on the Benjamini-Hochberg procedure to correct for multiplicity testing. Concepts are quantified to be enriched when it has an FDR/*P-value*  $< 0.05$  in the DES-Amyloidoses corpus compared to the complete article set. In this manner, the KB provides the user with the most topic-relevant concepts.

**2.2.2 Enriched pairs.** The enriched or topic-relevant concepts co-occur in literature with several other concepts. We classified concepts as co-occurring if they were mentioned within a 200-character distance in text. This co-occurrence of concepts may also be enriched; for example, the enriched concept may co-occur with the other concept 90% of the time. Thus, DES-Amyloidoses also provides users with the pairs of enriched concepts based on co-occurrence (or association) compared to the enriched concept's occurrence (the second concept in the pair may or may not be enriched). The concepts' co-occurrence signifies a potential association, but concepts in the enriched pair might not be directly associated. Nonetheless, enriched concept pairs increase the probability of an association between the two concepts existing.

**2.2.3 Semantic similarity.** Here, we used semantic similarity as a metric that establishes how close in meaning or relatedness two concepts are, based on their distribution within a text corpus. The semantic similarity relatedness can be in the form of hypernymy/hyponymy, antonymy, or synonymy. For example, liquid and water are semantically similar even though they are different concepts because water is a hyponym of liquid, and hence are more likely to be co-mentioned in the same context. We acquired the semantic similarity by first training a skip-gram Word2Vec model on the DES-Amyloidoses corpus, then calculating the cosine distance between concept embeddings, representing semantic similarity in DES. Therefore, semantic similarity represents a concept co-occurrence in DES, which might not be direct.

### 3. DES-Amyloidoses utilities and case study

DES-Amyloidoses provide users with a list of topic-specific enriched concepts and lists of concepts frequently mentioned in the same text as the enriched concept based on the amyloid-related literature, as these concepts may be directly or indirectly associated with the enriched concept. Users can explore these enriched concepts via multiple links built into the KB, including "Enriched Concepts", "Enriched Pairs", and "Semantic Similarity" (described in detail by [70, 74]).

Briefly, the "Enriched Concepts" link allows users to familiarize themselves with and explore the concepts enriched in the amyloid-related literature, such as APP, amyloid-beta, MAPT, cerebral, etc. The "Enriched Pairs" link allows the users to focus on a specific enriched concept of their interest, and explore other concepts (not enriched in the amyloid-related literature) that may be associated with the enriched concept, for example, NLRP3 and aortic sinus; NLRP3 and saturated fatty acid anion; NLRP3 and cytochalasin; NLRP3 and LRRFIP1, and so

forth. On the other hand, the “Semantic Similarity” link allows the users to explore concept co-occurrence that is not necessarily directly linked, as in the “Enriched Pairs” case; checking such co-occurring concepts for inferred biological association can be used to shortlist potential novel hypotheses.

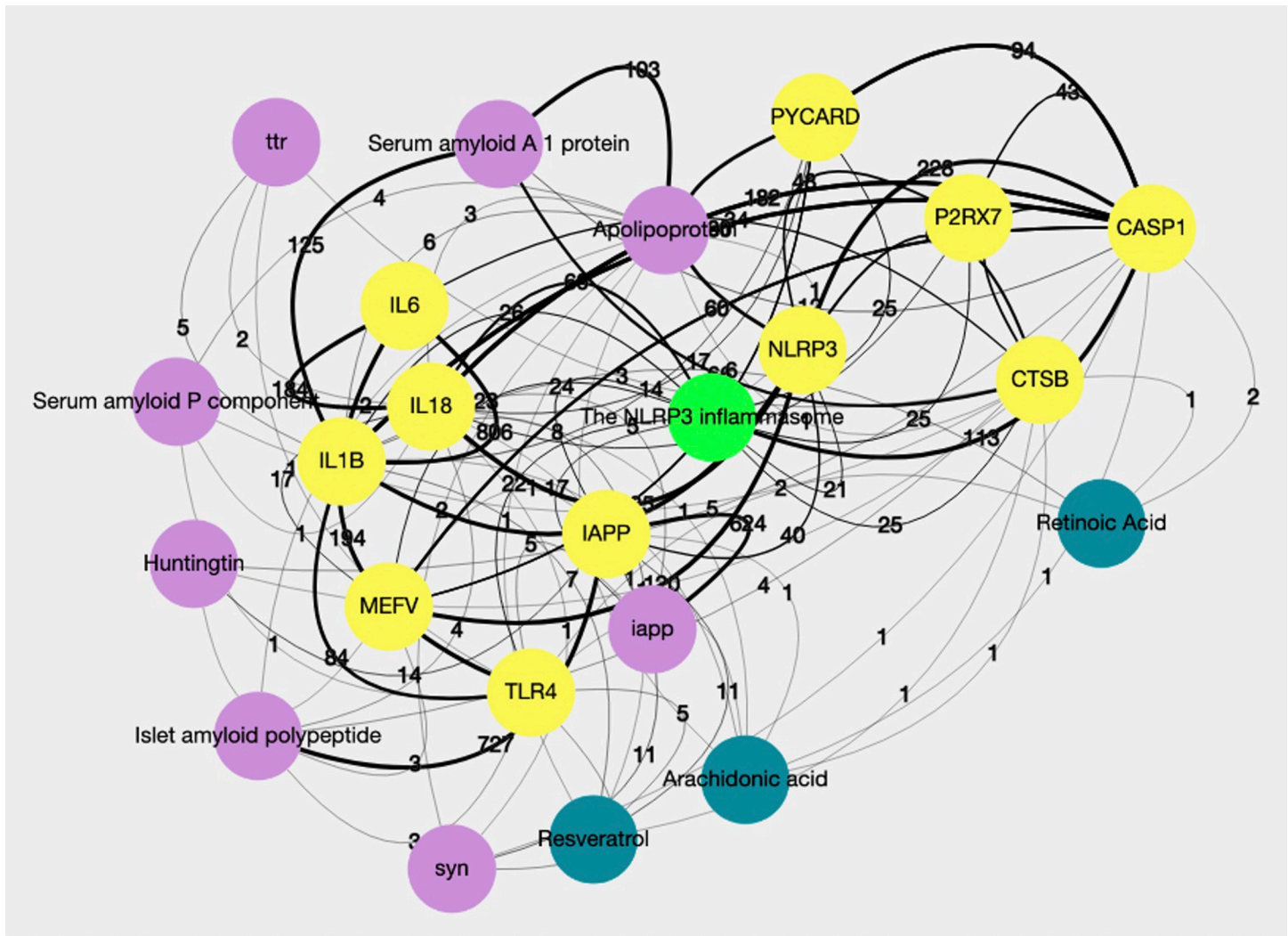
For each exploration, users can view enriched concepts in pre-compiled theme-based dictionaries and restrict concepts based on a specific term/concept, for example, using the text box to search for the enriched concepts that contain the term Amyloid, retrieves concepts such as “Conjunctival amyloidosis”, “paramyloidosis”, “amyloid precursor protein metabolic process”, and so forth. Concepts of interest can also be sorted using ranking options, including false discovery rate (FDR), density, KB frequency (KB\_FDR), and background frequency (BKB\_FDR) (see “Column visibility”), and results can be exported in excel or csv format via the “Export” link. Also, each concept is linked to a hover box from which users can generate a “Network” or retrieve “Term Co-occurrences”; generated networks can be saved in the json.txt format using the “Export Network” link.

### 3.1 Case study 1: Illustrating the usefulness of DES-Amyloidoses as a research support system: Progression of an Amyloid “network” in the pathogenesis of AD

Here we demonstrate the efficacy of DES-Amyloidoses in exploring inflammasome-amyloid associations, as amyloids have been demonstrated to activate the inflammasome to process Interleukin 1 beta (IL-1 $\beta$ ) [97]. For example, activation is induced in AD via A $\beta$  [98], in Diabetes type II (T2D) via IAPP [99], in PD via  $\alpha$ -synuclein [100] and in amyotrophic lateral sclerosis (ALS) through SOD1 [101].

We started this process by checking if the inflammasome pathway is an enriched concept in the amyloid literature. This was done by clicking the DES-Amyloidoses “Enriched Pairs” option (Step 1), which opens a page that lists associated terms from all dictionaries in two columns. In the first column, where users can specify the first dictionary (or concept A), we filtered by selecting the “Human Genes and Proteins (EntrezGene)” dictionary from the drop-down menu. Similarly, for the second dictionary (or concept B), we selected the “Pathways” dictionary from the drop-down menu. The “inflammasome” pathway is listed multiple times, and even the “The NLRP3 inflammasome” pathway is listed as one of the most significant pathways. Because Halle *et al.* [98] demonstrated that activation of the NALP3 inflammasome is an essential process in AD-related inflammation and tissue damage, we proceed by accessing the right-click menu (or hovering over) for the “The NLRP3 inflammasome” concept to generate a network (Step 2). On the network page, we selected “Amyloids”, “Human Genes and Proteins”, and “Lipids” in the “Choose Dictionaries” menu, then selected the “The NLRP3 inflammasome” node and used the right-click menu to ‘Expand’ the network with nodes from the selected dictionaries (Step 3). Using the same dictionaries, we performed a second round of network expansion on all nodes obtained in Step 3. We removed all nodes with two or fewer links (Fig 1, Step 4).

The final network comprises two sub-networks; one centered on the NLRP3 inflammasome node, while the other is centered on the amyloid, IAPP. The amyloid network, in this case, IAPP, works in concert with the activation of the inflammasome through NLRP3. Moreover, it also depicts an array of inflammation-related genes/proteins, and a direct association between NLRP3-CASP1 (inflammation-associated protein) and TRIM20 (MEFV) associated with innate regulation immunity suggests crosstalk between amyloids, innate immunity, and inflammation. Specifically, MEFV inhibits the NLRP3-CASP1 inflammasome pathway by directly binding to inflammasome components, including NLRP1, NLRP3, and CASP1. Also, it recruits autophagic machinery to execute degradation [102]. In this manner, autophagy



**Fig 1.** A depiction of the inflammasome-amyloid “network” involved in Alzheimer’s disease’s pathogenesis.

<https://doi.org/10.1371/journal.pone.0271737.g001>

controls the hub signaling machinery [102]. A similar direct association is depicted between NLRP3-CASP1 (inflammation-associated proteins) and Cathepsin B (CTSB). CTSB plays a crucial role in several physiological processes, one essential being driving degradation within the lysosome [103]. CTSB reduces the expression levels of lysosomal and autophagy-related proteins, thereby reducing the number of lysosomes and autophagosomes in the cell. It has further been demonstrated that CTSB is released from the lysosome with lysosomal damage, causing autophagy-lysosomal dysfunction and the activation of the NLRP3-CASP1 inflammasome pathway [98, 104, 105]. Thus, both IAPP and A $\beta$  induce NLRP3-CASP1 activation through a mechanism involving the released CTSB [99].

The precise mechanism of NLRP3-CASP1 activation is still debated, however, considering that 1/ CTSB inhibition prevents A $\beta$ -induced NLRP3-CASP1 activation, which reduces amyloid plaque load and improves memory in the AD brain of mouse models [106], and 2/ CTSB has been associated with several amyloids, this network provides users with a bird’s-eye view of amyloid-related literature. It suggests CTSB should be considered a potential therapeutic approach for treating AD wherein the inflammasome is targeted.

### 3.2 Case study 2: DES-Amyloidoses unveils the microRNA, possibly regulating the Amyloid “network” in the pathogenesis of AD

To identify the microRNAs possibly regulating the inflammasome-amyloid associations in AD, we explored the microRNAs semantically linked to the essential genes ("IAAP," "CTSB," "NLRP3," "PYCARD," and "CASP1") identified in the inflammasome-amyloid associations in AD (see case study 1). This was done by clicking the DES-Amyloidoses “Semantic Similarity” option, which opens a page with two columns. In the first column, we inserted the name of the gene of interest (or concept A), but for the second column, we selected the “Human microRNAs” dictionary from the drop-down menu (or concept B) (see Fig 2). Then, we repeated this process for all genes of interest and tabulated all microRNAs significantly linked to these genes (provided in Fig 2).

We identified 13 unique microRNA with cosine similarity above 0.7. However, we found no literature connecting these microRNA to the amyloid “network” despite this indirect association depicted by DES-Amyloidoses. Consequently, we used the microRNA Data Integration Portal (mirDIP) to search if these microRNAs are predicted to target our set of genes. As a result, we found 9 of the 13 unique microRNA (hsa-miR-6089, hsa-miR-3661, hsa-miR-299, hsa-miR-653, hsa-miR-129-1, hsa-miR-625, hsa-miR-1302, hsa-miR-489, hsa-miR-34a, and hsa-miR-1908) predicted to target all five genes, and a literature search showed all these microRNAs have differential expression linked to AD [107–110] (see Fig 3). Furthermore, one of the microRNAs, hsa-miR-1908, was experimentally validated to inhibit ApoE expression, which suggested that miR-1908 inhibits Aβ clearance by repressing ApoE expression [111].

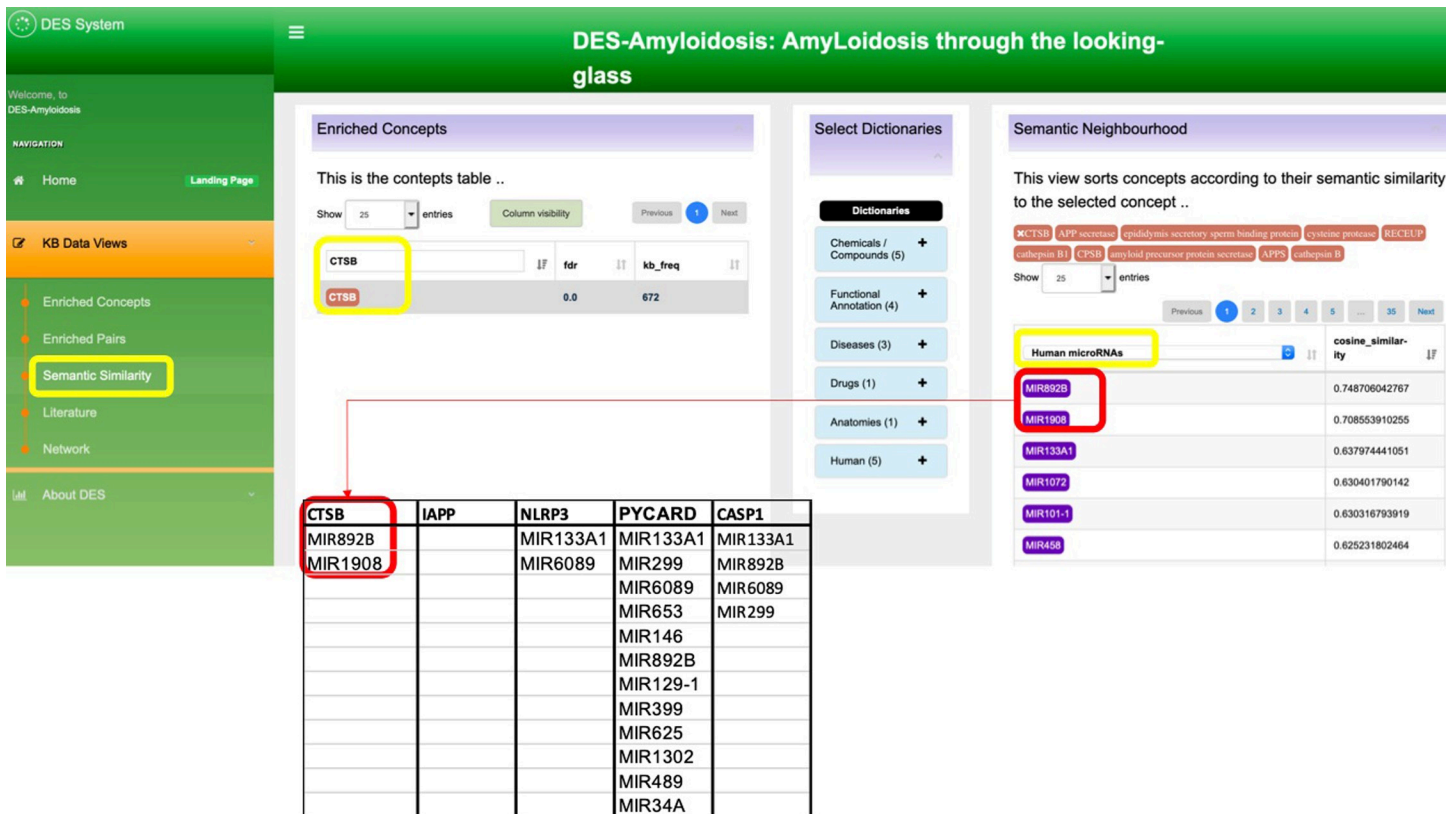


Fig 2. An illustration of how DES-Amyloidoses can be used to identify relationships between the concepts based on semantic similarity. The yellow square indicates the changes that were implemented, and the tabulation shows the microRNAs that were shortlisted for this process.

<https://doi.org/10.1371/journal.pone.0271737.g002>



CTSB	IAPP	NLRP3	PYCARD	CASP1	Differential expression		
					AD MCI marker	Preclinical AD	AD Blood Mononuclear Cells
MIR6089 <input type="checkbox"/>	MIR6089 <input type="checkbox"/>	MIR6089 <input type="checkbox"/>	MIR6089 <input type="checkbox"/>	MIR6089 <input type="checkbox"/>	no		
MIR3661 <input type="checkbox"/>	MIR3661 <input type="checkbox"/>	MIR3661 <input type="checkbox"/>	MIR3661 <input type="checkbox"/>	MIR3661 <input type="checkbox"/>	yes		
MIR299 <input type="checkbox"/>	MIR299 <input type="checkbox"/>	MIR299 <input type="checkbox"/>	hsa-miR-299-3p <input type="checkbox"/> hsa-miR-299-5p <input type="checkbox"/>	hsa-miR-299-3p <input checked="" type="checkbox"/> hsa-miR-299-5p <input type="checkbox"/>	yes		
MIR653 <input type="checkbox"/>	MIR653 <input type="checkbox"/>	MIR653 <input type="checkbox"/>	hsa-miR-653-5p <input type="checkbox"/> hsa-miR-653-3p <input type="checkbox"/>	MIR653 <input type="checkbox"/>	no		yes
MIR129-1 <input type="checkbox"/>	MIR129-1 <input type="checkbox"/>	MIR129-1 <input type="checkbox"/>	MIR129-1 <input type="checkbox"/>	MIR129-1 <input type="checkbox"/>	yes		
hsa-miR-625-5p <input checked="" type="checkbox"/> hsa-miR-625-3p <input type="checkbox"/>	MIR625 <input type="checkbox"/>	hsa-miR-625-5p <input type="checkbox"/> hsa-miR-625-3p <input type="checkbox"/>	hsa-miR-625-5p <input type="checkbox"/> hsa-miR-625-3p <input type="checkbox"/>	MIR625 <input type="checkbox"/>	yes		
MIR1302 <input type="checkbox"/>	MIR1302 <input type="checkbox"/>	MIR1302 <input type="checkbox"/>	MIR1302 <input type="checkbox"/>	MIR1302 <input type="checkbox"/>		yes	
hsa-miR-489-3p <input checked="" type="checkbox"/> hsa-miR-489-5p <input type="checkbox"/>	hsa-miR-489-3p <input type="checkbox"/> hsa-miR-489-5p <input type="checkbox"/>	MIR489 <input type="checkbox"/>	MIR489 <input type="checkbox"/>	hsa-miR-489-3p <input type="checkbox"/> hsa-miR-489-5p <input type="checkbox"/>			yes
hsa-miR-34a-5p <input checked="" type="checkbox"/> hsa-miR-34a-3p <input type="checkbox"/>	MIR34A <input type="checkbox"/>	MIR34A <input type="checkbox"/>	hsa-miR-34a-5p <input type="checkbox"/> hsa-miR-34a-3p <input type="checkbox"/>	MIR34A <input type="checkbox"/>			yes
hsa-miR-1908-5p <input type="checkbox"/> hsa-miR-1908-3p <input type="checkbox"/>	hsa-miR-1908-5p <input type="checkbox"/> hsa-miR-1908-3p <input type="checkbox"/>	hsa-miR-1908-5p <input type="checkbox"/> hsa-miR-1908-3p <input type="checkbox"/>	hsa-miR-1908-5p <input type="checkbox"/> hsa-miR-1908-3p <input type="checkbox"/>	hsa-miR-1908-5p <input type="checkbox"/> hsa-miR-1908-3p <input type="checkbox"/>			yes

mirDIP scores:  Top 5% high  Top 1/3 medium  Bottom 2/3 low

Fig 3. The microRNAs predicted to target the essential genes with the mirDIP scores indicated. AD MCI marker [110]; Preclinical AD [107]; AD Blood Mononuclear Cells [108, 109].

<https://doi.org/10.1371/journal.pone.0271737.g003>

### 4. Discussion

The idea that the amyloids present a system is not new [112–114]. The reason is that they have networks that suggest interrelations with other biological networks and environmental stressors that can induce metabolic changes that may impair homeostatic defenses during the lifetime of humans. Furthermore, changes in the interrelations with these other networks may cause differences in patterns, heterogeneity, age of onset, disease progression, and divergent patient phenotypes. Some phenotypes have emerged due to different Aβ conformations [115] and the seeding capabilities of the amyloids [38, 116], which adds to this complexity. For example, IAPP was identified in human cerebral Aβ deposits, and Aβ fibrils were found to seed IAPP in vivo as efficiently as hproIAPP [116], which offers a possible molecular link as to

why epidemiological studies suggest patients with type 2 diabetes have an almost twofold increased risk of developing AD [117].

Specific amyloid converting endotrophic triggers have not yet been pinpointed, despite genetic mutations linked to some disorders' etiology. The reason is that sensitivity towards environmental pathogens (e.g., pesticides), reactive oxygen species, or metals characterizes amyloid aggregation, which partially explains the idiopathic amyloid disorders, is more frequent than familial cases. Fig 1 suggests that IAPP being more prone to A $\beta$  Amyloidosis than AA amyloidosis may be a consequence of the divergence in its interaction with the NLRP3-inflammasome that can sense and respond to dysfunction triggered by environmental stressors [114].

Clearance of amyloids by phagocytosis is a needed physiological process, but it can adversely perturb cellular homeostasis. Specifically, phagocytosis of amyloid peptides, like A $\beta$  and IAPP, still may lead to the activation of the innate immunity activator, NLRP3 [118]. A recent paper by Cai and colleagues [119] showed that pattern recognition receptors and prion could replace NLRP3 and ASC, respectively, in inflammasome signaling. This may indicate that amyloids have a more intrinsic role in inflammatory processes than previously realized, and they may be working with other signaling proteins to shift perturbed homeostatic mechanisms to typical values, which suggests a protective role. On the other hand, these amyloids may independently aggravate inflammation in neurodegeneration disorders such as AD by activating caspase 1, which then cleaves pro-IL-1 $\beta$  and pro-IL-18 into their mature, secreted forms resulting in neuronal cell death [120]. A multitarget approach may be a promising therapy strategy; as the case studies indicate, targeting Cathepsin B in concert with an inflammasome-amyloid network associated microRNA/s may reset the mechanism altered in AD.

## 5. Concluding remarks

The notion that amyloids comprise a “Network” that can be defined as a system is an essential jump in understanding protein folding diseases. By defining the network/system in which a disease is presented through an integrative perspective from the genotype to the phenotype, it shall be possible to discern important “hub” proteins and pathways, as shown through the “inflammasome-amyloid hub”. With DES-Amyloidoses we presented an “amyloid system” and its interacting network based on the literature and data mining approaches compiled into a KB. This system enables a novel way to interrogate information about amyloids and associated diseases. We hope the two case studies shared demonstrate how users may find DES-Amyloidoses to be a valuable tool for supporting amyloidoses-related research questions. Furthermore, we intend to update the KB biannually to ensure the KB contents remain current.

## Acknowledgments

This work is part of the collaboration between the Laboratory of Radiobiology and Molecular Genetics, Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia and King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia.

## Author Contributions

**Conceptualization:** Vladan P. Bajic, Vladimir B. Bajic, Magbubah Essack.

**Data curation:** Rozaimi Razali, Lada Zivkovic, Mahmut Uludag.

**Formal analysis:** Vladan P. Bajic, Katja Lakota, Lada Zivkovic, Biljana Spremo-Potparevic.

**Software:** Adil Salhi, Aleksandar Radovanovic, Mahmut Uludag, Benoit Marchand.

**Supervision:** Vladimir B. Bajic.

**Visualization:** Faroug Tifratene.

**Writing – original draft:** Vladan P. Bajic, Olaa Motwalli, Magbubah Essack.

**Writing – review & editing:** Takashi Gojobori, Esma R. Isenovic, Magbubah Essack.

## References

1. Goedert M: Tau protein and the neurofibrillary pathology of Alzheimer's disease. *Trends Neurosci* 1993, 16(11):460–465. [https://doi.org/10.1016/0166-2236\(93\)90078-z](https://doi.org/10.1016/0166-2236(93)90078-z) PMID: 7507619
2. Kosik KS, Joachim CL, Selkoe DJ: Microtubule-associated protein tau (tau) is a major antigenic component of paired helical filaments in Alzheimer disease. *Proc Natl Acad Sci U S A* 1986, 83(11):4044–4048. <https://doi.org/10.1073/pnas.83.11.4044> PMID: 2424016
3. Walti MA, Ravotti F, Arai H, Glabe CG, Wall JS, Bockmann A, et al: Atomic-resolution structure of a disease-relevant Aβ(1–42) amyloid fibril. *Proc Natl Acad Sci U S A* 2016, 113(34):E4976–4984. <https://doi.org/10.1073/pnas.1600749113> PMID: 27469165
4. Colvin MT, Silvers R, Ni QZ, Can TV, Sergeev I, Rosay M, et al: Atomic Resolution Structure of Monomorphic Aβ(42) Amyloid Fibrils. *J Am Chem Soc* 2016, 138(30):9663–9674. <https://doi.org/10.1021/jacs.6b05129> PMID: 27355699
5. Xu S: Cross-beta-sheet structure in amyloid fiber formation. *J Phys Chem B* 2009, 113(37):12447–12455. <https://doi.org/10.1021/jp903106x> PMID: 19705845
6. Knowles TP, Vendruscolo M, Dobson CM: The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* 2014, 15(6):384–396. <https://doi.org/10.1038/nrm3810> PMID: 24854788
7. Chiti F, Dobson CM: Protein misfolding, functional Amyloid, and human disease. *Annu Rev Biochem* 2006, 75:333–366. <https://doi.org/10.1146/annurev.biochem.75.101304.123901> PMID: 16756495
8. Cowan AJ, Skinner M, Seldin DC, Berk JL, Lichtenstein DR, O'Hara CJ, et al: Amyloidosis of the gastrointestinal tract: a 13-year, single-center, referral experience. *Haematologica* 2013, 98(1):141–146. <https://doi.org/10.3324/haematol.2012.068155> PMID: 22733017
9. Said SM, Sethi S, Valeri AM, Leung N, Cornell LD, Fidler ME, et al: Renal amyloidosis: origin and clinicopathologic correlations of 474 recent cases. *Clin J Am Soc Nephrol* 2013, 8(9):1515–1523. <https://doi.org/10.2215/CJN.10491012> PMID: 23704299
10. Jacobs R, Maasdorp E, Malherbe S, Loxton AG, Stanley K, Van Der Spuy G, et al: Diagnostic potential of novel salivary host biomarkers as candidates for the immunological diagnosis of tuberculosis disease and monitoring of tuberculosis treatment response. *PloS one* 2016, 11(8):e0160546. <https://doi.org/10.1371/journal.pone.0160546> PMID: 27487181
11. Rambaran RN, Serpell LC: Amyloid fibrils. *Prion* 2008, 2(3):112–117. <https://doi.org/10.4161/pri.2.3.7488> PMID: 19158505
12. Sipe JD, Benson MD, Buxbaum JN, Ikeda SI, Merlini G, Saraiva MJ, et al: Amyloid fibril proteins and Amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 Nomenclature Guidelines. *Amyloid* 2016, 23(4):209–213. <https://doi.org/10.1080/13506129.2016.1257986> PMID: 27884064
13. Schleegeer M, Deckert-Gaudig T, Deckert V, Velikov KP, Koenderink G, Bonn M: Amyloids: From molecular structure to mechanical properties. *Polymer* 2013, 54(10):2473–2488.
14. Eisenberg D, Jucker M: The amyloid state of proteins in human diseases. *Cell* 2012, 148(6):1188–1203. <https://doi.org/10.1016/j.cell.2012.02.022> PMID: 22424229
15. Espargaro A, Busquets MA, Estelrich J, Sabate R: Key Points Concerning Amyloid Infectivity and Prion-Like Neuronal Invasion. *Front Mol Neurosci* 2016, 9:29. <https://doi.org/10.3389/fnmol.2016.00029> PMID: 27147962
16. Clavaguera F, Hench J, Goedert M, Tolnay M: Invited review: Prion-like transmission and spreading of tau pathology. *Neuropathology and applied neurobiology* 2015, 41(1):47–58. <https://doi.org/10.1111/nan.12197> PMID: 25399729
17. Hyman BT: Tau propagation, different tau phenotypes, and prion-like properties of tau. *Neuron* 2014, 82(6):1189–1190. <https://doi.org/10.1016/j.neuron.2014.06.004> PMID: 24945760

18. Polymenidou M, Cleveland DW: Prion-like spread of protein aggregates in neurodegeneration. *Journal of Experimental Medicine* 2012, 209(5):889–893. <https://doi.org/10.1084/jem.20120741> PMID: 22566400
19. Hansen C, Angot E, Bergström A-L, Steiner JA, Pieri L, Paul G, et al:  $\alpha$ -Synuclein propagates from mouse brain to grafted dopaminergic neurons and seeds aggregation in cultured human cells. *The Journal of clinical investigation*, 2011, 121(2):715–725. <https://doi.org/10.1172/JCI43366> PMID: 21245577
20. Goedert M, Falcon B, Clavaguera F, Tolnay M: Prion-like mechanisms in the pathogenesis of tauopathies and synucleinopathies. *Current neurology and neuroscience reports* 2014, 14(11):1–11. <https://doi.org/10.1007/s11910-014-0495-z> PMID: 25218483
21. Herva ME, Spillantini MG: Parkinson's disease as a member of Prion-like disorders. *Virus research*, 2015, 207:38–46. <https://doi.org/10.1016/j.virusres.2014.10.016> PMID: 25456401
22. Kane MD, Lipinski WJ, Callahan MJ, Bian F, Durham RA, Schwarz RD, et al: Evidence for seeding of  $\beta$ -amyloid by intracerebral infusion of Alzheimer brain extracts in  $\beta$ -amyloid precursor protein-transgenic mice. *Journal of Neuroscience* 2000, 20(10):3606–3611. <https://doi.org/10.1523/JNEUROSCI.20-10-03606.2000> PMID: 10804202
23. Nath S, Agholme L, Kurudenkandy FR, Granseth B, Marcusson J, Hallbeck M: Spreading of neurodegenerative pathology via neuron-to-neuron transmission of  $\beta$ -amyloid. *Journal of Neuroscience* 2012, 32(26):8767–8777. <https://doi.org/10.1523/JNEUROSCI.0615-12.2012> PMID: 22745479
24. Walker LC, Diamond MI, Duff KE, Hyman BT: Mechanisms of protein seeding in neurodegenerative diseases. *JAMA neurology* 2013, 70(3):304–310. <https://doi.org/10.1001/jamaneurol.2013.1453> PMID: 23599928
25. Ren P-H, Lauckner JE, Kachirskaja I, Heuser JE, Melki R, Kopito RR: Cytoplasmic penetration and persistent infection of mammalian cells by polyglutamine aggregates. *Nature cell biology* 2009, 11(2):219–225. <https://doi.org/10.1038/ncb1830> PMID: 19151706
26. Trevino RS, Lauckner JE, Sourigues Y, Pearce MM, Bousset L, Melki R, et al: Fibrillar structure and charge determine the interaction of polyglutamine protein aggregates with the cell surface. *Journal of Biological Chemistry* 2012, 287(35):29722–29728. <https://doi.org/10.1074/jbc.M112.372474> PMID: 22753412
27. Polymenidou M, Cleveland DW: The seeds of neurodegeneration: prion-like spreading in ALS. *Cell* 2011, 147(3):498–508. <https://doi.org/10.1016/j.cell.2011.10.011> PMID: 22036560
28. Münch C, O'Brien J, Bertolotti A: Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. *Proceedings of the National Academy of Sciences* 2011, 108(9):3548–3553. <https://doi.org/10.1073/pnas.1017275108> PMID: 21321227
29. Furukawa Y, Kaneko K, Watanabe S, Yamanaka K, Nukina N: A seeding reaction recapitulates intracellular formation of Sarkosyl-insoluble transactivation response element (TAR) DNA-binding protein-43 inclusions. *Journal of Biological Chemistry* 2011, 286(21):18664–18672. <https://doi.org/10.1074/jbc.M111.231209> PMID: 21454603
30. Nonaka T, Masuda-Suzukake M, Arai T, Hasegawa Y, Akatsu H, Obi T, et al: Prion-like properties of pathological TDP-43 aggregates from diseased brains. *Cell reports* 2013, 4(1):124–134. <https://doi.org/10.1016/j.celrep.2013.06.007> PMID: 23831027
31. Kfoury N, Holmes BB, Jiang H, Holtzman DM, Diamond MI: Trans-cellular propagation of tau aggregation by fibrillar species. *Journal of Biological Chemistry* 2012, 287(23):19440–19451. <https://doi.org/10.1074/jbc.M112.346072> PMID: 22461630
32. Yanamandra K, Kfoury N, Jiang H, Mahan TE, Ma S, Maloney SE, et al: Anti-tau antibodies that block tau aggregate seeding in vitro markedly decrease pathology and improve cognition in vivo. *Neuron* 2013, 80(2):402–414. <https://doi.org/10.1016/j.neuron.2013.07.046> PMID: 24075978
33. Frost B, Jacks RL, Diamond MI: Propagation of tau misfolding from the outside to the inside of a cell. *Journal of Biological Chemistry* 2009, 284(19):12845–12852. <https://doi.org/10.1074/jbc.M808759200> PMID: 19282288
34. Iba M, Guo JL, McBride JD, Zhang B, Trojanowski JQ, Lee VM-Y: Synthetic tau fibrils mediate transmission of neurofibrillary tangles in a transgenic mouse model of Alzheimer's-like tauopathy. *Journal of Neuroscience* 2013, 33(3):1024–1037. <https://doi.org/10.1523/JNEUROSCI.2642-12.2013> PMID: 23325240
35. Freundt EC, Maynard N, Clancy EK, Roy S, Bousset L, Sourigues Y, et al: Neuron-to-neuron transmission of  $\alpha$ -synuclein fibrils through axonal transport. *Annals of neurology* 2012, 72(4):517–524. <https://doi.org/10.1002/ana.23747> PMID: 23109146
36. Luk KC, Kehm VM, Zhang B, O'Brien P, Trojanowski JQ, Lee VM: Intracerebral inoculation of pathological  $\alpha$ -synuclein initiates a rapidly progressive neurodegenerative  $\alpha$ -synucleinopathy in mice.

- Journal of Experimental Medicine 2012;jem. 20112457. <https://doi.org/10.1084/jem.20112457> PMID: 22508839
37. Eisele YS, Obermüller U, Heilbronner G, Baumann F, Kaeser SA, Wolburg H, et al: Peripherally applied A $\beta$ -containing inoculates induce cerebral  $\beta$ -amyloidosis. *Science* 2010, 330(6006):980–982. <https://doi.org/10.1126/science.1194516> PMID: 20966215
  38. Oskarsson ME, Paulsson JF, Schultz SW, Ingelsson M, Westermark P, Westermark GT: In vivo seeding and cross-seeding of localized amyloidosis: a molecular link between type 2 diabetes and Alzheimer disease. *The American journal of pathology* 2015, 185(3):834–846. <https://doi.org/10.1016/j.ajpath.2014.11.016> PMID: 25700985
  39. N Fawver J, Ghiwot Y, Koola C, Carrera W, Rodriguez-Rivera J, Hernandez C, et al: Islet amyloid polypeptide (IAPP): a second amyloid in Alzheimer's disease. *Current Alzheimer Research* 2014, 11(10):928–940. <https://doi.org/10.2174/1567205011666141107124538> PMID: 25387341
  40. Hu R, Zhang M, Chen H, Jiang B, Zheng J: Cross-seeding interaction between  $\beta$ -amyloid and human islet amyloid polypeptide. *ACS chemical neuroscience* 2015, 6(10):1759–1768. <https://doi.org/10.1021/acscchemneuro.5b00192> PMID: 26255739
  41. Kindy MS, Yu J, Guo J-T, Zhu H: Apolipoprotein serum amyloid A in Alzheimer's disease. *Journal of Alzheimer's Disease* 1999, 1(3):155–167. <https://doi.org/10.3233/jad-1999-1303> PMID: 12214001
  42. Fein JA, Sokolow S, Miller CA, Vinters HV, Yang F, Cole GM, et al: Co-localization of amyloid beta and tau pathology in Alzheimer's disease synaptosomes. *The American journal of pathology* 2008, 172(6):1683–1692. <https://doi.org/10.2353/ajpath.2008.070829> PMID: 18467692
  43. Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L: Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nature biotechnology* 2004, 22(10):1302–1306. <https://doi.org/10.1038/nbt1012> PMID: 15361882
  44. Thompson MJ, Sievers SA, Karanicolas J, Ivanova MI, Baker D, Eisenberg D: The 3D profile method for identifying fibril-forming segments of proteins. *Proceedings of the National Academy of Sciences of the United States of America* 2006, 103(11):4074–4078. <https://doi.org/10.1073/pnas.0511295103> PMID: 16537487
  45. Goldschmidt L, Teng PK, Riek R, Eisenberg D: Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proceedings of the National Academy of Sciences* 2010, 107(8):3487–3492. <https://doi.org/10.1073/pnas.0915166107> PMID: 20133726
  46. Trovato A, Seno F, Tosatto SC: The PASTA server for protein aggregation prediction. *Protein Engineering Design and Selection* 2007, 20(10):521–523. <https://doi.org/10.1093/protein/gzm042> PMID: 17720750
  47. Kim C, Choi J, Lee SJ, Welsh WJ, Yoon S: NetCSSP: web application for predicting chameleon sequences and amyloid fibril formation. *Nucleic acids research* 2009, 37(suppl 2):W469–W473. <https://doi.org/10.1093/nar/gkp351> PMID: 19468045
  48. Garbuzynskiy SO, Lobanov MY, Galzitskaya OV: FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics* 2010, 26(3):326–332. <https://doi.org/10.1093/bioinformatics/btp691> PMID: 20019059
  49. O'donnell CW, Waldispühl J, Lis M, Halfmann R, Devadas S, Lindquist S, et al: A method for probing the mutational landscape of amyloid structure. *Bioinformatics* 2011, 27(13):i34–i42. <https://doi.org/10.1093/bioinformatics/btr238> PMID: 21685090
  50. Bryan AW, O'Donnell CW, Menke M, Cowen LJ, Lindquist S, Berger B: STITCHER: Dynamic assembly of likely amyloid and prion  $\beta$ -structures from secondary structure predictions. *Proteins: Structure, Function, and Bioinformatics* 2012, 80(2):410–420. <https://doi.org/10.1002/prot.23203> PMID: 22095906
  51. Tsois AC, Papandreou NC, Iconomidou VA, Hamodrakas SJ: A consensus method for the prediction of 'aggregation-prone' peptides in globular proteins. *PloS one* 2013, 8(1):e54175. <https://doi.org/10.1371/journal.pone.0054175> PMID: 23326595
  52. de Groot NS, Pallarés I, Avilés FX, Vendrell J, Ventura S: Prediction of "hot spots" of aggregation in disease-linked polypeptides. *BMC Structural Biology* 2005, 5(1):18.
  53. Tartaglia GG, Cavalli A, Pellarin R, Caffisch A: Prediction of aggregation rate and aggregation-prone segments in polypeptide sequences. *Protein Science* 2005, 14(10):2723–2734. <https://doi.org/10.1110/ps.051471205> PMID: 16195556
  54. Yoon S, Welsh WJ: Detecting hidden sequence propensity for amyloid fibril formation. *Protein Science* 2004, 13(8):2149–2160. <https://doi.org/10.1110/ps.04790604> PMID: 15273309
  55. Pawlicki S, Le Behec A, Delamarche C: AMYPdb: a database dedicated to amyloid precursor proteins. *BMC Bioinformatics* 2008, 9:273. <https://doi.org/10.1186/1471-2105-9-273> PMID: 18544157

56. Thangakani AM, Nagarajan R, Kumar S, Sakthivel R, Velmurugan D, Gromiha MM: CPAD, Curated Protein Aggregation Database: A Repository of Manually Curated Experimental Data on Protein and Peptide Aggregation. *PLoS One* 2016, 11(4):e0152949. <https://doi.org/10.1371/journal.pone.0152949> PMID: 27043825
57. Bodi K, Prokaeva T, Spencer B, Eberhard M, Connors LH, Seldin DC: AL-Base: a visual platform analysis tool for the study of amyloidogenic immunoglobulin light chain sequences. *Amyloid* 2009, 16(1):1–8. <https://doi.org/10.1080/13506120802676781> PMID: 19291508
58. The UniProt C: UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 2017, 45(D1):D158–D169. <https://doi.org/10.1093/nar/gkw1099> PMID: 27899622
59. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al: The Protein Data Bank. *Nucleic Acids Res* 2000, 28(1):235–242. <https://doi.org/10.1093/nar/28.1.235> PMID: 10592235
60. [Internet] P: Bethesda (MD): National Library of Medicine (US). In.; 1946.
61. Thangakani AM, Kumar S, Nagarajan R, Velmurugan D, Gromiha MM: GAP: towards almost 100 percent prediction for beta-strand-mediated aggregating peptides with distinct morphologies. *Bioinformatics* 2014, 30(14):1983–1990. <https://doi.org/10.1093/bioinformatics/btu167> PMID: 24681906
62. Lu X, Brickson CR, Murphy RM: TANGO-Inspired Design of Anti-Amyloid Cyclic Peptides. *ACS Chem Neurosci* 2016, 7(9):1264–1274. <https://doi.org/10.1021/acschemneuro.6b00150> PMID: 27347598
63. Louros N, Konstantoulea K, De Vleeschouwer M, Ramakers M, Schymkowitz J, Rousseau F: WALTZ-DB 2.0: an updated database containing structural information of experimentally determined amyloid-forming peptides. *Nucleic Acids Res* 2020, 48(D1):D389–D393. <https://doi.org/10.1093/nar/gkz758> PMID: 31504823
64. Salhi A, Essack M, Radovanovic A, Marchand B, Bougouffa S, Antunes A, et al: DESM: portal for microbial knowledge exploration systems. *Nucleic acids research* 2015:gkv1147. <https://doi.org/10.1093/nar/gkv1147> PMID: 26546514
65. Essack M, Radovanovic A, Bajic VB: Information exploration system for sickle cell disease and repurposing of hydroxyfasudil. *PLoS One* 2013, 8(6):e65190. <https://doi.org/10.1371/journal.pone.0065190> PMID: 23762313
66. Kaur M, Radovanovic A, Essack M, Schaefer U, Maqungo M, Kibler T, et al: Database for exploration of functional context of genes implicated in ovarian cancer. *Nucleic Acids Research* 2009, 37(Database issue):D820–823. <https://doi.org/10.1093/nar/gkn593> PMID: 18790805
67. Sagar S, Kaur M, Dawe A, Seshadri SV, Christoffels A, Schaefer U, et al: DDESC: Dragon database for exploration of sodium channels in human. *BMC Genomics* 2008, 9:622. <https://doi.org/10.1186/1471-2164-9-622> PMID: 19099596
68. Dawe AS, Radovanovic A, Kaur M, Sagar S, Seshadri SV, Schaefer U, et al: DESTAF: a database of text-mined associations for reproductive toxins potentially affecting human fertility. *Reproductive Toxicology* 2012, 33(1):99–105. <https://doi.org/10.1016/j.reprotox.2011.12.007> PMID: 22198179
69. Kwofie SK, Radovanovic A, Sundararajan VS, Maqungo M, Christoffels A, Bajic VB: Dragon exploratory system on hepatitis C virus (DESHCV). *Infection, Genetics and Evolution* 2011, 11(4):734–739. <https://doi.org/10.1016/j.meegid.2010.12.006> PMID: 21194573
70. Salhi A, Essack M, Alam T, Bajic VP, Ma L, Radovanovic A, et al: DES-ncRNA: A knowledgebase for exploring information about human micro and long noncoding RNAs based on literature-mining. *RNA Biol* 2017, 14(7):963–971. <https://doi.org/10.1080/15476286.2017.1312243> PMID: 28387604
71. Essack M, Radovanovic A, Schaefer U, Schmeier S, Seshadri SV, Christoffels A, et al: DDEC: Dragon database of genes implicated in esophageal cancer. *BMC Cancer* 2009, 9:219. <https://doi.org/10.1186/1471-2407-9-219> PMID: 19580656
72. Maqungo M, Kaur M, Kwofie SK, Radovanovic A, Schaefer U, Schmeier S, et al: DDPC: Dragon Database of Genes associated with Prostate Cancer. *Nucleic Acids Research* 2011, 39(Database issue):D980–985. <https://doi.org/10.1093/nar/gkq849> PMID: 20880996
73. Bajic VB, Veronika M, Veladandi PS, Meka A, Heng M-W, Rajaraman K, et al: Dragon Plant Biology Explorer. A text-mining tool for integrating associations between genetic and biochemical entities with genome annotation and biochemical terms lists. *Plant physiology* 2005, 138(4):1914–1925. <https://doi.org/10.1104/pp.105.060863> PMID: 16172098
74. Essack M, Salhi A, Stanimirovic J, Tifratene F, Bin Raies A, Hungler A, et al: Literature-Based Enrichment Insights into Redox Control of Vascular Biology. *Oxid Med Cell Longev* 2019, 2019:1769437. <https://doi.org/10.1155/2019/1769437> PMID: 31223421
75. Salhi A, Negrao S, Essack M, Morton MJL, Bougouffa S, Razali R, et al: DES-TOMATO: A Knowledge Exploration System Focused On Tomato Species. *Sci Rep* 2017, 7(1):5968. <https://doi.org/10.1038/s41598-017-05448-0> PMID: 28729549

76. Maglott D, Ostell J, Pruitt KD, Tatusova T: Entrez Gene: gene-centered information at NCBI. *Nucleic acids research* 2005, 33(suppl 1):D54–D58. <https://doi.org/10.1093/nar/gki031> PMID: 15608257
77. Consortium U: Activities at the universal protein resource (UniProt). *Nucleic acids research* 2014, 42(D1):D191–D198.
78. Hastings J, de Matos P, Dekker A, Ennis M, Harsha B, Kale N, et al: The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Res* 2013, 41(Database issue):D456–463. <https://doi.org/10.1093/nar/gks1146> PMID: 23180789
79. Cotter D, Maer A, Guda C, Saunders B, Subramaniam S: LMPD: LIPID MAPS proteome database. *Nucleic Acids Res* 2006, 34(Database issue):D507–510. <https://doi.org/10.1093/nar/gkj122> PMID: 16381922
80. Sud M, Fahy E, Cotter D, Brown A, Dennis EA, Glass CK, et al: LMSD: LIPID MAPS structure database. *Nucleic Acids Res* 2007, 35(Database issue):D527–532. <https://doi.org/10.1093/nar/gkl838> PMID: 17098933
81. Kale NS, Haug K, Conesa P, Jayseelan K, Moreno P, Rocca-Serra P, et al: MetaboLights: An Open-Access Database Repository for Metabolomics Data. *Curr Protoc Bioinformatics* 2016, 53:14.13.11–14.13.18. <https://doi.org/10.1002/0471250953.bi1413s53> PMID: 27010336
82. Wishart D, Arndt D, Pon A, Sajed T, Guo AC, Djoumbou Y, et al: T3DB: the toxic exposome database. *Nucleic Acids Res* 2015, 43(Database issue):D928–934. <https://doi.org/10.1093/nar/gku1004> PMID: 25378312
83. Gene Ontology C: Gene Ontology Consortium: going forward. *Nucleic Acids Res* 2015, 43(Database issue):D1049–1056. <https://doi.org/10.1093/nar/gku1179> PMID: 25428369
84. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M: KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 1999, 27(1):29–34. <https://doi.org/10.1093/nar/27.1.29> PMID: 9847135
85. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al: The Reactome pathway Knowledgebase. *Nucleic Acids Res* 2016, 44(D1):D481–487. <https://doi.org/10.1093/nar/gkv1351> PMID: 26656494
86. Morgat A, Coissac E, Coudert E, Axelsen KB, Keller G, Bairoch A.: UniPathway: a resource for the exploration and annotation of metabolic pathways. *Nucleic Acids Res* 2012, 40(Database issue):D761–769. <https://doi.org/10.1093/nar/gkr1023> PMID: 22102589
87. Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S, et al: The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res* 2005, 33(Database issue):D284–288. <https://doi.org/10.1093/nar/gki078> PMID: 15608197
88. Kibbe WA, Arze C, Felix V, Mitra E, Bolton E, Fu G, et al: Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Res* 2015, 43(Database issue):D1071–1078. <https://doi.org/10.1093/nar/gku1011> PMID: 25348409
89. Kohler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Ayme S, et al: The Human Phenotype Ontology in 2017. *Nucleic Acids Res* 2017, 45(D1):D865–D876. <https://doi.org/10.1093/nar/gkw1039> PMID: 27899602
90. Kuhn M, Letunic I, Jensen LJ, Bork P: The SIDER database of drugs and side effects. *Nucleic Acids Res* 2016, 44(D1):D1075–1079. <https://doi.org/10.1093/nar/gkv1075> PMID: 26481350
91. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al: DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 2018, 46(D1):D1074–D1082. <https://doi.org/10.1093/nar/gkx1037> PMID: 29126136
92. Maglott D, Ostell J, Pruitt KD, Tatusova T: Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res* 2011, 39(Database issue):D52–57. <https://doi.org/10.1093/nar/gkq1237> PMID: 21115458
93. Yates B, Braschi B, Gray KA, Seal RL, Tweedie S, Bruford EA: Genenames.org: the HGNC and VGNC resources in 2017. *Nucleic Acids Res* 2017, 45(D1):D619–D625. <https://doi.org/10.1093/nar/gkw1033> PMID: 27799471
94. Maglott D, Ostell J, Pruitt KD, Tatusova T: Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res* 2005, 33(Database issue):D54–58. <https://doi.org/10.1093/nar/gki031> PMID: 15608257
95. Schmeier S, Alam T, Essack M, Bajic VB: Tcof-DB v2: update of the database of human and mouse transcription co-factors and transcription factor interactions. *Nucleic Acids Res* 2017, 45(D1):D145–D150. <https://doi.org/10.1093/nar/gkw1007> PMID: 27789689
96. Wei CH, Harris BR, Kao HY, Lu Z: tmVar: a text mining approach for extracting sequence variants in biomedical literature. *Bioinformatics* 2013, 29(11):1433–1439. <https://doi.org/10.1093/bioinformatics/btt156> PMID: 23564842

97. Masters SL, O'Neill LA: Disease-associated amyloid and misfolded protein aggregates activate the inflammasome. *Trends in molecular medicine* 2011, 17(5):276–282. <https://doi.org/10.1016/j.molmed.2011.01.005> PMID: 21376667
98. Halle A, Hornung V, Petzold GC, Stewart CR, Monks BG, Reinheckel T, et al: The NALP3 inflammasome is involved in the innate immune response to amyloid- $\beta$ . *Nature immunology* 2008, 9(8):857–865. <https://doi.org/10.1038/ni.1636> PMID: 18604209
99. Masters SL, Dunne A, Subramanian SL, Hull RL, Tannahill GM, Sharp FA, et al: Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1 [beta] in type 2 diabetes. *Nature immunology* 2010, 11(10):897–904. <https://doi.org/10.1038/ni.1935> PMID: 20835230
100. Codolo G, Plotegher N, Pozzobon T, Brucale M, Tessari I, Bubacco L, et al: Triggering of Inflammasome by Aggregated  $\alpha$ -Synuclein, an Inflammatory Response in Synucleinopathies. *PLoS one* 2013, 8(1):e55375. <https://doi.org/10.1371/journal.pone.0055375> PMID: 23383169
101. Meissner F, Molawi K, Zychlinsky A: Mutant superoxide dismutase 1-induced IL-1 $\beta$  accelerates ALS pathogenesis. *Proceedings of the National Academy of Sciences* 2010, 107(29):13046–13050. <https://doi.org/10.1073/pnas.1002396107> PMID: 20616033
102. Kimura T, Jain A, Choi SW, Mandell MA, Schroder K, Johansen T, et al: TRIM-mediated precision autophagy targets cytoplasmic regulators of innate immunity. *J Cell Biol* 2015, 210(6):973–989. <https://doi.org/10.1083/jcb.201503023> PMID: 26347139
103. Man SM, Kanneganti T-D: Regulation of lysosomal dynamics and autophagy by CTSB/cathepsin B. *Autophagy* 2016, 12(12):2504–2505. <https://doi.org/10.1080/15548627.2016.1239679> PMID: 27786577
104. Chiu HW, Chen CH, Chang JN, Chen CH, Hsu YH: Far-infrared promotes burn wound healing by suppressing NLRP3 inflammasome caused by enhanced autophagy. *J Mol Med (Berl)* 2016, 94(7):809–819.
105. Hornung V, Bauernfeind F, Halle A, Samstad EO, Kono H, Rock KL, et al: Silica crystals and aluminum salts activate the NALP3 inflammasome through phagosomal destabilization. *Nat Immunol* 2008, 9(8):847–856. <https://doi.org/10.1038/ni.1631> PMID: 18604214
106. Hook VY, Kindy M, Hook G: Inhibitors of cathepsin B improve memory and reduce  $\beta$ -amyloid in transgenic Alzheimer disease mice expressing the wild-type, but not the Swedish mutant,  $\beta$ -secretase site of the amyloid precursor protein. *Journal of Biological Chemistry* 2008, 283(12):7745–7753. <https://doi.org/10.1074/jbc.M708362200> PMID: 18184658
107. Baker KR, Rice L: The amyloidoses: clinical features, diagnosis and treatment. *Methodist Debaquey Cardiovasc J* 2012, 8(3):3–7. <https://doi.org/10.14797/mdcj-8-3-3> PMID: 23227278
108. Schipper HM, Maes OC, Chertkow HM, Wang E: MicroRNA expression in Alzheimer blood mononuclear cells. *Gene Regul Syst Bio* 2007, 1:263–274. <https://doi.org/10.4137/grsb.s361> PMID: 19936094
109. Wang M, Qin L, Tang B: MicroRNAs in Alzheimer's Disease. *Front Genet* 2019, 10:153. <https://doi.org/10.3389/fgene.2019.00153> PMID: 30881384
110. Keller A SC, Meese E, Kappel A, Backes C, Leidinger P, inventors; Siemens AG, assignee.: Diagnostic miRNA markers for Alzheimer. United States patent US 10,138,520 2018 Nov 27.
111. Wang Z, Qin W, Wei CB, Tang Y, Zhao LN, Jin HM, et al: The microRNA-1908 up-regulation in the peripheral blood cells impairs amyloid clearance by targeting ApoE. *Int J Geriatr Psychiatry* 2018, 33(7):980–986. <https://doi.org/10.1002/gps.4881> PMID: 29635818
112. Falsone SF: The yin and yang of amyloid aggregation. *Future Sci OA* 2015, 1(2):FSO40. <https://doi.org/10.4155/fso.15.40> PMID: 28031869
113. Falsone A, Falsone SF: Legal but lethal: functional protein aggregation at the verge of toxicity. *Front Cell Neurosci* 2015, 9:45. <https://doi.org/10.3389/fncel.2015.00045> PMID: 25741240
114. Petrof Chapa RD, Emery MA, Fawver JN, Murray IV: Amyloids as Sensors and Protectors (ASAP) hypothesis. *J Alzheimers Dis* 2012, 29(3):503–514. <https://doi.org/10.3233/JAD-2012-112015> PMID: 22330832
115. Qiang W, Yau WM, Lu JX, Collinge J, Tycko R: Structural variation in amyloid-beta fibrils from Alzheimer's disease clinical subtypes. *Nature* 2017, 541(7636):217–221. <https://doi.org/10.1038/nature20814> PMID: 28052060
116. Westermark GT, Fandrich M, Lundmark K, Westermark P: Noncerebral Amyloidoses: Aspects on Seeding, Cross-Seeding, and Transmission. *Cold Spring Harb Perspect Med* 2018, 8(1).
117. Westermark P: Amyloid in the islets of Langerhans: thoughts and some historical aspects. *Ups J Med Sci* 2011, 116(2):81–89. <https://doi.org/10.3109/03009734.2011.573884> PMID: 21486192



118. Lee HM, Kim JJ, Kim HJ, Shong M, Ku BJ, Jo EK: Upregulated NLRP3 inflammasome activation in patients with type 2 diabetes. *Diabetes* 2013, 62(1):194–204. <https://doi.org/10.2337/db12-0420> PMID: 23086037
119. Cai X, Chen J, Xu H, Liu S, Jiang QX, Halfmann R, et al: Prion-like polymerization underlies signal transduction in antiviral immune defense and inflammasome activation. *Cell* 2014, 156(6):1207–1222. <https://doi.org/10.1016/j.cell.2014.01.063> PMID: 24630723
120. Pennisi M, Crupi R, Di Paola R, Ontario ML, Bella R, Calabrese EJ, et al: Inflammasomes, hormesis, and antioxidants in neuroinflammation: Role of NLRP3 in Alzheimer disease. *J Neurosci Res* 2017, 95(7):1360–1372. <https://doi.org/10.1002/jnr.23986> PMID: 27862176