Review

A Bibliometric Insight of Genetic Factors in ASD: Emerging Trends and New Developments

Kang Wang [1], Weicheng Duan [1], Yijie Duan [1], Yuxin Yu [2], Xiuyi Chen [3], Yinhui Xu [2], Haihong Chen [4], Hongzhi Huang [1] and Bo Xiong [1,*]

[1] Department of Forensic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China; wangkjustice@hust.edu.cn (K.W.); duanweic@hust.edu.cn (W.D.); duanyijie@hust.edu.cn (Y.D.); huang1995@hust.edu.cn (H.H.)

[2] The First Clinical College, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China; yuyuxin@hust.edu.cn (Y.Y.); xuyinhui@hust.edu.cn (Y.X.)

[3] Key Laboratory of Environment and Health (HUST), Ministry of Education, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China; chenxiuyi@hust.edu.cn

[4] School of Health Policy & Management, Nanjing Medical University, Nanjing 211166, China; chenhaihong@njmu.edu.cn

[*] Correspondence: bxiong@hust.edu.cn; Tel.: +86-27-83691052

Abstract: Autism spectrum disorder (ASD) cases have increased rapidly in recent decades, which is associated with various genetic abnormalities. To provide a better understanding of the genetic factors in ASD, we assessed the global scientific output of the related studies. A total of 2944 studies published between 1997 and 2018 were included by systematic retrieval from the Web of Science (WoS) database, whose scientific landscapes were drawn and the tendencies and research frontiers were explored through bibliometric methods. The United States has been acting as a leading explorer of the field worldwide in recent years. The rapid development of high-throughput technologies and bioinformatics transferred the research method from the traditional classic method to a big data-based pipeline. As a consequence, the focused research area and tendency were also changed, as the contribution of de novo mutations in ASD has been a research hotspot in the past several years and probably will remain one into the near future, which is consistent with the current opinions of the major etiology of ASD. Therefore, more attention and financial support should be paid to the deciphering of the de novo mutations in ASD. Meanwhile, the effective cooperation of multi-research centers and scientists in different fields should be advocated in the next step of scientific research undertaken.

Keywords: bibliometric; ASD; genetics; quantitative analysis; research frontiers

1. Introduction

Autism spectrum disorder (ASD) is considered a multifactorial syndrome which is characterized by three core characteristics, including qualitative impairment in interests and activities, social interaction, and communication paired with repetitive behaviors and may be accompanied by multiple neurological symptoms, such as intellectual and developmental disability (IDD), communication, speech, and language disorder, attention deficit and hyperactivity disorder, and so on [1]. The prevalence of ASD, which affects more than 1% of children globally, has been steadily increasing in recent years [2,3]. In the United States, 78% of families with autistic children were reported to have health care expenditures for their child, where 34% spent more than 3% of their income [4]. Thus, the families will undergo heavy economic burdens once their offspring is diagnosed with ASD [5]. Besides that, a considerable burden has been added to the public health care system and the government by ASD patients throughout their life. For children under 5 years old, ASD was ranked among the 20 leading causes of disability. Among the 5–14 age group of children,

ASD was listed as the fourth leading cause of disability among mental disorders [6,7]. Special education services and parental job absences are the largest cost components for families with ASD children. During adulthood, medical costs are even higher compared with children. Also, residential nursing, supportive living accommodation, and loss of individual productivity contribute to the high burden and costs [8].

The developmental trajectory of the nervous system can be affected by many potential causes at different areas and stages. Given the temporal and spatial complexity of the disorder, examples of pathogenic mechanisms including social deprivation, genetic and metabolic defects [9], immune dysregulation [10], nutritional factors [11], and toxic or environmental factors have been reported [12]. With an estimated heritability of up to 80–90%, genetic factors are a major cause of ASD [13,14]. A series of genetic studies, including cytogenetics, linkage associations, genome-wide association studies (GWAS), and whole genome or exome sequencing, revealed that the genetic architecture of ASD is complicated and heterogeneous [3,15]. With the development and application of next generation sequencing (NGS), whole genome sequencing (WGS), and whole exon sequencing (WES), progress has been made in understanding the genetic causes of ASD. Hundreds of candidate genes of ASDs have been identified, such as, *KMT5B*, *FMR1*, *NAA15*, and *CHD8*. [16–18].

In this study, we aimed to conduct a bibliometric network analysis on the genetic research on ASD, which involved an objective measurement in scientific literature evaluation and aggregating the opinions of multiple scholars working in the field to mitigate researcher bias in reviews of the scientific literature [19]. The main goal of this analysis is to understand the structure, the current state of the art, and the future directions of the genetic studies in the ASD literature through a scientometrics approach. Utilizing statistical and mathematical methods, this study would provide a quantitative and qualitative analysis of the related literatures which can give a general overview on genetic factors in ASD and help grasp the research frontiers for future development [20,21].

2. Materials and Methods

The Web of Science (WoS) database was used for searching the related literature. An initial database was built from WoSCC to retrieve the literature with the followingsearching strategy: ((Topic = ('Autistic Disorder') OR ('Disorder, Autistic') OR ('Disorders, Autistic') OR ('Kanner's Syndrome') OR ('Kanner Syndrome') OR ('Kanners Syndrome') OR ('Autism, Infantile') OR ('Infantile Autism') OR ('Autism') OR ('Autism, Early Infantile') OR ('Early Infantile Autism') OR ('Infantile Autism, Early')). To ensure searching accuracy, the entry searching terms of autism were obtained from a standardized Medical Subject Headings (MeSH) list of searching terms from the National Library of Medicine. This search strategy identified 60,933 records over the whole previous year up to September 16, 2019. Next, we excluded 4771 records published in 2019, narrowed the retrieved literatures to the WoS genetics heredity category, and identified 4336 selected records for further analysis. To detect the original discoveries and minimize the bias as much as possible, we restricted the document types to original articles without restrictions based on language and finally obtained 2944 science literature examples for in-depth study and analysis (Figure 1). Raw data from the WoSCC were initially downloaded, checked, and verified by two independent investigators with different backgrounds. Any divergence was reconciled through discussion, and finally agreements were reached.

VOSViewer [22,23] 1.6.9 and CiteSpace [24,25] 5.5 were used to conduct statistical analyses of the scientific literature in genetic studies of ASD and to convert these references into visualized graphs. We drew the maps of annual and accumulated publication numbers, identified the contributions and collaborations between countries, performed co-citation analyses on references, and analyzed co-occurrence terms and burst keywords to detect the hotspots and frontiers.
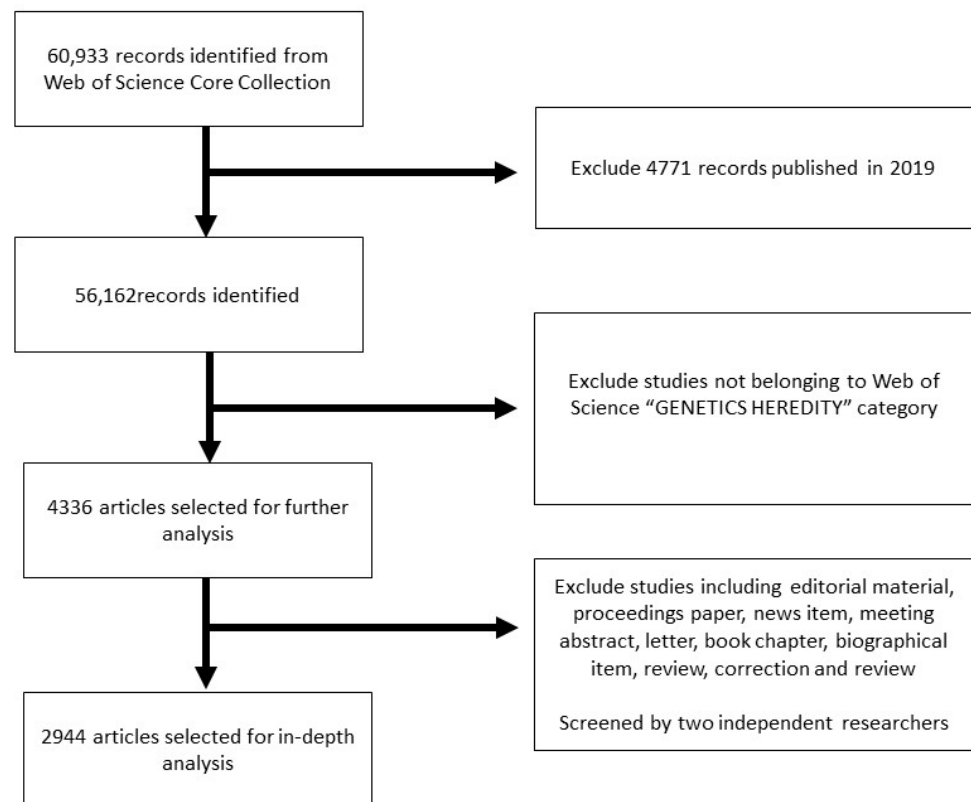
Figure 1. The flow chart of the methodology.

3. Results

3.1. Analysis of Quantity and Growth Trend of Annual Publications

We performed a thorough search of the literature regarding genetic studies of ASD and retrieved 2944 items in total. The quantity of annually published studies was continuously increasing before 2014 and relatively stable from 2015 to 2018, with the average number of annually published literature as approximately 134 (Figure 2). In the 22-year period, the growth rate fluctuated, while the annual number of publications increased over time and the publication number increased nearly 17-fold from 17 records in 1997 to 295 in 2018.
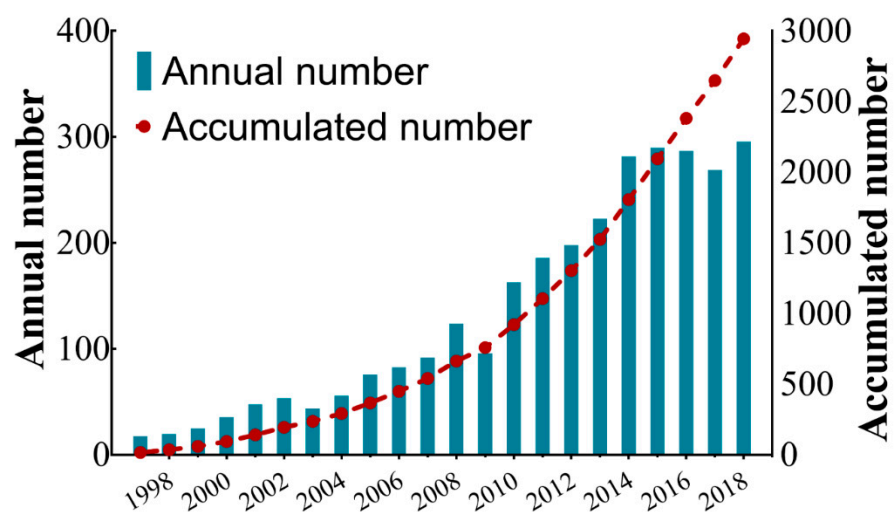

Figure 2. Accumulated number and annual number of genetic factors in autism spectrum disorder (ASD) publications from 1997 to 2018.

3.2. Leading Countries and Institutions

We next evaluated the research activities regarding ASD genetics in different countries. The retrieved literature was contributed by over 70 countries, and an intensive cooperation pattern was clearly observed (Figure 3). The number of publications from a country or region is a sensitive indicator that reflects the attention placed and the research strength in the specific research area. In this regard, the United States participated in the greatest number of studies based on the results and maintained close cooperation with other countries with the highest H-index of 130, followed by England (77) and Canada (60) (Table 1). The average number of citations per article for publications in this field was 38.89, and the H-index was 145. The most effervescent research groups were mainly coming from North America and Europe. In addition, East Asia and Oceania contributed a number of research achievements (Figure 4). The top three countries of the citation/article ratios were Sweden (91.20), Germany (68.76), and the Netherlands (63.20). Furthermore, for Canada (55.283%), France (61.728%), and Sweden (58.197%), the major achievements were made by their core institutions. Nine of the top ten most productive countries were categorized as high-income countries, which produced approximately two-thirds (71.5%) of the documents in the field of genetic factors in ASD. These countries played a major role in ASD research and maintained a high degree of collaborations with other countries and regions.
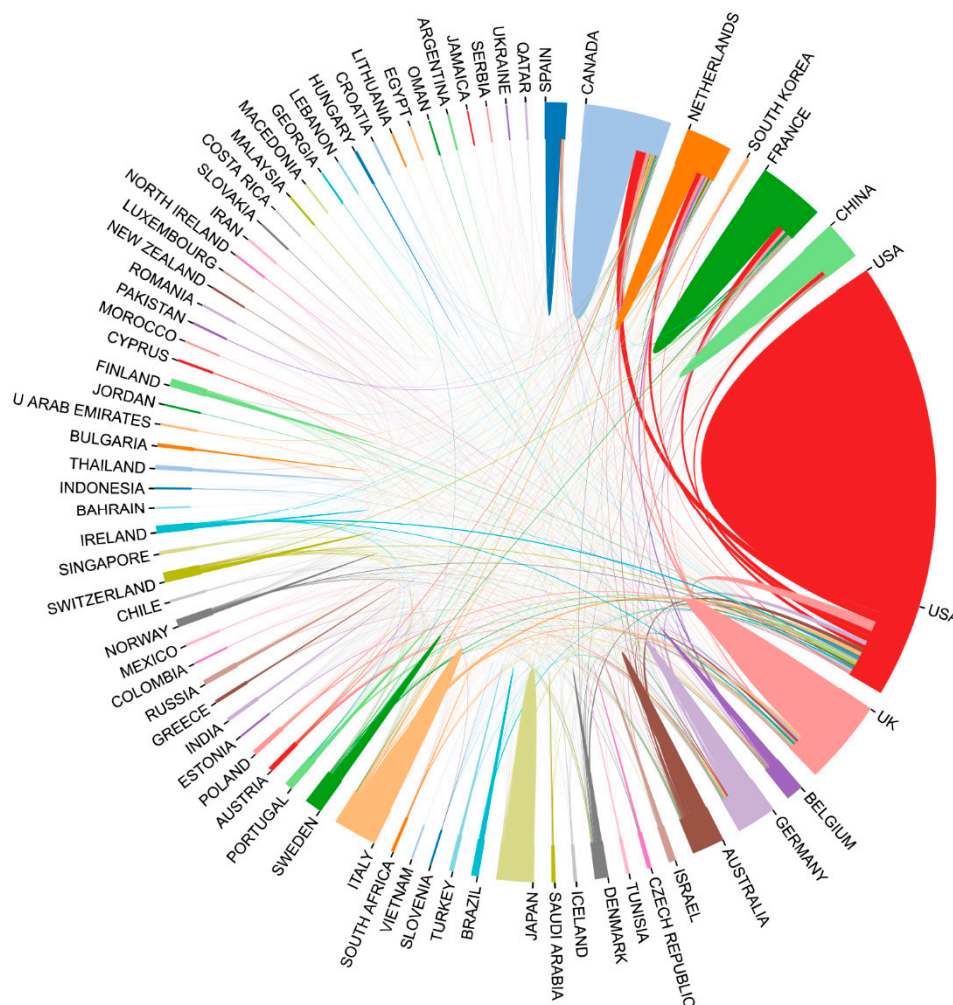

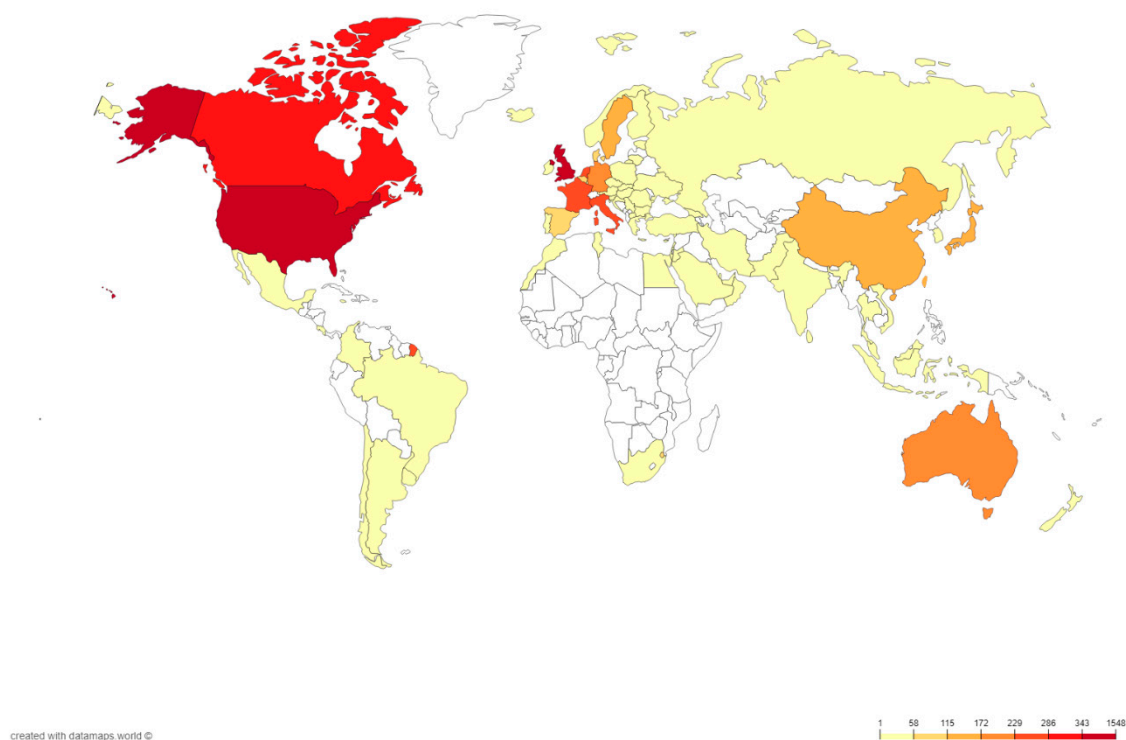**Figure 3.** Cooperation between contributed countries.

Figure 4. Geographical distribution of genetic factors in ASD publications.

Table 1. Top countries and country institutions for autism genetics research.

Country	Articles	Citations	H-Index	Citations per Article	Top Country Institution	Top Institution Articles (%)
USA	1548	79,352	130	51.26	University of California System	304 (19.638%)
England	485	25,568	77	52.72	University of London	186 (38.351%)
Canada	315	17,817	60	56.56	University of Toronto	174 (55.283%)
Netherlands	250	15,799	56	63.20	Radboud University Nijmegen	98 (39.200%)
France	243	14,159	46	58.27	Institut National de La Sante et de La Recherche Medicale Inserm	150 (61.728%)
Italy	242	13,335	54	55.10	IRCCS Fondazione Stella Maris	32 (13.223%)
Germany	223	15,333	54	68.76	Helmholtz Association	52 (23.318%)
Australia	176	8247	45	46.86	University of Melbourne	39 (22.159%)
China	148	3169	29	21.41	Fudan University	22 (14.856%)
Sweden	122	11,126	40	91.20	Karolinska Institutet	71 (58.197%)

3.3. Most Active Journals and Highly Cited Publications

Being important for exchanging, disseminating, and inheriting scientific findings, academic journals play key roles in the kingdom of science. According to our analyses, more than 120 scientific journals have published literature related to genetic studies of ASD. Based on Bradford's distribution method [26,27], 3 core journals were highlighted (*Molecular Autism*, the *American Journal of Medical Genetics Part A*, and the *Journal of Intellectual Disability Research*), and the top 10 most productive journals were identified (Table 2). We retrieved the impact factor (2018), quartile, and Web of Science categories of these journals from the JCR database. Two of the three core journals were located in JCR quartile one, which are considered high-quality scientific publications in the JCR evaluation system. Different from the other core journals, the *Journal of Intellectual Disability Research*, which is an interdisciplinary journal combining biomedical research with social issues, was included in the social science citation index (SSCI) database. As for the top 10 highly cited articles

(Table 3), they were all published in two top-notch journals: *Nature Genetics* and the *American Journal of Human Genetics*, which both had a relatively high impact factor compared with other scientific journals in the field. Interestingly, all top 10 highly cited articles came from the United States, which was consistent with its contribution in the field.

Table 2. The 10 most active journals that published articles on autism genetics research. SJR: SCImago Journal Rank.

Journal	Published Numbers (%)	IF 2018	SJR 2018	JCR Quartile	Categories
Mol Autism	373 (12.67%)	5.712	2.53	Q1	Genetics and Heredity; Neurosciences
Am J Med Genet A	304 (10.33%)	2.179	1.11	Q3	Genetics and Heredity
J Intell Disabil Res	282 (9.58%)	1.941	0.98	Q1	Education, Special (SSCI); Rehabilitation (SSCI)
Am J Med Genet B	192 (6.52%)	3.123	1.56	Q2	Genetics and Heredity; Psychiatry
Hum Mol Genet	176 (5.98%)	4.544	3.10	Q1	Genetics and Heredity; Biochemistry and Molecular Biology
Am J Hum Genet	137 (4.65%)	9.924	6.97	Q1	Genetics and Heredity
Eur J Hum Genet	114 (3.87%)	3.650	1.84	Q2	Genetics and Heredity; Biochemistry and Molecular Biology
J Med Genet	83 (2.82%)	5.899	3.02	Q1	Genetics and Heredity
Am J Med Genet	78 (2.65%)	NA	NA	NA	Genetics and Heredity
Eur J Med Genet	73 (2.48%)	2.022	0.90	Q3	Genetics and Heredity

Table 3. The characteristics of highly cited articles.

Rank	Total Citations	Article Title	Journal	Published Year	Country	IF 2018
1	2901	Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2 [28]	Nat Genet	1999	USA	25.455
2	1984	A general framework for estimating the relative pathogenicity of human genetic variants [29]	Nat Genet	2014	USA	25.455
3	1218	Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies [30]	Am J Hum Genet	2010	USA	9.924
4	1219	Structural variation of chromosomes in autism spectrum disorder [31]	Am J Hum Genet	2008	USA	9.924
5	1082	Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism [32]	Nat Genet	2003	USA	25.455
6	984	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs [33]	Nat Genet	2013	USA	25.455
7	925	Mapping autism risk loci using genetic linkage and chromosomal rearrangements [34]	Nat Genet	2007	USA	25.455
8	847	Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders [35]	Nat Genet	2007	USA	25.455
9	703	A copy number variation morbidity map of developmental delay [36]	Nat Genet	2011	USA	25.455
10	692	Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations [37]	Nat Genet	2011	USA	25.455

3.4. Development Skeleton and Scientific Landscapes of Genetic Factors in Autism

The co-citation knowledge map refers to a network of co-citation publications, which is defined as the frequency at which two documents are cited together. Clusters of co-cited documents provide insights into the specialty structure of knowledge [38]. Therefore, we performed co-citation analysis to detect the literature with high co-citation frequencies in genetic studies of ASD to draw scientific landscapes, as well as to explore the historical development trends and correlation between the literature. According to the results of co-citation analysis, the documents could be clustered into 10 main groups, and a timeline map was generated by CiteSpace (Figure 5). The map indicated that most clusters were concentrated in the period from 2007 to 2016, namely in the latest decade. The earlier studies were mainly devoted to linkage disequilibrium (#3) and ethnicity (#7). Recent research focuses were shifted to de novo mutations (#0), microarray (#1), array comparative genome hybridization (CGH) (#2), termination of pregnancy (#5), gene set analysis (#8), and social behavior (#9), which are involved in the etiology and the intervention of ASD. As one of the most important achievements and powerful tools, studies on the database (#4) had been carried out during the time. Due to the genetic complexity and diversity among different countries, regions, and ethnicities, various databases were built to document the complex causes of ASD.
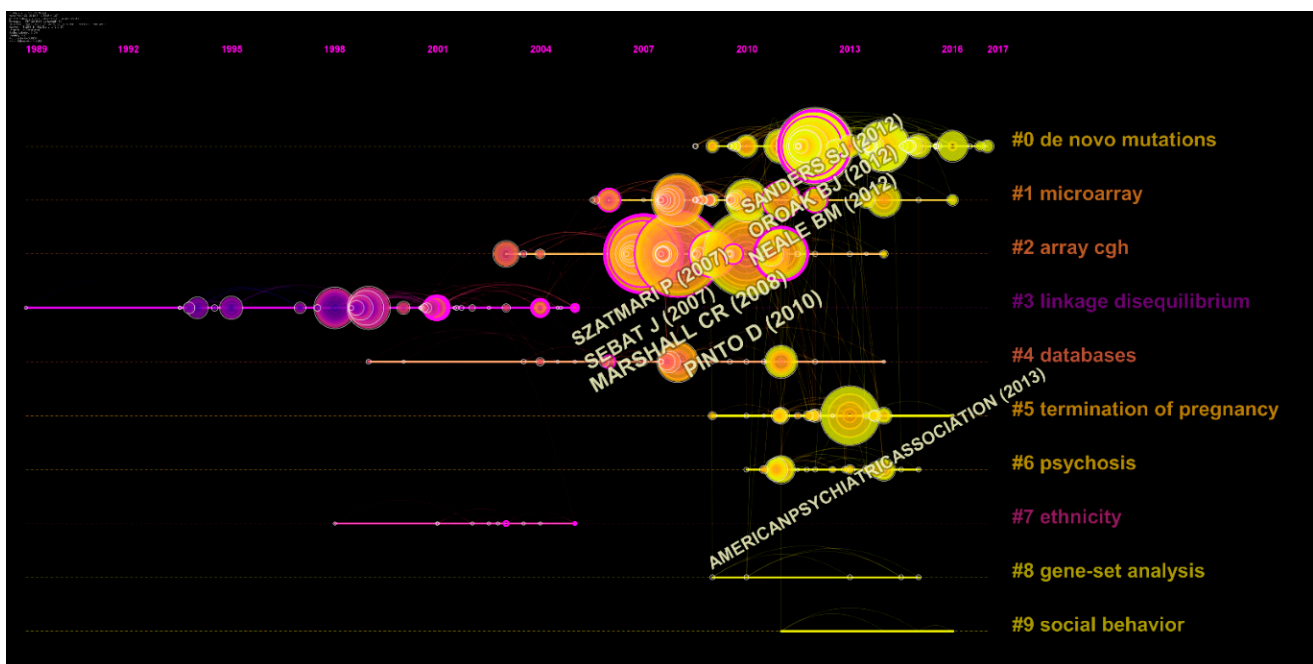

Figure 5. Co-cited references timeline map based on CiteSpace. Co-cited references that are commonly cited in the literature of genetic factors in ASD are clustered and identified by CiteSpace. Nodes on the map represent referenced documents. The shift of research concerns was reflected by the trend of co-cited literature subjects. Years are arranged horizontally at the top, and the label of each cluster is shown at the end of the cluster's timeline.

We then performed bibliometric mapping with the VOSviewer software after excluding repeated and irrelevant items (Figure 6). Through the bibliometric analysis, the terms were automatically divided into four clusters. The red cluster was the largest group, whose terms were more focused on the common clinical symptoms and characteristics of ASD. The distinct clinical heterogeneity, gender, genetics, and comorbidity are recognized as the contributing factors. Clusters in green, blue, and yellow were tightly linked to each other due to their closer physical distance on the graph and were mainly related to ASD-related genetic abnormities, including single nucleotide polymorphisms (SNPs), biomarkers, deletion, or copy number variation (CNVs). The cluster in blue mainly includes the terms

associated with mutations. The terms including CNVs, deletion, and duplication were gathered in the green cluster, and the cluster in yellow mainly covered the terms related to chromosome linkage analysis or SNPs.
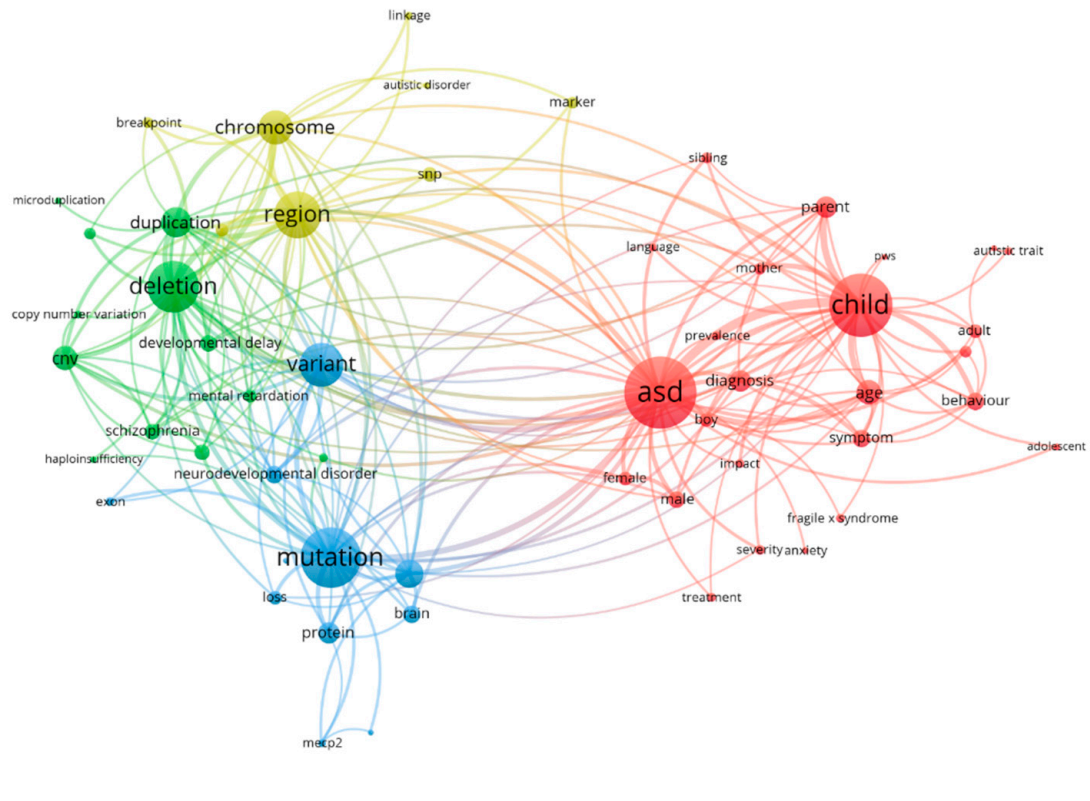

Figure 6. Co-occurrence network of terms of genetic factors in ASD publications from 1997 to 2018. The most common terms used in the scientific literature were investigated to identify relationships between the extracted terms and new aspects and the applied technologies. Terms were automatically exacted from titles and abstracts and divided into four clusters by the natural language processing techniques of VOSviewer.

3.5. Visualization Analysis of Focus Transfer and Research Frontiers

To obtain the frontier topics and research tendency, keywords with the strongest citation bursts in the scientific literature were analyzed and visualized in a keywords burst map by the CiteSpace algorithm-dependent analytical tool, aligned based on their time of appearance (Figure 7). The keywords burst map contained two key points: the strength and beginning or ending year of the burst. The former represented the intensity of the burst, and the latter not only included the duration of the burst time but also revealed the transfer of research focus. Thirty keywords with the strongest citation bursts were included. Before 2000, burst keywords were mainly related to classical research techniques such as twin study (beginning in 1997) and family history study (beginning in 1997). Then, the genomic screen technique (beginning in 2000) and array comparative genome hybridization (array CGH) (beginning in 2007) were gradually developed and extensively applied in the first decade of the twenty-first century. In the latest decade, genome-wide association study (GWAS) (beginning in 2011) started to emerge. As for focused areas, researchers transferred from linkage disequilibrium phenomenon (beginning in 1999) to microdeletion (beginning in 2009) and structural variation (beginning in 2009), and the duration was eight years and six years, respectively. Nowadays, copy number variation (beginning in 2011) and de novo mutation (beginning in 2013) have become the most popular research topics. It was noteworthy that the keyword autism began in 2004 and ended in 2008 with a duration of

5 years, while the keyword autism spectrum disorder appears on the map and continues to the present, which might indicate that the studies were no longer focused on the specific neurological disorder autism, but a set of disorders with the umbrella term autism spectrum disorder. This transformation revealed a deeper understanding and notable progress on the study of ASD.

Top 30 Keywords with the Strongest Citation Bursts

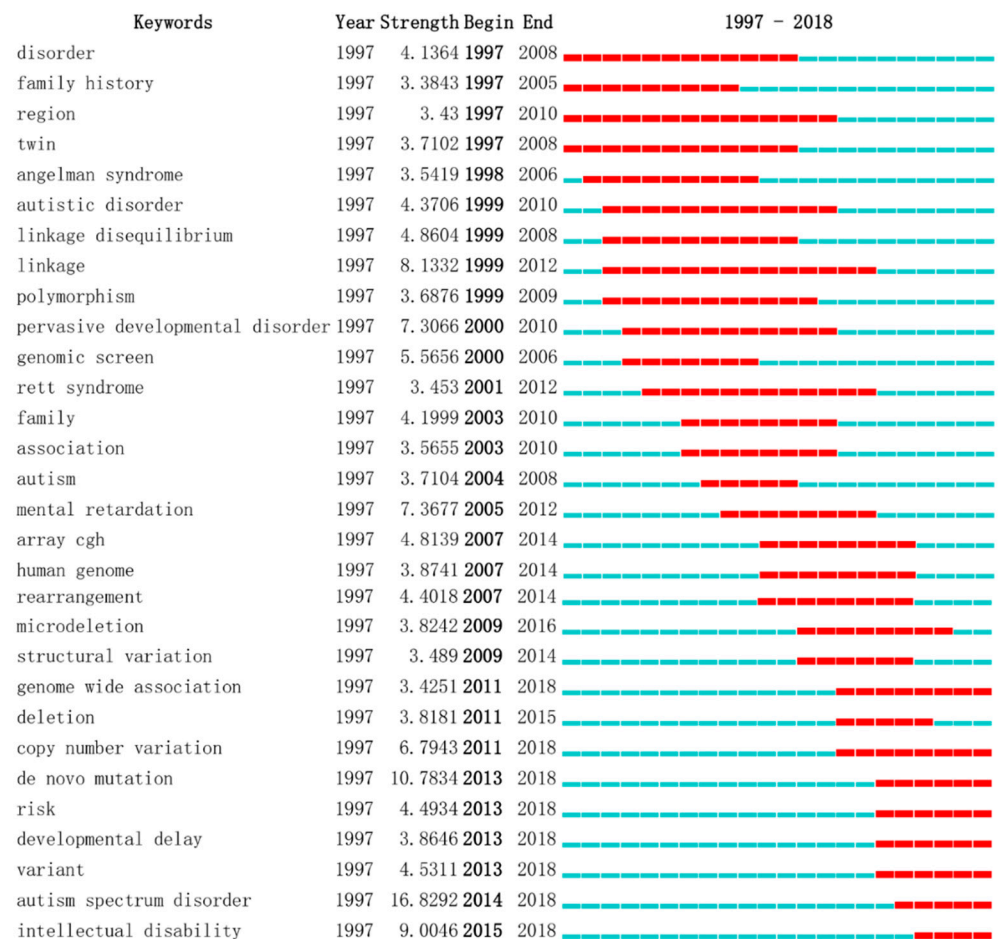| Keywords	Year	Strength	Begin	End	1997 – 2018
disorder	1997	4.1364	1997	2008	
family history	1997	3.3843	1997	2005	
region	1997	3.43	1997	2010	
twin	1997	3.7102	1997	2008	
angelman syndrome	1997	3.5419	1998	2006	
autistic disorder	1997	4.3706	1999	2010	
linkage disequilibrium	1997	4.8604	1999	2008	
linkage	1997	8.1332	1999	2012	
polymorphism	1997	3.6876	1999	2009	
pervasive developmental disorder	1997	7.3066	2000	2010	
genomic screen	1997	5.5656	2000	2006	
rett syndrome	1997	3.453	2001	2012	
family	1997	4.1999	2003	2010	
association	1997	3.5655	2003	2010	
autism	1997	3.7104	2004	2008	
mental retardation	1997	7.3677	2005	2012	
array cgh	1997	4.8139	2007	2014	
human genome	1997	3.8741	2007	2014	
rearrangement	1997	4.4018	2007	2014	
microdeletion	1997	3.8242	2009	2016	
structural variation	1997	3.489	2009	2014	
genome wide association	1997	3.4251	2011	2018	
deletion	1997	3.8181	2011	2015	
copy number variation	1997	6.7943	2011	2018	
de novo mutation	1997	10.7834	2013	2018	
risk	1997	4.4934	2013	2018	
developmental delay	1997	3.8646	2013	2018	
variant	1997	4.5311	2013	2018	
autism spectrum disorder	1997	16.8292	2014	2018	
intellectual disability	1997	9.0046	2015	2018	

Figure 7. Top 30 keywords with the strongest citation bursts. Keywords with the strongest citation bursts in the scientific literature were analyzed and visualized in the keyword's burst map. Each short line represents a year, and a line in red stands for the burst detection years. Keywords with red lines extending to the latest year can indicate the research frontiers a short period of time in the future.

4. Discussion

In this study, we performed a systematic bibliometric assessment of the literature regarding genetic studies of ASD from 1997 to 2018. Previous related studies were mostly focused on ASD rather than its genetic factors and limited publication sources to only one specific country such as Spain, Brazil, or the United States [39–41]. One of the previous studies pointed out that the genetic fields got the most attention and received high numbers of citations within all ASD aspects [42]. Here, we aimed to obtain the landscape overview and the variation trend of the literature regarding the genetic factors of ASD. According to the results, the accumulated publication number had been growing steadily since 1997, when the first such literature appeared. Compared with the lower publication growth of research in forensic anthropology, the increase rate exhibited a relatively higher level,

which indicated the importance of this field [43]. However, the annual publication number has not clearly increased in the last five years. The slow growth of financial investments or limited research positions might partially explain this. In terms of both the contribution number and publication quality, North America could be regarded as an undisputed leading region in this field, followed by European countries and two Asian countries, which was consistent with the medical big data research [44]. Interestingly, there were no publications from other countries except the United States on the top 10 most cited articles list, indicating that the US is leading ASD research. As a fast-developing country, China was listed on the top 10 productive countries. However, the H-index and citations per article were relatively low, indicating a requirement to improve research quality. Similar issues were also observed by bibliometric analysis in other areas, such as IL-35 research and hypertension [45,46].

To explore the branch development venation, the development of citation structures, size of clusters, and scientific landscapes over the past few decades were detected (Figures 5–7). Sequencing technology and post-processing were leading a revolution in biological investigation due to the explosive development of biotechnology, data science tools, and interdisciplinary life science [47]. This led to the massive increase in databases that have been powerful resources for researchers, such as the autism genetic database [48], Simons Simplex Collection [49], and MSSNG [50]. More co-cited references have transferred from array CGH to microarray and finally to de novo mutations from 2007 to 2017 (Figure 5). Early studies clustered in array CGH and microarray mainly focused on CNVs, microdeletions, or microduplications. Many novel loci and ASD genes were identified, such as *SHANK2*, *ANKRD11*, and *DLGAP2* [31,51–53]. Finally, studies turned to de novo mutations. Various studies, which were performed using whole-exome sequencing or the multiplex targeted sequencing technique, revealed that de novo mutations were strongly associated with ASD [54,55]. These academic advances promoted genetic studies and pointed out new directions for researchers.

Using natural language processing techniques, VOSviewer extracts terms from the corpus file, where a term is defined as a sequence of nouns or noun phrases that can be found in a sentence. Over the past 22 years, deletion, variants, chromosomes, mutation, and duplication have become active research topics related to genetic factors of ASD. The four clusters displayed three major genetic research aspects and one major clinical aspect. CNVs, such as deletion or duplication, mainly affect the gene expression levels, which are supposed to be tightly related to developmental delay disease and schizophrenia [56,57]. Another research direction focused on gene mutations or variants. Several key ASD candidate genes were identified by targeted sequencing or WES to play an important role in the pathogenesis, such as *CHD8*, *WAC*, and *POGZ* [16,17,58,59]. There is no single biological marker, single gene, single region of the brain, or pathophysiological mechanism responsible for ASD [60]. Thus, multiple genetic factors were explored to give an account for it.

Different from the terms exacted from an article title or an abstract automatically, keywords are taken from the author-supplied keyword list of a publication. By defining some filed-in details and providing a common vocabulary and a tool with which the evolution of research can be studied, keywords play an important role in scientific research [61]. Burst keyword shifts imply the transformation of research focus and inspiration on research frontiers. The shifting of research hotspots implicates the transformation of the bias of fond assistance, and researchers should have put more spirit and mind to it so as to avoid detours in their careers. In the development of genome research, the appearance of omics is a watershed and of a significance that totally changed the research methods. However, we failed to catch the early research emphases that have not continued into the post-omics era because of the late start date. The keywords that burst until current day indicated the research frontiers and emerging trends of the publications of genetic factors in autism, including genome-wide association, copy number variation, developmental delay, intellectual disability, and so on. Autism spectrum disorder and de novo mutation

were marked as the highest burst strengths along with them. Compared to the burst keyword autism appearing in 2004 and ending in 2008, the autism spectrum disorder burst from 2014 may indicate that the perception of autism transferred from a single disease to a kind of spectrum disorder. At the same time, researchers gained deeper insights into similar development disabilities, with developmental delay rising starting in 2013. With the appearance and rapid development of various high-throughput large-scale sequencing technologies, multiple studies have confirmed the contribution of de novo mutations of many single genes to the risk for autism spectrum disorders [54,62–64]. Transcriptomics studies underlining these high-risk gene knockouts in animals provided potential networks and possible transcriptional mechanisms for the etiology of ASD. Some of them also play a significant role in post-translational modifications or cellular energy metabolism. Their abnormalities can lead to neurodevelopmental disorders and cause psychiatric symptoms. Thus, it can be predicted that de novo mutations will continue to be a research focus in the near future. Finally, the linkage between ASD and intellectual disability has been brought to the attention of research communities since 2015. Increasing numbers of overlapping genes of ASD with other neurological disorders have been discovered recently, and the interactions of these genes and diseases indicate future interdisciplinary research directions. In summary, the burst keywords mainly reflected that the ASD research focused on genomic abnormalities. In comparison, the bibliometric myocardial infarction field mainly focused on microRNA with similar analyses [65].

5. Conclusions

We constructed a series of scientific maps of core journals, most contributions from countries or regions, co-cited references, and burst keywords to show the theme evolution and emerging trends of the ASD research field. As a highly heterogeneous disorder, a major ongoing research trend was to identify the de novo mutations related to the disease. Sequencing technology development and the era of omics led to the rapid growth of data volume. At this juncture, the field of ASD has the potential to draw inspiration and resources from breakthroughs in life science and IT technologies to make dramatic progress toward the understanding of the genetic neurobiology of brain science affecting children. Collaborations between scientists in different fields, such as clinical psychologists, data engineers, and brain scientists, will engender the expertise necessary for leveraging the power of large data sets to further our understanding of ASD.

6. Limitations

There were some limitations to this bibliometric work. First, the content of the science literature database is constantly changing from time to time, and numerous burgeoning scientific achievements were employed. Thus, the dataset on the published date should be broadened. However, compared with the original predominant dataset, its influence on the results was acceptable. Second, the *American Journal of Medical Genetics* had been divided into three journals—the *American Journal of Medical Genetics-part A*, the *American Journal of Medical Genetics-part B*, and the *American Journal of Medical Genetics-part C* by Wiley-Liss Inc.—several years ago. Here, we recognized them as four independent journals and calculated the published articles respectively. This method resulted in blank impact factors, SJR scores, and JCR quartiles of the *American Journal of Medical Genetics* in 2018 (Table 2).

References

1. Association, A.P. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*; American Psychiatric Pub: Washington, DC, USA, 2013; ISBN 0-89042-557-4.
2. Lai, M.-C.; Lombardo, M.V.; Baron-Cohen, S. Autism. *Lancet* **2014**, *383*, 896–910. [CrossRef]
3. Grove, J.; Ripke, S.; Als, T.D.; Mattheisen, M.; Walters, R.K.; Won, H.; Pallesen, J.; Agerbo, E.; Andreassen, O.A.; Anney, R.; et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 431–444. [CrossRef] [PubMed]
4. Parish, S.; Thomas, K.; Rose, R.; Kilany, M.; McConville, R. State Insurance Parity Legislation for Autism Services and Family Financial Burden. *Intellect. Dev. Disabil.* **2012**, *50*, 190–198. [CrossRef] [PubMed]
5. Zablotsky, B.; Kalb, L.G.; Freedman, B.; Vasa, R.; Stuart, E.A. Health Care Experiences and Perceived Financial Impact among Families of Children with an Autism Spectrum Disorder. *Psychiatr. Serv.* **2014**, *65*, 395–398. [CrossRef] [PubMed]
6. Murray, C.J.L.; Ezzati, M.; Flaxman, A.D.; Lim, S.; Lozano, R.; Michaud, C.; Naghavi, M.; Salomon, J.A.; Shibuya, K.; Vos, T.; et al. GBD 2010: Design, definitions, and metrics. *Lancet* **2012**, *380*, 2063–2066. [CrossRef]
7. Baxter, A.J.; Brugha, T.S.; Erskine, H.E.; Scheurer, R.W.; Vos, T.; Scott, J.G. The epidemiology and global burden of autism spectrum disorders. *Psychol. Med.* **2015**, *45*, 601–613. [CrossRef]
8. Buescher, A.V.S.; Cidav, Z.; Knapp, M.; Mandell, D.S. Costs of Autism Spectrum Disorders in the United Kingdom and the United States. *JAMA Pediatr.* **2014**, *168*, 721–728. [CrossRef]
9. Sahin, M.; Sur, M. Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. *Science* **2015**, *350*. [CrossRef]
10. Onore, C.; Careaga, M.; Ashwood, P. The role of immune dysfunction in the pathophysiology of autism. *Brain Behav. Immun.* **2012**, *26*, 383–392. [CrossRef]
11. Bibel, B.M. Autism & Diet: What You Need To Know. *Libr. J.* **2009**, *134*, 95.
12. Hallmayer, J.; Cleveland, S.; Torres, A.; Phillips, J.; Cohen, B.; Torigoe, T.; Miller, J.; Fedele, A.; Collins, J.; Smith, K.; et al. Genetic heritability and shared environmental factors among twin pairs with autism. *Arch. Gen. Psychiatry* **2011**, *68*, 1095–1102. [CrossRef] [PubMed]
13. Van Loo, K.M.J.; Martens, G.J.M. Genetic and environmental factors in complex neurodevelopmental disorders. *Curr. Genom.* **2007**, *8*, 429–444. [CrossRef]
14. Ronald, A.; Hoekstra, R.A. Autism spectrum disorders and autistic traits: A decade of new twin studies. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **2011**, *156*, 255–274. [CrossRef] [PubMed]
15. Murdoch, J.D.; State, M.W. Recent developments in the genetics of autism spectrum disorders. *Curr. Opin. Genet. Dev.* **2013**, *23*, 310–315. [CrossRef]
16. Stessman, H.A.F.; Xiong, B.; Coe, B.P.; Wang, T.; Hoekzema, K.; Fenckova, M.; Kvarnung, M.; Gerdts, J.; Trinh, S.; Cosemans, N.; et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat. Genet.* **2017**, *49*, 515–526. [CrossRef]
17. Bernier, R.; Golzio, C.; Xiong, B.; Stessman, H.A.; Coe, B.P.; Penn, O.; Witherspoon, K.; Gerdts, J.; Baker, C.; Vulto-van Silfhout, A.T.; et al. Disruptive CHD8 mutations define a subtype of autism early in development. *Cell* **2014**, *158*, 263–276. [CrossRef]
18. Budimirovic, D.B.; Schlageter, A.; Filipovic-Sadic, S.; Protic, D.D.; Bram, E.; Mahone, E.M.; Nicholson, K.; Culp, K.; Javanmardi, K.; Kemppainen, J.; et al. A Genotype-Phenotype Study of High-Resolution FMR1 Nucleic Acid and Protein Analyses in Fragile X Patients with Neurobehavioral Assessments. *Brain Sci.* **2020**, *10*, 694. [CrossRef]
19. Zupic, I.; Cater, T. Bibliometric Methods in Management and Organization. *Organ. Res. Methods* **2015**, *18*, 429–472. [CrossRef]
20. Zyoud, S.H.; Waring, W.S.; Al-Jabi, S.W.; Sweileh, W.M. Bibliometric profile of global scientific research on digoxin toxicity (1849–2015). *Drug Chem. Toxicol.* **2018**, 1–7. [CrossRef]
21. Kamdem, J.P.; Duarte, A.E.; Rodrigues Lima, K.R.; Teixeira Rocha, J.B.; Hassan, W.; Barros, L.M.; Roeder, T.; Tsopmo, A. Research trends in food chemistry: A bibliometric review of its 40 years anniversary (1976–2016). *Food Chem.* **2019**, *294*, 448–457. [CrossRef]
22. Van Eck, N.J.; Waltman, L. Visualizing Bibliometric Networks. In *Measuring Scholarly Impact: Methods and Practice*; Ding, Y., Rousseau, R., Wolfram, D., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 285–320. ISBN 978-3-319-10377-8.
23. Perianes-Rodriguez, A.; Waltman, L.; van Eck, N.J. Constructing bibliometric networks: A comparison between full and fractional counting. *J. Informetr.* **2016**, *10*, 1178–1195. [CrossRef]
24. Chen, C. Searching for intellectual turning points: Progressive knowledge domain visualization. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5303–5310. [CrossRef] [PubMed]

25. Chen, C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 359–377. [CrossRef]

26. Brookes, B.C. Bradford's law and the bibliography of science. *Nature* **1969**, *224*, 953–956. [CrossRef] [PubMed]

27. Desai, N.; Veras, L.; Gosain, A. Using Bradford's law of scattering to identify the core journals of pediatric surgery. *J. Surg. Res.* **2018**, *229*, 90–95. [CrossRef] [PubMed]

28. Amir, R.E.; Van den Veyver, I.B.; Wan, M.; Tran, C.Q.; Francke, U.; Zoghbi, H.Y. Rett Syndrome Is Caused by Mutations in X-Linked MECP2, Encoding Methyl-CpG-Binding Protein 2. *Nat. Genet.* **1999**, *23*, 185–188. [CrossRef]

29. Kircher, M.; Witten, D.M.; Jain, P.; O'Roak, B.J.; Cooper, G.M.; Shendure, J. A General Framework for Estimating the Relative Pathogenicity of Human Genetic Variants. *Nat. Genet.* **2014**, *46*, 310–315. [CrossRef]

30. Miller, D.T.; Adam, M.P.; Aradhya, S.; Biesecker, L.G.; Brothman, A.R.; Carter, N.P.; Church, D.M.; Crolla, J.A.; Eichler, E.E.; Epstein, C.J.; et al. Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. *Am. J. Hum. Genet.* **2010**, *86*, 749–764. [CrossRef]

31. Marshall, C.R.; Noor, A.; Vincent, J.B.; Lionel, A.C.; Feuk, L.; Skaug, J.; Shago, M.; Moessner, R.; Pinto, D.; Ren, Y.; et al. Structural Variation of Chromosomes in Autism Spectrum Disorder. *Am. J. Hum. Genet.* **2008**, *82*, 477–488. [CrossRef]

32. Jamain, S.; Quach, H.; Betancur, C.; Råstam, M.; Colineaux, C.; Gillberg, I.C.; Soderstrom, H.; Giros, B.; Leboyer, M.; Gillberg, C.; et al. Mutations of the X-Linked Genes Encoding Neuroligins NLGN3 and NLGN4 Are Associated with Autism. *Nat. Genet.* **2003**, *34*, 27–29. [CrossRef]

33. Cross-Disorder Group of the Psychiatric Genomics Consortium; Lee, S.H.; Ripke, S.; Neale, B.M.; Faraone, S.V.; Purcell, S.M.; Perlis, R.H.; Mowry, B.J.; Thapar, A.; Goddard, M.E.; et al. Genetic Relationship between Five Psychiatric Disorders Estimated from Genome-Wide SNPs. *Nat. Genet.* **2013**, *45*, 984–994. [CrossRef] [PubMed]

34. Autism Genome Project Consortium; Szatmari, P.; Paterson, A.D.; Zwaigenbaum, L.; Roberts, W.; Brian, J.; Liu, X.-Q.; Vincent, J.B.; Skaug, J.L.; Thompson, A.P.; et al. Mapping Autism Risk Loci Using Genetic Linkage and Chromosomal Rearrangements. *Nat. Genet.* **2007**, *39*, 319–328. [CrossRef] [PubMed]

35. Durand, C.M.; Betancur, C.; Boeckers, T.M.; Bockmann, J.; Chaste, P.; Fauchereau, F.; Nygren, G.; Rastam, M.; Gillberg, I.C.; Anckarsäter, H.; et al. Mutations in the Gene Encoding the Synaptic Scaffolding Protein SHANK3 Are Associated with Autism Spectrum Disorders. *Nat. Genet.* **2007**, *39*, 25–27. [CrossRef] [PubMed]

36. Cooper, G.M.; Coe, B.P.; Girirajan, S.; Rosenfeld, J.A.; Vu, T.H.; Baker, C.; Williams, C.; Stalker, H.; Hamid, R.; Hannig, V.; et al. A Copy Number Variation Morbidity Map of Developmental Delay. *Nat. Genet.* **2011**, *43*, 838–846. [CrossRef]

37. O'Roak, B.J.; Deriziotis, P.; Lee, C.; Vives, L.; Schwartz, J.J.; Girirajan, S.; Karakoc, E.; Mackenzie, A.P.; Ng, S.B.; Baker, C.; et al. Exome Sequencing in Sporadic Autism Spectrum Disorders Identifies Severe de Novo Mutations. *Nat. Genet.* **2011**, *43*, 585–589. [CrossRef]

38. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *42*, 676–684. [CrossRef]

39. Belinchón Carmona, M.; Boada Muñoz, L.; García de Andrés, E.; Fuentes Biggi, J.; Posada de la Paz, M. Trends in studies on autism in Spain: Publications and authorship networks (1974–2007). *Psicothema* **2010**, *22*, 242–249.

40. Teixeira, M.C.T.V.; Mecca, T.P.; Velloso, R.D.L.; Bravo, R.B.; Ribeiro, S.H.B. Brazilian scientific output on autism spectrum disorders. *Rev. Assoc. Med. Bras. (1992)* **2010**, *56*, 607–614. [CrossRef]

41. Whyatt, C.P.; Torres, E.B. Autism Research: An Objective Quantitative Review of Progress and Focus between 1994 and 2015. *Front. Psychol.* **2018**, *9*, 1526. [CrossRef]

42. Sweileh, W.M.; Al-Jabi, S.W.; Sawalha, A.F.; Zyoud, S.H. Bibliometric profile of the global scientific research on autism spectrum disorders. *Springerplus* **2016**, *5*, 1480. [CrossRef]

43. Lei, G.; Liu, F.; Liu, P.; Zhou, Y.; Jiao, T.; Dang, Y.-H. Worldwide tendency and focused research in forensic anthropology: A bibliometric analysis of decade (2008–2017). *Leg. Med.* **2019**, *37*, 67–75. [CrossRef] [PubMed]

44. Liao, H.; Tang, M.; Luo, L.; Li, C.; Chiclana, F.; Zeng, X.-J. A Bibliometric Analysis and Visualization of Medical Big Data Research. *Sustainability* **2018**, *10*, 166. [CrossRef]

45. Cai, X.; Zhou, C.; Zhou, L.; Xu, Q. A bibliometric analysis of IL-35 research from 2009 to 2018. *PeerJ* **2019**, *7*, e7992. [CrossRef] [PubMed]

46. Devos, P.; Menard, J. Bibliometric analysis of research relating to hypertension reported over the period 1997–2016. *J. Hypertens.* **2019**, *37*, 2116–2122. [CrossRef]

47. Van Heyningen, V. Genome sequencing-the dawn of a game-changing era. *Heredity* **2019**, *123*, 58–66. [CrossRef]

48. Matuszek, G.; Talebizadeh, Z. Autism genetic database (AGD): A comprehensive database including autism susceptibility gene-CNVs integrated with known noncoding RNAs and fragile sites. *BMC Med. Genet.* **2009**, *10*, 102. [CrossRef]

49. Fischbach, G.D.; Lord, C. The Simons Simplex Collection: A Resource for Identification of Autism Genetic Risk Factors. *Neuron* **2010**, *68*, 192–195. [CrossRef]

50. Yuen, R.K.C.; Merico, D.; Bookman, M.; Howe, J.L.; Thiruvahindrapuram, B.; Patel, R.V.; Whitney, J.; Deflaux, N.; Bingham, J.; Wang, Z.; et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat. Neurosci.* **2017**, *20*, 602–611. [CrossRef]

51. Weiss, L.A.; Shen, Y.; Korn, J.M.; Arking, D.E.; Miller, D.T.; Fossdal, R.; Saemundsen, E.; Stefansson, H.; Ferreira, M.A.R.; Green, T.; et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N. Engl. J. Med.* **2008**, *358*, 667–675. [CrossRef]

52. Pinto, D.; Pagnamenta, A.T.; Klei, L.; Anney, R.; Merico, D.; Regan, R.; Conroy, J.; Magalhaes, T.R.; Correia, C.; Abrahams, B.S.; et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **2010**, *466*, 368–372. [CrossRef]

53. Sebat, J.; Lakshmi, B.; Malhotra, D.; Troge, J.; Lese-Martin, C.; Walsh, T.; Yamrom, B.; Yoon, S.; Krasnitz, A.; Kendall, J.; et al. Strong association of de novo copy number mutations with autism. *Science* **2007**, *316*, 445–449. [CrossRef] [PubMed]

54. Sanders, S.J.; Murtha, M.T.; Gupta, A.R.; Murdoch, J.D.; Raubeson, M.J.; Willsey, A.J.; Ercan-Sencicek, A.G.; DiLullo, N.M.; Parikshak, N.N.; Stein, J.L.; et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **2012**, *485*, 237–241. [CrossRef] [PubMed]

55. O'Roak, B.J.; Vives, L.; Fu, W.; Egertson, J.D.; Stanaway, I.B.; Phelps, I.G.; Carvill, G.; Kumar, A.; Lee, C.; Ankenman, K.; et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* **2012**, *338*, 1619–1622. [CrossRef]

56. Drakesmith, M.; Parker, G.D.; Smith, J.; Linden, S.C.; Rees, E.; Williams, N.; Owen, M.J.; van den Bree, M.; Hall, J.; Jones, D.K.; et al. Genetic risk for schizophrenia and developmental delay is associated with shape and microstructure of midline white-matter structures. *Transl. Psychiatry* **2019**, *9*, 102. [CrossRef] [PubMed]

57. Marshall, C.R.; Howrigan, D.P.; Merico, D.; Thiruvahindrapuram, B.; Wu, W.; Greer, D.S.; Antaki, D.; Shetty, A.; Holmans, P.A.; Pinto, D.; et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat. Genet.* **2017**, *49*, 27–35. [CrossRef]

58. DeSanto, C.; D'Aco, K.; Araujo, G.C.; Shannon, N.; Study, D.; Vernon, H.; Rahrig, A.; Monaghan, K.G.; Niu, Z.; Vitazka, P.; et al. *WAC* loss-of-function mutations cause a recognisable syndrome characterised by dysmorphic features, developmental delay and hypotonia and recapitulate 10p11.23 microdeletion syndrome. *J. Med. Genet.* **2015**, *52*, 754–761. [CrossRef]

59. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* **2015**, *519*, 223–228. [CrossRef]

60. Santangelo, S.L.; Tsatsanis, K. What is known about autism—Genes, brain, and behavior. *Am. J. Pharm.* **2005**, *5*, 71–92. [CrossRef]

61. Barki, H.; Rivard, S.; Talbot, J. A Keyword Classification Scheme for Is Research Literature—An Update. *MIS Q.* **1993**, *17*, 209–226. [CrossRef]

62. Alonso-Gonzalez, A.; Rodriguez-Fontenla, C.; Carracedo, A. De novo Mutations (DNMs) in Autism Spectrum Disorder (ASD): Pathway and Network Analysis. *Front. Genet.* **2018**, *9*, 406. [CrossRef]

63. Duan, W.; Wang, K.; Duan, Y.; Chu, X.; Ma, R.; Hu, P.; Xiong, B. Integrated Transcriptome Analyses Revealed Key Target Genes in Mouse Models of Autism. *Autism Res.* **2020**, *13*, 352–368. [CrossRef] [PubMed]

64. Rahman, M.R.; Petralia, M.C.; Ciurleo, R.; Bramanti, A.; Fagone, P.; Shahjaman, M.; Wu, L.; Sun, Y.; Turanli, B.; Arga, K.Y.; et al. Comprehensive Analysis of RNA-Seq Gene Expression Profiling of Brain Transcriptomes Reveals Novel Genes, Regulators, and Pathways in Autism Spectrum Disorder. *Brain Sci.* **2020**, *10*, 747. [CrossRef] [PubMed]

65. Zhou, H.; Tan, W.; Qiu, Z.; Song, Y.; Gao, S. A bibliometric analysis in gene research of myocardial infarction from 2001 to 2015. *PeerJ* **2018**, *6*, e4354. [CrossRef] [PubMed]