World Journal of Materials Science
ISSN: 2960-0146
DOI: 10.61784/wjms240168

APPLICATION PROGRESS OF MATERIALS GENOME TECHNOLOGY IN THE FIELD OF NEW ENERGY MATERIALS

Kevin Ortner
University of British Columbia, Vancouver, British Columbia V6T 1Z2, Canada.

Abstract: Materials genome integrates high-throughput computing, high-throughput preparation, high-throughput detection and database systems of materials. It is a "paradigm revolution" in materials research and development. With its profound scientific connotation and significant application potential, it will accelerate New materials discovery and applications. This article focuses on the use of materials genome in the research and development of new energy materials to shorten the "discovery-development-production-application" cycle of new energy materials. It introduces the internationally representative Materials Project and OQMD two material genome platforms, as well as some important the application of materials genome computing technologies, such as material conformation characterization, high-throughput computing and screening, machine learning, neural network technology, optimization algorithms and new high-throughput preparation and characterization technologies, in the research and development of new energy materials, and the next step the development of materials genome puts forward prospects, such as developing high-precision high-throughput computing, using artificial intelligence to develop high-throughput experimental systems and platforms, generating material big data, and making full use of material big data through intelligent computing to create computing and experiments. The integrated materials genome big data artificial intelligence system accelerates the discovery and application of new energy materials.
Keywords: Materials genome; New energy materials; High-throughput computing; High-throughput experiment

1 CURRENT STATUS OF DEVELOPMENT AND APPLICATION OF MATERIALS GENOMICS TECHNOLOGY

Materials are the foundation and forerunner of all industries. The time span from research and development to putting materials into the market is extremely long. The key is that materials research and development have long relied too much on scientific intuition and trial-and-error experimental experience accumulation, and the preparation process is long and full of variables. It has become a common pursuit of all countries in the world to change the way of research and development of materials and improve the speed of materials from discovery to application. In June 2011, the United States launched the Materials Genome Initiative (MGI). In December 2014, the Science and Technology Committee directly under the President of the United States issued an upgraded version of the Materials Genome Initiative Strategic Plan [1]. Its core content is to establish high-throughput materials calculation methods, high-throughput materials experimental methods and materials databases. An important change is to implement "the discovery of new materials starts with design", focusing on understanding, design and calculation at the atomic and molecular levels. New materials, and collect the correlation between the structures and properties of existing materials through a database to guide the design and development of new materials. Its purpose and significance are: ① Accelerate the material research and development process through high-throughput screening of new materials; ② Transform the new material research and development paradigm to save manpower and material resources; ③ Accelerate human understanding of the nature and laws of materials, and then use the materials they know to The essence and laws guide the research and development of new materials such as design, preparation and testing, so as to further understand the nature and laws of materials; ④ Build a reliable materials genome database (new material big data) to realize resource sharing and accelerate the discovery and application of new materials. It should be said that a cover article [4] published in the main issue of Nature in May 2016 that used experimental waste data to guide the successful discovery of materials through machine learning and data mining is the best explanation of the effectiveness of the application of materials genome technology [2].

Shorten the "discovery-development-production-application" cycle of new materials In this period, reducing the human and material costs in material research and development is the key to implementing material The fundamental purpose of the genome project is also to realize China's development in the field of new materials. In line with the inherent demand for leapfrog development, starting in 2016, our country for the first time Materials genome engineering and technology is included in the national key research and development plan, among which It mainly includes three aspects of research: research and development of common key technologies; Application of typical materials and construction of software and hardware platforms. Domestic experts mentioned It is pointed out that the Chinese version of the materials genome project must be developed around “accelerated application”. Choose energy and environmental materials, marine engineering materials, military Materials and biomedical materials are related to national security and energy security and people’s health and welfare, which are urgently needed by the country and have a certain foundation. Demonstration of materials to achieve results as soon as possible for further promotion and popularization to accumulate experience in the entire material field [3]. Among them, the new energy materials field Effective use of

materials genome technology in the field to guide, accelerate research and development and application It is of great significance.
Currently, the more famous materials genome large-scale high-throughput computing platforms and materials databases include the Materials Project jointly developed by Lawrence Berkeley National Laboratory (LBNL) and MIT, Duke's Aflow.org, and Northwest's OQMD.

1.1 Materials Project

Berkeley, located less than 100 kilometers north of Silicon Valley in California, USA, has many research groups and teachers who use advanced computer technology to help materials research and development and production, and have obtained many materials calculation methods and research results. Among these research groups, CEDER The research group is the most influential. In 2017, due to his research on applied computational materials design and energy storage materials (mainly lithium battery materials) and technology, CEDER Elected as a member of the American Academy of Engineering. The materialsproject.org platform of his and PERSSON research groups has great influence around the world, and its positioning is to become Google in the field of materials research. The Materials Project contains a database that stores a large amount of information (nearly 60,000 crystal structures), which can store the results of high-throughput material property calculations, such as various calculation information, including band state density information, and battery materials. Charge-discharge curves, phase diagrams, etc., all materials calculation researchers can download structures, search material properties, view material phase diagrams, and even use the platform of this website to search for unknown materials on this developed platform. In addition, this website also opens a database interface. Using this database, you can search and filter materials by writing code. Under this working mode, the number of repeated calculations of materials is reduced. When doing material calculations, the number of human operations is reduced, the computer intelligent judgment and design are increased, and the material calculation speed is greatly improved.
In the early days, CEDER and others mainly used the Monte Carlo method [4-5] to calculate metal alloy oxide materials with fixed crystal structure sites, and published many articles using calculations to study the arrangement of oxygen atoms in metal alloy oxides [6-9]. Then, CEDER et al. [10-14] did a lot of research on the structure and phase of lithium battery materials. In these research works, they used a cluster expansion method developed from the Monte Carlo method of metallurgical calculations. : This method is to use some known density functional calculation results to fit the total energy of the system, and then use the fitting results to predict the total energy of the unknown system. The method itself relies on density functional theory calculations, but also uses some statistical concepts to accelerate ordinary density functional calculations. It is a tool similar to the current material genome calculation method. Its core is to sacrifice part of the accuracy in exchange for faster calculation speed, and then screen materials in faster calculations. from 2003 Since the year, CEDER [15-17] have successively published some papers with titles such as "data mining" and "high-throughput computing". It can be said that this is the initial germination of the idea of using computer technology to accelerate the development of traditional materials industry. For now, this platform is relatively leading in the development of materials genome technology. They have a mature materials database, some relatively mature materials calculation methods and accumulation of experience in materials screening mechanisms, and many mature researchers studying materials genome. There are also It is influenced by the world's most advanced computer ideas next to Silicon Valley in the Bay Area, and there have been many successful cases related to the use of materials genome technology in the field of new energy materials [18-25]. However, there are still many bottlenecks that need to be broken before it can truly guide experiments systematically and accelerate experiments.

1.2 Wolverton(OQMD)

CHRIS, professor of materials and mechanics at Northwestern University WOLVERTON The research group is an interdisciplinary research group whose members have very diverse backgrounds, including materials, physics, Chemistry, Mechanical Engineering and Mathematics. Their research content is also very broad, including battery materials, hydrogen storage materials, solar materials, thermoelectric materials and machine learning data mining. The WOLVERTON research group has abundant computing resources, including a 1008-core Linux cluster. In addition, they also cooperate with Northwest High-Performance Computing Systems, which has a large amount of machine time available for use. WOLVERTON was inspired by CEDER's Materials Inspired by the Project, we established the “Open Quantum Materials Database” (the open quantum materials database, OQMD) and is freely available at http://oqmd.org [26]. This is a database of thermodynamics and structure of materials based on density functional theory (DFT) calculations. This database has a friendly web interface suitable for small-scale access, and also provides an API interface to download the entire database. The database now contains 471,857 entries and includes some completely new structures. In addition to providing the crystal structure, energy, space group, formation energy, data source, energy band and other properties of the material, OQMD also clearly displays the phase diagram of the material, which is a major feature of this database.
In addition, the research group obtained data from thousands of DFT A machine learning model is constructed in the computational database (OQMD). The model can be used to predict the thermodynamic stability of any compound without any other input, and the calculation time is 6 orders of magnitude less than that of DFT. WOLVERTON et al. [27] used this model to scan approximately 160 Ten thousand candidate combinations of new ternary compounds, 4500

were successfully predicted A new type of stable material. Also using machine learning methods, WOLVERTON The research group studied the force-driven mechanism of dopant stability in zirconia. They created a cluster ranking modeling (CRM) automated method to discover powerful chemical descriptors in large property databases, and applied CRM to Stability studies of zirconia dopants. CRM is a general method that can operate on both experimental and computational data to identify the electronic structure characteristics of doped oxides and provide a good prediction of the stability of the doped oxides when they are dissolved in zirconia. [26]. They also used a high-throughput computational method called elemental substitution to use DFT 378 XYZ types were calculated (X=Cr, Mn, Fe, Co, Ni, Ru, Rh, Y=Ti, V, Cr, Mn, Fe, Ni, Z=Al, Ga, In, Si, Ge, Sn , P, As, Sb) electronic structure, magnetic properties and structural stability of half-Heusler alloys. Through calculations, the phase diagram was obtained, and new thermodynamically stable phases and dozens of semiconductors, semimetals and near-semimetals with negative formation energies were predicted [26]. It should be said that the research team currently has certain leading advantages in the fields of database construction, material genome technology method development and application.

Domestic scientific research units and universities that have carried out relevant research in this area include Shanghai Institute of Materials Genetic Engineering, Ningbo Institute of Materials Genetics, Chinese Academy of Sciences, School of New Materials, Peking University Shenzhen Graduate School, Beijing Computational Science Research Center, University of Electronic Science and Technology of China, etc. However, there is a big gap with advanced countries such as the United States, such as the lack of independently developed high-throughput material calculation programs, the lack of high-throughput material detection equipment and a relatively practical and complete material gene database [28]. The team of Academician Chen Liquan of the Institute of Physics, Chinese Academy of Sciences independently researched, developed and compiled an automated high-throughput calculation method and software platform in China, and obtained the copyright. Through supercomputers, it can select lithium-containing data from 300,000 pieces of data in the inorganic crystallography database. Materials, through fast "bond valence sum" method and high-precision first-principles molecular dynamics method, calculate the electronic structure, three-dimensional ion conduction channel, and ion migration activation energy of the material, thereby establishing a database of electrolytes and electrode materials. Through this database, a data mining method is further established, which is beneficial to the screening of new solid electrolyte materials. This high-throughput calculation method has predicted a sulfide electrolyte and obtained experimental verification [29]. Under the leadership of Academician Zhang Tongyi, Shanghai Institute of Materials and Genetic Engineering has carried out a lot of basic work in base construction such as database construction, integrated computing and software development, structural and physical property characterization, service and failure. Researcher Xiang Xiaodong of the Ningbo Institute of Materials and Genetic Engineering (now a professor at Southern University of Science and Technology) has developed high-throughput combined materials experiments and in-situ real -time high-throughput combined materials experimental technologies, including universality based on the development of synchrotron radiation large science facilities In-situ real-time high-throughput material composition/structure characterization technology, which is based on adjustable pulsed infrared laser and synchrotron radiation micro-beam white light X-ray, can explore the material structure-composition-process covering process parameters such as time, temperature, and environmental atmosphere. Correlation. Based on other micro-area characterization probes or spectroscopy testing tools, a series of functionally rich in-situ real-time high-throughput characterization technologies can also be developed, thereby giving full play to the role of high-throughput combined material preparation and characterization technology as a "new materials search engine" potential. 2014 The new generation of combined material chip technology developed by his team in 2016 changed the synthesis and screening of materials from "one pot and one stir-fry" to "thousands of pots and stir-fried at the same time", and the efficiency was rapidly increased by 1,000 to 100,000 times, thereby minimizing the time to develop new materials. Compressed to one week [30]. Professor Ma Yanming's team from Jilin University [31-32] developed relevant theories and simulation methods for predicting material structure based on chemical components, and combined first-principles calculations and high-pressure experimental measurements to explore the novel physics of matter under limited conditions such as high pressure. and chemical properties, design and synthesize new unconventional high-pressure phase multifunctional materials, reveal the intrinsic relationship between structure and macroscopic properties, and lay a knowledge base for the development of new physical theories. Based on the classification retrieval idea of crystal symmetry, combined with the particle swarm multi-objective optimization algorithm and the introduction of the bonding characteristic matrix, the research team proposed and developed CALYPSO (crystal structural analysis by particle swarm optimization) structure prediction method, based on which the CALYPSO structure prediction program with independent intellectual property rights was developed. The input quantities of the CALYPSO software package are chemical components and external conditions (such as pressure). The structure of the substance can be reasonably determined through the calculation of structural evolution and total energy, and the structural design of functional materials (such as superhard materials, etc.) can be carried out as needed. . CALYPSO The software can not only carry out structural research on three-dimensional crystals, but also carry out structural research on two-dimensional layered materials, two-dimensional surface reconstruction and zero-dimensional clusters. In the future, it can also carry out other hot research on rich structural phenomena (such as interfaces, transition states , chemical reactions, etc.).

In addition, materials genome research is also booming in other countries and regions around the world. In response to the problem of finding high symmetry points and K-point paths in structures encountered in high-throughput calculations, scientists from Lausanne, Switzerland, and a research group from the University of Tokyo in Japan [33] published a paper on finding structural symmetry on materialscloud.org. Software that facilitates high-throughput computational research on materials genomes. In this software package, the research group conducted a detailed

analysis of the symmetry, symmetry operations and other properties of the structure, and it exists in the form of code. Additionally, at materialscloud.org There is also a lot of information about pseudopotentials and download channels on the website. Although this website does not have direct materials genome work, these basic works also provide a good foundation and platform for materials genome work, accelerating materials genome research. In Singapore, large-scale high-throughput experimental preparation work has emerged. For example, for the anode material of lithium battery materials, the research group of Singapore University of Technology and Design [34] used MoSx (2<x<3) combined with multi-walled carbon nanotubes, more than 1000 mA·h/g anode material. The research group of Nanyang Polytechnic [35] prepared Pt-MoS2 High-throughput preparation methods were also used in the materials, and were reported in Nature Published in Communication. Nanyang Technological Technology collaborated with a research group from Stanford University [36] to conduct high-throughput preparation of Sn-Ge alloys. In terms of computing, the research group of Nanyang Polytechnic[37] Using existing data modeling, materials suitable for carbon monoxide catalysis were screened. In the study of impurity diffusion and activation energy of some elements in face-centered cubic materials, Singapore's High Performance Computing Institute used high-throughput computing methods to analyze the properties of this type of material [38]. In Japan, research on materials genome has been greatly developed in various fields of materials. The National Institute of Advanced Industrial Science and Technology in Japan used materials calculation methods to study sugar alcohols that do not exist in nature but can store a lot of heat. materials [39]. In addition, Japan's Asia-Pacific Research and Technology Center collaborated with Germany's Fraunhofer Institute for Mechanics of Materials to screen a type of permanent magnet material using high-throughput calculation methods [5].

It can be seen that there is research on materials genomics all over the world, covering not only traditional inorganic crystal materials, but also organic materials. The methods used are not only widely used in high-throughput calculations and high-throughput screening, but also have their own characteristics in high-throughput experimental preparation.

2 APPLICATION OF MATERIALS GENOME TECHNOLOGY IN THE FIELD OF NEW ENERGY MATERIALS

Judging from the current situation, no typical successful case has been found that systematically applies materials genome technology to guide material research and development in the field of new energy materials. However, it is not uncommon to see cases where a single or combination of methods using materials genome technology is used to guide and accelerate the development of new energy materials. CEDER The research group [18] has one of the most representative articles that searches for materials that can photocatalyze water. In all material databases, he first searched for materials with a stable structure based on the phase diagram to ensure that this material can be stably prepared. Then, based on the properties of the photocatalytic water material, he searched for materials with a band gap between 1.3 and 3.6. materials between eV, and then compare the energy band positions of the materials to ensure that the conduction band and valence band of the materials are at the appropriate positions, and then verify through experiments whether these searched materials can be used as materials for photocatalytic water splitting, and finally Sixteen materials with these photocatalytic possibilities were found. Likewise, WOLVERTON When studying how to suppress the degradation of the cathode of lithium-ion batteries, a key factor affecting battery life, a high-throughput screening method was used to first determine the required physical properties of the coating, and then select 130,000 properties from the OQMD database. Contains O The structures were screened, and 30 most likely coating candidate materials were finally screened out [40].

In the field of new energy materials research, the use of materials genome high-throughput computing technology to guide materials research and development is still in the preliminary development stage. Currently, the known theoretical application guidance technologies mainly include the following categories.

(1) Material conformation characterization (representation) It was first used in the fields of chemistry and biology to study the chain expression of polymers and the characterization of bond lengths and angles between atoms. As early as 1985, SMILES notation [41] established a chemical "language" through molecular graph theory methods and successfully encoded chemical molecules. However, due to the periodic repeatability of the crystal structure and the symmetry operations of atoms in the crystal structure, the requirements for the uniqueness of the representation of the crystal structure have increased. Until recently, more successful representation techniques for the crystal structure have been reported. SCHÜTT [42] used the pair radial distribution function as the crystal structure representation and used this representation as the input of the machine learning model to successfully predict the Fermi level of the large system (about 100 atoms). Density of energy states. FABER [43] used the traditional coulomb matrix as the representation of the crystal structure under the condition of considering periodicity, and successfully predicted the formation energy of the material.

(2) High-throughput calculation and screening and screening) In recent decades, with the development of density functional theory (density functional theory functional With the increasing maturity of quantum mechanics theory, scientists can accurately and effectively calculate the electronic interactions between atoms in materials from the perspective of quantum mechanics, thereby predicting a series of properties of materials and guiding material design theoretically. With powerful high-performance parallel computing capabilities and efficient algorithms, high-throughput computing calculation) and high-throughput screening (high-throughput screening) driven materials big data technology emerged at the historic moment. Obtaining a large number of material properties through high-throughput calculations and screening out qualified materials through high-throughput calculations have become an important means for

materials scientists to discover new materials. CEDER [44] successfully predicted the relationship between safety and charge-discharge voltage by high-throughput calculation of the relationship between charge-discharge voltage and theoretical capacity of thousands of compounds, and then screened conditions such as electron migration ability, stability and safety. relationship, and predicted ideal battery cathode materials. CASTELLI [45] used conditions such as stability, energy band position, and bandgap width to evaluate the cubic perovskite structure containing a total of 52 High-throughput screening of 5,400 semiconductor oxides and nitrogen oxides of various metal elements successfully identified 10 existing ideal oxides and 5 ideal nitrogen oxides, and predicted 9 unknown combination materials as high-performance materials. Efficient photolysis water material. MADSEN [46] used a high-throughput method to automatically search for new thermoelectric materials among 570 antimony element-containing complexes in the Inorganic Crystal Structure Database (ICSD), and found that LiZnSb with Zintl phase is an ideal n-type thermoelectric materials.

(3) Machine learning In recent years, artificial intelligence has been increasingly used in various fields in order to discover phenomena and laws that cannot be discovered using traditional means. CEDER [23] used the Bayesian probability statistics method, using the crystal structure data in the inorganic crystal structure database as the training set, and the crystal configuration and element type as the input of the model, consuming a small amount of computing resources. 209 new ternary oxides were discovered. CORMA et al. [47] used the support vector machine method and used the synthesis variables in the zeolite synthesis process as parameters of the model (concentration of the initial gel, reaction process, temperature and time, etc.), Accurately predict the structural properties and thermodynamic properties of synthesized products.

(4) Neural network Neural network technology is an algorithmic mathematical model that imitates the behavioral characteristics of neurons in animal brains and performs distributed parallel information processing. It achieves the purpose of processing complex connections by adjusting the (non-)linear relationship between a large number of internal nodes. In the field of materials science, there are also complex relationships between various variables that need to be explored, such as the structure-activity relationship of crystals, the relationship between reaction variables and products during material synthesis, etc. MAYYAS et al. [48] used forward propagation technology and used the mass percentage of copper and the volume ratio of silicon carbide as the input of the model.

Density, porosity, and hardness were used as model outputs to successfully predict the physical properties of the compound based on the structural properties of aluminum-copper-based silicon carbide. MORADI [49] through artificial neural network (artificial neural network) and multiple linear regression models (multiple linear Regression model), using reflection parameters such as electrolyte concentration, temperature and applied voltage as the input of the model, successfully predicted the mesopore spacing of nano-mesoporous anodic aluminum oxide. SCHERAGA [50] used artificial neural network technology to generate a force field for liquid water based on the multi-body polarization effect. After Monte Carlo simulation (Monte Carlo simulation), the force field of liquid water was generated. Carlo simulation), which is in good agreement with the experimental values.

(5) Optimization algorithm Optimization algorithms have also become an important technical method for materials scientists to search for new compounds in chemical space for material discovery and design. Common optimization algorithms such as genetic algorithm, evolutionary algorithm and particle swarm algorithm have been applied to find the stable crystal structure with the lowest energy. WANG [51] used a differential evolutionary algorithm to find low-energy stable cluster structures, and successfully discovered dozens of cobalt elemental clusters and lead elemental clusters while consuming a small amount of computing resources. HO [52] used an adaptive genetic algorithm to discover multiple phases of magnesium silicate during the decomposition process, and verified this discovery through experiments. MA et al. [31] developed the structure search program CALYPSO based on the particle swarm algorithm to find the lowest energy structure, and successfully predicted the stable phases of silica and calcium carbonate under high pressure conditions.

Although various methods or combinations of methods currently have successful application cases in the field of new energy materials research, they are still far from reaching the original intention of the Materials Genome Project. The main manifestations are: First, the system is not strong. Of course, this is different from the Materials Genome technology. It is still in the preliminary stage of development and cannot yet achieve a high degree of aggregation and system application of tools, means, methods, data, platforms, etc. Second, various technological developments still need to be strengthened. For example: ① For the research and development of materials with non-single performance requirements, we cannot yet fully evaluate the material properties of systems with complex performance requirements through calculations; for lithium batteries, their performance needs to be excellent in all aspects (capacity, charge and discharge speed, stability, etc.), we still cannot use material genome high-throughput calculations to intuitively predict what kind of materials can make batteries with good performance; ② For the development of new materials, our current material search is still based on the existing crystal database. Looking for materials with certain properties among the materials that can be prepared, or just doing some work similar to chemical doping, it is difficult to achieve revolutionary results; although there are already some computational software that use genetic algorithms to search for material structures , but these software consume a lot of computing resources and cannot stably guarantee that new materials will be found. There is currently no computing technology that uses these software for high-throughput screening; ③ For computational simulation of experiments, we cannot yet predict the experimental process through materials genome technology. At present, we do not have a mature method for computational simulation experiments. It is more important to use To interpret and verify experimental results through computational simulation, it is necessary

to integrate high-throughput calculations and high-throughput experiments to give full play to the systematic advantages of materials genome technology.

3 CASE INTRODUCTION OF THE APPLICATION PROGRESS OF MATERIALS GENOME TECHNOLOGY IN THE FIELD OF NEW ENERGY MATERIALS

Since the "Xiangshan Science Conference" in December 2011, my country has begun to attach importance to the development of materials science systems engineering, and in 2016, for the first time during the 13th Five-Year Plan period, it established a national key special plan for "Materials Genetic Engineering Key Technologies and Support Platforms". Among them, the "Research and Development of All-Solid-State Lithium Batteries and Key Materials Based on Materials Genome Technology" project was officially approved in August 2016. The project is led by Professor Pan Feng from Peking University Shenzhen Graduate School, and 11 domestic units with good research foundations in the field of materials computing and lithium batteries participate in the research and development [28]. As the project lead unit, the School of New Materials of Peking University Shenzhen Graduate School has achieved certain accumulation in the field of new energy materials genomic technology and applications, and has formed a systematic R&D roadmap. , the "Peking University New Materials" public service materials big data platform website has been built to build a new energy materials genome public information and data service platform and a new energy material testing and evaluation public service platform based on the concepts of cooperation, openness, and sharing. Using these technologies and resources, the team has initially achieved a series of results in the field of new energy materials and materials genome technology applications. In terms of high-throughput computing platform construction, the team has independently built more than 100Tflops Computing power of GPU+CPU New high-throughput computing cluster. Independently developed high-throughput material computing software packages and new algorithms, including improvements and improvements to independently developed GPU-based The first-principles materials calculation program PWmat has higher calculation efficiency than the widely used CPU-based first-principles materials calculation program VASP; a set of structure search programs SGO was developed; a set of more advanced Program WKM for accurate calculation of electronic structures. The overall high-throughput computing platform can achieve ≥102 levels of concurrent high-throughput calculations. Currently, ≥4,000 materials have been calculated through high-throughput around energy storage materials and photovoltaic materials.

In terms of high-throughput computing, the team is currently focusing on the mechanism research of materials (Figure 1), looking for structural genes that determine material properties, and enriching the relevant mechanical properties and physical models of materials in the database. For example, through high-throughput calculations of dozens of ternary layered material structures, they systematically and completely revealed the lithium ion deintercalation and transport mechanism of ternary NMC layered cathode materials for the first time, and discovered that lithium in the ternary material (NMC) system The relationship between ion diffusion dynamics and the initial component composition of ternary materials provides important clues and theoretical guidance for the future optimization and design of lithium ion kinetic properties of ternary materials and the preparation of high-performance lithium battery materials [53]. This article was selected by Science Network as one of the top ten most watched papers in domestic natural sciences in 2015. Recently, through large-scale first principles calculations and experimental verification, they found that the stability of the ternary layered cathode material is related to the most unstable oxygen in the lattice structure, and the stability of oxygen is determined by its basic coordination unit. , by establishing a model, it systematically revealed the regulation of oxygen stability by factors such as the lithium content, the content and valence state of transition metal elements, Ni/Li anti-site defects in layered materials, and provided a basis for the stability of ternary material lithium-ion batteries in the future. It provides important clues and theoretical guidance for the optimization of sex [54]. In addition, integrating theoretical calculations and experiments, the team For the first time in the field of thin film solar cells and materials, a physical model of nanofunctional interface passivation solar cells was proposed, which solved the problem of the high barrier nature of CdTe material making it difficult to form a continuous conductive channel. ALD-Al2O3 back contact CdTe cells [55-56] and P-type semiconductor back electrode CdTe cells [57] were theoretically designed and experimentally prepared successfully, achieving an increase in solar cell efficiency of more than 10%. In addition to its guiding significance for solar cells, this physical model also has direct reference value for other optoelectronic materials and devices, and a CdS photoresist device was successfully prepared [58]. The team also collaborated with City University of Hong Kong and the Institute of Chemistry, Chinese Academy of Sciences, to make progress in research on the interface between TiO2 and dyes in dye-sensitized solar cells (DSSCs) through first principles and experiments [59].

In terms of high-throughput detection, the team developed high-throughput detection methods from electrode active materials, battery pole pieces to full batteries. They started from the characteristics of the cyclic voltammogram curve of single-particle nanolithium iron phosphate and established an electrochemical model of nano-lithium iron phosphate single particle to quickly analyze and characterize the diffusion and interface charge transfer process of lithium ions in nano-lithium iron phosphate nanoparticles. , to find out the main factors limiting the electrochemical performance of single-particle electrode materials [60]. By innovatively applying 3D printing technology to the research of battery electrodes, 3D printing can be used to control the thickness and structure of the electrodes, reduce the impact of solution diffusion, and quickly test the performance of battery electrodes with different characteristics. Analysis[61]. Furthermore, by simulating the operating conditions of electric vehicles when the current is used as the excitation source and the voltage as the response signal, the response of the battery under such a large signal is analyzed. The

results show that this method can well characterize the SOC and SOH of the battery. Using Multi-channel testing equipment can achieve high-throughput and rapid detection of full battery performance [62]. In addition, they cooperate with relevant international and domestic large scientific facilities to jointly develop high-throughput detection methods. For example, using high-throughput detection methods such as fast laser spectroscopy, synchrotron radiation hard X-ray structure determination, soft X-ray absorption spectroscopy, and the application of large scientific equipment, they discovered for the first time that Cu2+ spontaneously changes into Zn0.5Cd0.5S after doping. Cu+ , and promotes S Generation of defects. At the same time, Cu+ and S defects will accept photogenerated holes and photoelectrons respectively, thus improving the separation ability of photogenerated carriers. Combined with the property of Cu doping to improve visible light absorption, the photocatalytic hydrogen production performance of Cu-doped samples for water splitting is the highest at 2.8 times that of the undoped samples. times, reaching 21.4 mmol/ (h·g) , at 428 nm led The quantum efficiency under light reaches 18.8%. The research results successfully explained Cu for the first time Doping improves Zn-Cd-S The mechanism of the system's photocatalytic performance is the world's first-class photocatalytic water splitting and hydrogen production performance [63].

In terms of database construction, a material gene database containing structural data and partial performance data of more than 300,000 crystal materials has been initially built. The electronic structure calculation of the database materials uses the most accurate HSE algorithm currently, which is unique in the world. .

Judging from the team's case study, the systematic use of materials genome technology and paradigms to conduct research on new energy materials and devices has indeed brought higher efficiency to the team's scientific research breakthroughs. Professor Pan Feng, the leader of the team, has therefore become an Elsevier China Highly Cited Scholar in 2015 and 2016. In future work, the team will devote more energy to high-throughput computing and screening, the construction of large materials databases, and the development of advanced intelligent technologies such as machine learning and neural network technology for data mining, so as to deepen materials genome technology. Integrate into the entire process of new energy material research and development and accelerate the research and development process of new materials.

4 EXHIBITION SEE

Currently, there are many successful application cases using single methods or method combinations to accelerate research in the field of new energy materials using materials genome technology. However, there is still much room for improvement before the systematic application of materials genome technology. The ultimate goal of materials genome is to complete the "on-demand design" of materials through theoretical simulation and calculation and realize full digital intelligent manufacturing of materials. It is a new model of materials development and is essentially a new and efficient research method, culture and concept. . We believe that there are mainly the following development paths.

4.1 Improve the Technical Level of all Aspects of Material Genome

Efforts should be increased to independently develop various aspects of the materials genome, such as high-throughput computing, high-throughput experimental technology and the construction of a database platform for new energy materials big data, with a focus on improving the technical level of each aspect. For example, establishing electroacoustic interaction models and ion transport models; developing new efficient computing packages and large-scale parallel computing programs based on first principles; establishing cross-scale multi-physics models from atoms to devices for simulation. Design and develop an intelligent high-throughput material preparation center based on combinatorial chemical methods, a high-throughput preparation center based on microfluidic reactors, an in-situ preparation and testing system based on multi-channel electrochemical deposition, and a high-throughput material preparation center based on source and substrate control. High-flux material preparation technology such as flux magnetron sputtering preparation equipment. Based on 3D printing, we develop high-throughput device preparation technology for lithium batteries, thin-film photovoltaic devices, water electrolysis and fuel cell electrodes. Optimize and develop new basic detection methods and high-throughput detection methods for energy storage devices from single particles of materials, pole pieces, batteries to systems. Establish a structured database system for new energy materials, use data regression and machine learning algorithms to conduct data mining analysis, study their structure-activity relationships, accelerate the prediction and design of new original devices and key materials, and select materials and device structures with excellent performance. At the same time, it is necessary to accelerate the construction of large scientific devices with high precision, high throughput and extreme conditions (ultra-high pressure, supercritical, rapid cooling) with different temporal and spatial resolutions, break through the limits of traditional analysis, reveal the essence of materials, and provide information for the discovery of new materials. Support platform for high-precision and high-throughput detection.

4.2 Systematically Integrate Technologies from all Aspects of Material Genome

We should effectively integrate and achieve a high degree of aggregation of various resources in the process of developing various tools, means, methods, and databases for new energy materials genome technology, and integrate various elements into a big data artificial intelligence system through the systematic application of materials genome technology. , build a platform, establish standards, establish a complete data sharing mechanism, generate new energy material big data through open and shared Internet technology, develop artificial intelligence algorithms, software and

hardware systems suitable for learning and analyzing material big data, and improve material discovery and application speed.

4.3 The Government Leads the Construction of Major Public Scientific Research Service Platforms such as Large Scientific Installations

The materials genome platform should reflect the characteristics of public services. High-precision and high-throughput detection in materials genome technology requires the government to lead the construction of large scientific facilities and other major public scientific research platforms, combined with the Internet and big data Characteristics, we believe that the non-profit public service platform is the structure that can best maximize its efficiency. In order to ensure that the public service platform can better play the role of cooperation, openness and sharing, the ideal should be a non-profit public service platform for the New Energy Materials Genome Research Institute led by the government and co-constructed by one or several scientific research institutions. This public service platform can as a new collaborative innovation scientific research entity that integrates and manages various distributed resources in the future.

4.4 Market Application-Oriented Service Innovation Across the Entire Industry Chain

The Materials Genome Project is a market-and application-oriented material The new concept of research and development emphasizes the integration and collaboration of all aspects of materials "discovery-development-production-application". Therefore, we should use material-based Integrate resources through group technologies and cultivate new types of open and collaborative cooperation model, innovate service paradigms and organizational structures, and organize government, industry, academia, A collaborative innovation alliance integrating research and application across the entire industry chain, with a non-profit With the Public Service Platform Research Institute as the core, government guidance, and enterprises giving full play to and social capital through multiple channels to build a large infrastructure platform and deploy it effectively resources, carry out international cooperation and exchanges, serve and lead key new capabilities Source material research and development, solving common industrial problems, and providing systems for enterprises solutions to promote the leap-forward development of the entire industry chain of new energy materials.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] National Science and Technology Council. Materials genome initiative strategic plan. USA, 2014.
[2] RACCUGLIA P, ELBERT K C, ADLER P D F. Machine-learning-assisted materials discovery using failed experiments. Nature, 2016, 533 (7601): 73-76.
[3] Wang Hong, Xiang Yong, Xiang Xiaodong. Materials genome -a new model for materials research and development. Science and Technology Herald, 2015, 33(10): 13-19.
[4] DREYSSE H, CEDER G, FONTAINE D D. Determination of effective pair interactions and segregation behavior at alloy surfaces. Vacuum,1990, 41(1): 446-448.
[5] CEDER G, HUANG P, MENON S. Ab initio calculation of the Cu-Pd one-dimensional long period superstructure phase diagram. Acta Metallurgica Et Materialia,1990, 38(11): 2299-2308.
[6] FONTAINE DD, CEDER G, ASTA M. Thermodynamics of oxygen ordering in YBa2Cu3Oz. Journal of the Less-Common Metals, 1990, 164(11): 108-123.
[7] CEDER G, ASTA M, CARTER W C. Phase diagram and low-temperature behavior of oxygen ordering in YBa2Cu3Oz using ab initio interactions. Physical Review B Condensed Matter, 1990, 41(13). https://doi.org/10.1103/PhysRevB.41.8698.
[8] CEDER G. A computational study of oxygen ordering in YBa2Cu3Oz and its relation to superconductivity. Molecular Simulation, 1994,12(2): 141-153.
[9] FONTAINE D D, CEDER G, ASTA M. Low-temperature long-range oxygen order in YBa2Cu3Oz. Nature, 1990, 343(6258): 544-546.
[10] VEN A V D, MARIANETTI C, MORGAN D. Phase transformations and volume changes in spinel LixMn2O4. Solid State Ionics, 2000, 135(1): 21-32.
[11] VEN A V D, CEDER G. Lithium diffusion mechanisms in layered intercalation compounds. Journal of Power Sources, 2001(97/98): 529-531.
[12] DOMPABLO M E, MARIANETTI C, VEN A V D. Jahn-Teller mediated ordering in layered LixMO2 compounds. Physical Review B Condensed Matter, 2001, 63(63). https://doi.org/10.1103/PhysRevB.63.144107.
[13] VAN D V A, CEDER G, ASTAM. First-principles theory of ionic diffusion with nondilute carriers. Physical Review B, 2001, 64(18):607-611.
[14] DOMPABLO M E, VEN A V D, CEDER G. First-principles calculations of lithium ordering and phase stability on LixNiO2.Physical Review B, 2002, 66(6): 340-351.

[15] CURTAROLO S, MORGAN D, PERSSON K. Predicting crystal structures with data mining of quantum calculations. Physical Review Letters, 2003, 91(13): doi: 10.1103/PhysRevLett.91.135503.
[16] MORGAN D. Data mining approach to ab-initio prediction of crystal structure. MRS Proceedings, 2003, 804: 1-6.
[17] MORGAN D, CEDER G, CURTAROLO S. High-throughput and data mining with abinitio methods. Measurement Science & Technology, 2004, 16(16): 296-301.
[18] WU Y, LAZIC P, HAUTIER G. First principles high throughput screening of oxynitrides for water-splitting photocatalysts. Energy & Environmental Science, 2012, 6(1): 157-168.
[19] MUELLER T, HAUTIER G, JAIN A. Evaluation of tavorite-structured cathode materials for lithium-ion batteries using high-throughput computing. Chemistry of Materials, 2011, 23(17): 3854-3862.
[20] CEDER G, MORGAN D, FISCHER C. Data-mining-driven quantum mechanics for the prediction of structure. MRS Bulletin, 2006, 31(12): 981-985.
[21] JAIN A, HAUTIER G, MOORE C J. A high-throughput infrastructure for density functional theory calculations. Computational Materials Science, 2011, 50(8): 2295-2310.
[22] HAUTIER G, JAIN A, ONG S P. Phosphates as lithium-ion battery cathodes: An evaluation based on high-throughput ab initio calculations. Chemistry of Materials, 2011, 23(15): 3495-3508.
[23] HAUTIER G, FISCHER C C, JAIN A. Finding nature's missing ternary oxide compounds using machine learning and density functional theory. Chemistry of Materials, 2010, 22(12): 3762-3767.
[24] HAUTIER G, FISCHER C, EHRLACHER V. Data mined ionic substitutions for the discovery of new compounds. Inorganic Chemistry, 2011, 50(2): doi: 10.1021/ic102031h.
[25] YANG L, CEDER G. Data-mined similarity function between material compositions. Physical Review B, 2013, 88(22): 330-339.
[26] MEREDIG B, WOLVERTON C. Dissolving the periodic table in cubic zirconia: Data mining to discover chemical trends. Chemistry of Materials, 2014, 26(6): 1985-1991.
[27] MEREDIG B, AGRAWAL A, KIRKLIN S. Combinatorial screening for new materials in unconstrained composition space with machine learning. Physical Review B, 2014, 89(9): 82-84.
[28] Lin Hai, Linyuan, Pan Feng. "Research on all-solid-state lithium-ion batteries based on materials genome" was approved by the national project. Energy Storage Science and Technology, 2016, 5(6): 922-925.
[29] WANG X, XIAO R, LI H. Oxysulfide LiAlSO: A lithium superionic conductor from first principles. Physical Review Letters, 2017, 118(19): doi: https://doi.org/10.1103/PhysRevLett.118.195901.
[30] XIANG X, WANG H, XIANG Y. Applications of combinatorial material chip technology in research and development of new materials. Science & Technology Review, 2015, 33(10): 64-78.
[31] WANG Y, LU J, ZHU L. Crystal structure prediction via particle swarm optimization. Physics, 2010, 82(9): 7174-7182.
[32] WANG Y, LU J, ZHU L. CALYPSO: A method for crystal structure prediction. Computer Physics Communications, 2012, 183(10): 2063-2070.
[33] HINUMA Y, PIZZI G, KUMAGAI Y. Band structure diagram paths based on crystallography. Computational Materials Science, 2017, 128: 140-184.
[34] SHI Y, WANG Y, WONG J I. Self-assembly of hierarchical MoSx/CNT nanocomposites (2<x<3): Towards high performance anode materials for lithium ion batteries. Scientific Reports, 2013, 3: doi: 10.1038/srep02169.
[35] HUANG X, ZENG Z, BAO S. Solution-phase epitaxial growth of noble metal nanostructures on dispersible single-layer molybdenum disulfide nanosheets. Nature Communications, 2013, 4: 1444-1452.
[36] FAN S, LIM L Y, TAY Y Y. Rapid fabrication of a novel Sn-Ge alloy: Structure-property relationship and its enhanced lithium storage properties. Journal of Materials Chemistry A, 2013, 1(46): 14577-14585.
[37] YAN W, CHEN Y, YANG Y. Development of high performance catalysts for CO oxidation using data-based modeling. Catalysis Today, 2011, 174 (1): 127-134.
[38] ZENG Y, BAI K. High-throughput prediction of activation energy for impurity diffusion in fcc metals of Group I and VIII. Journal of Alloys and Compounds, 2015, 624 : 201-209.
[39] INAGAKI T, ISHIDA T. Computational design of non-natural sugar alcohols to increase thermal storage density: Beyond existing organic phase change materials. Journal of the American Chemical Society, 2016, 138(36): 11810 -11819.
[40] AYKOL M, KIM S, HEGDE V I. High-throughput computational design of cathode coatings for Li-ion batteries. Nature Communications, 2016, 7: 13779-13790.
[41] WEININGER D SMILES. A chemical language and information system. 1. Introduction to methodology and encoding rules. Journal of Chemical Information & Computer Sciences, 1988, 28 (1): 31-36.
[42] SCHUTT K T, GLAWE H, BROCKHERDE F. How to represent crystal structures for machine learning: towards fast prediction of electronic properties. Physical Review B, 2013, 89(20): 163-168.
[43] FABER F, LINDMAA A, VON LILIENFELD O A. Crystal structure representations for machine learning models of formation energies. International Journal of Quantum Chemistry, 2015, 115(16): 1094-1101.
[44] CEDER G. Opportunities and challenges for first-principles materials design and applications to Li battery materials. MRS Bulletin, 2010, 35(9): 693-701.
[45] CASTELLI I E, OLSEN T, DATTA S. Computational screening of perovskite metal oxides for optimal solar light capture. Energy & Environmental Science, 2012, 5(2): 5814-5819.

[46] MADSEN G K H. Automated search for new thermoelectric materials: The case of LiZnSb. Journal of the American Chemical Society, 2006, 128(37): 12140-12146.
[47] SERRA J M, BAUMES L A, MOLINER M. Zeolite synthesis modelling with support vector machines: A combinatorial approach. Combinatorial Chemistry & High Throughput Screening, 2007, 10(1):13-24.
[48] HASSAN A M, ALRASHDAN A, HAYAJNEH M T. Prediction of density, porosity and hardness in aluminum-copper-based composite materials using artificial neural network. Journal of Materials Processing Technology, 2009, 209(2): 894-899.
[49] AKBARPOUR H, MOHAJERI M, MORADI M. Investigation on the synthesis conditions at the interpore distance of nanoporous anodic aluminum oxide: A comparison of experimental study, artificial neural network, and multiple linear regression. Computational Materials Science, 2013, 79: 75-81.
[50] CHO K H, NO K T, SCHERAGA H A. A polarizable force field for water using an artificial neural network. Journal of Molecular Structure, 2002, 641(1): 77-91.
[51] CHEN Z, JIANG X, LI J. PDECO: Parallel differential evolution for clusters optimization. Journal of Computational Chemistry, 2013, 34(12): 1046-1059.
[52] WU S Q, JI M, WANG C Z. An adaptive genetic algorithm for crystal structure prediction. Journal of Physics Condensed Matter An Institute of Physics Journal, 2014, 26(3): doi: 10.1088/0953-8984/26/3/035402.
[53] WEI Y, ZHENG J, CUI S. Kinetics tuning of Li-ion diffusion in layered Li(NixMnyCoz)O2. Journal of the American Chemical Society, 2015, 137(26): 8364-8367.
[54] ZHENG J, LIU T, HU Z. Tuning of thermal stability in layered Li(NixMnyCoz)O2. Journal of the American Chemical Society, 2016, 138(40): 13326.
[55] LIANG J, LIN Q, LI H. Rectification and tunneling effects enabled by Al2O3 atomic layer deposited on back contact of CdTe solar cells. Applied Physics Letters, 2015, 107(1): 699-7539.
[56] LIN Q, SU Y, ZHANG M J. A novel p-type and metallic dual-functional Cu-Al2O3 ultra-thin layer as the back electrode enabling high performance of thin film solar cells. Chemical Communications, 2016, 52(71): 10708-10711.
[57] ZHANG M J, LIN Q, YANG X. Novel p-type conductive semiconductor nanocrystalline film as the back electrode for high-performance thin film solar cells. Nano Letters, 2016, 16(2): 1218-1223.
[58] PEI Y, PEI R, LIANG X. CdS-nanowires flexible photo-detector with Ag-nanowires electrode based on non-transfer process. Scientific Reports, 2016, 6: doi: 10.1038/srep21551.
[59] ZHENG J, ZHANG K, FANG Y. How to optimize the interface between photosensitizers and TiO2 nanocrystals with molecular engineering to enhance performances of dye-sensitized solar cells?. ACS Applied Materials & Interfaces, 2015, 7(45): 66-76.
[60] HU J, LI W, DUAN Y. Single-particle performances and properties of LiFePO4 nanocrystals for Li-ion batteries. Advanced Energy Materials, 2017, 7(5): doi: 10.1002/aenm.201601894.
[61] HU J, JIANG Y, CUI S. 3D-Printed Cathodes of LiMn1-xFexPO4 nanocrystals achieve both ultrahigh rate and high capacity for advanced lithium-ion battery. Advanced Energy Materials, 2016, 6(18): 1600856.
[62] QIAO R X, ZHANG M J, LIU Y D. A novel real-time state-of-health and state-of-charge co-estimation method for LiFePO4 battery. Chinese Physics Letters, 2016, 33(7): 182-185.
[63] MEI Z, ZHANG B, ZHENG J. Tuning Cu dopant of Zn0.5Cd0.5S nanocrystals enables high-performance photocatalytic H2 evolution from water splitting under visible-light irradiation. Nano Energy, 2016, 26: 405-416.