ISSN 1674-1056
http://cpb.iphy.ac.cn
http://iopscience.iop.org/cpb
中国物理 B
Chinese Physics
中国物理
Chinese Physics B

High-throughput design of functional materials using materials genome approach

Kesong Yang(杨可松)

Citation: Chin. Phys. B . 2018, 27(12): 128103. **doi:** 10.1088/1674-1056/27/12/128103
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Combinatorial synthesis and high-throughput characterization of copper-oxide superconductors

J Wu, A T Bollinger, X He, I Božović

Chin. Phys. B . 2018, 27(11): 118102. **doi:** 10.1088/1674-1056/27/11/118102

The effects of combining alloying elements on the elastic properties of *γ*-Ni in Ni-based superalloy: High-throughput first-principles calculations

Baokun Lu(路宝坤), Chongyu Wang(王崇愚)

Chin. Phys. B . 2018, 27(7): 077104. **doi:** 10.1088/1674-1056/27/7/077104

Germanene nanomeshes:Cooperative effects of degenerate perturbation and uniaxial strain on tuning bandgap

Yan Su(苏燕), Xinyu Fan(范新宇)

Chin. Phys. B . 2017, 26(10): 108101. **doi:** 10.1088/1674-1056/26/10/108101

Geometrically induced *π*-band splitting in graphene superlattices

Yanpei Wei(魏艳佩), Tiantian Jia(贾甜甜), Gang Chen(陈刚)

Chin. Phys. B . 2017, 26(2): 028103. **doi:** 10.1088/1674-1056/26/2/028103

Enhancing visible-light photocatalytic activity of *α*-Bi_2O_3 via non-metal N and S doping

Shang Jun, Gao Yuan, Hao Wei-Chang, Jing Xi, Xin Hui-Ju, Wang Liang, Feng Hai-Feng, Wang Tian-Min

Chin. Phys. B . 2014, 23(3): 038103. **doi:** 10.1088/1674-1056/23/3/038103

TOPICAL REVIEW — Physics research in materials genome

High-throughput design of functional materials using materials genome approach

Kesong Yang(杨可松)[1,2,†]

[1] *Department of NanoEngineering, University of California San Diego, La Jolla, California* 92093-0448, *USA*
[2] *Center for Memory and Recording Research, University of California San Diego, La Jolla, California* 92093-401, *USA*

(Received 23 July 2018; revised manuscript received 18 September 2018; published online 10 November 2018)

High-throughput computational materials design provides one efficient solution to accelerate the discovery and development of functional materials. Its core concept is to build a large quantum materials repository and to search for target materials with desired properties via appropriate materials descriptors in a high-throughput fashion, which shares the same idea with the materials genome approach. This article reviews recent progress of discovering and developing new functional materials using high-throughput computational materials design approach. Emphasis is placed on the rational design of high-throughput screening procedure and the development of appropriate materials descriptors, concentrating on the electronic and magnetic properties of functional materials for various types of industrial applications in nanoelectronics.

Keywords: high-throughput, first-principles, materials genome, functional materials

PACS: 81.05.Zx, 71.15.–m, 31.15.es, 73.20.–r **DOI:** 10.1088/1674-1056/27/12/128103

1. Introduction

Advanced materials have always played critical roles in economic growth and human well-being, with various types of industrial applications ranging from energy production, conversion, and storage to electronic information technology, and continue to do so.[1,2] Nevertheless, owning to increasing demand for high-performance materials, the development of advanced materials is facing great challenges today and it is becoming crucial to accelerate discovery and development of advanced materials systems. In this background, the Materials Genome Initiative (MGI) was launched by the United States in 2011 with a goal to discover, develop, and deploy advanced materials in a double speed at a fractional cost.[2] Its central idea is to accelerate the material discovery and development by integrating computational, experimental, and data informatics tools, corresponding to three basic elements: computations, experiments, and digital data (or materials information). Among the three components, computations to predict materials properties, behaviors, and even performances play a fundamental role in the MGI, with an ideal goal to enable real-world materials discovery and development but with a reduced cost and time compared to traditional experimental design and tests.

As an emerging area of materials science, high-throughput computational materials design has attracted increasing attention, and it has been used for materials discovery and design.[3–19] Its core concept is to build a large quantum materials repository containing the calculated materials properties of existing and hypothetical materials,[20–23] and then search for target materials with the desired properties across the materials repository. This concept involves two consequent steps: i) large-scale electronic structure calculations using high-throughput computational tools, and ii) processing, collection, and storage of materials information for the entire materials repository.[20,24]

These two steps involve two basic components of the MGI: computational tools and digital data. Once target materials are identified after screening the quantum material repository, these materials will need to be further validated from experiments. Therefore, in this sense, the high-throughput computational design shares exactly same idea with the MGI; *i.e.*, to achieve an accelerated discovery and design of materials. Accordingly, it is also reasonable to call the high-throughput computational materials design approach as materials genome approach. So far, a number of efforts have been devoted to developing computational tools[24–29] and quantum materials repositories.[20–23] For example, several software frameworks have been developed to enable high-throughput calculations of inorganic crystalline materials and analysis of materials properties, including AFLOW,[20] pymatgen,[24,25] the Atomic Simulation Environment (ASE),[26] and MatCloud.[28] Several large-scale quantum materials databases produced from density functional theory (DFT) calculations, including AFLOWLIB,[20] Materials Project,[22,30] Open Quantum Materials Database (OQMD),[21] and Computational Materials Repository,[23] have also been available online.

As shown from its schematic workflow in Fig. 1, the high-throughput computational materials design is a systematic and comprehensive project, involving several consequent steps.

†Corresponding author. E-mail: kesong@ucsd.edu

© 2018 Chinese Physical Society and IOP Publishing Ltd

Besides the two steps mentioned above (*i.e.*, the generation of quantum materials repository and the management of materials information), one more critical step is to search for target materials using appropriate and effective materials descriptors (or *materials genes* in the materials genome approach). Actually, the key to a successful high-throughput computational materials design is development of accessible materials descriptors. The materials descriptors could be the any identifiable materials characters such as elements and structures or a combination of several materials characters that are closely related to the desired materials properties. The key idea of the materials descriptors is that they can be used to identify target materials in a high-throughput fashion via an efficient screening algorithm. In this review, we focus on the procedure of discovering and developing new materials using materials genome design approach. An emphasis is on the development of appropriate materials descriptors and the rational design of high-throughput screening procedure, concentrating on the electronic and magnetic properties of inorganic functional materials for various types of applications in nanoelectronics.

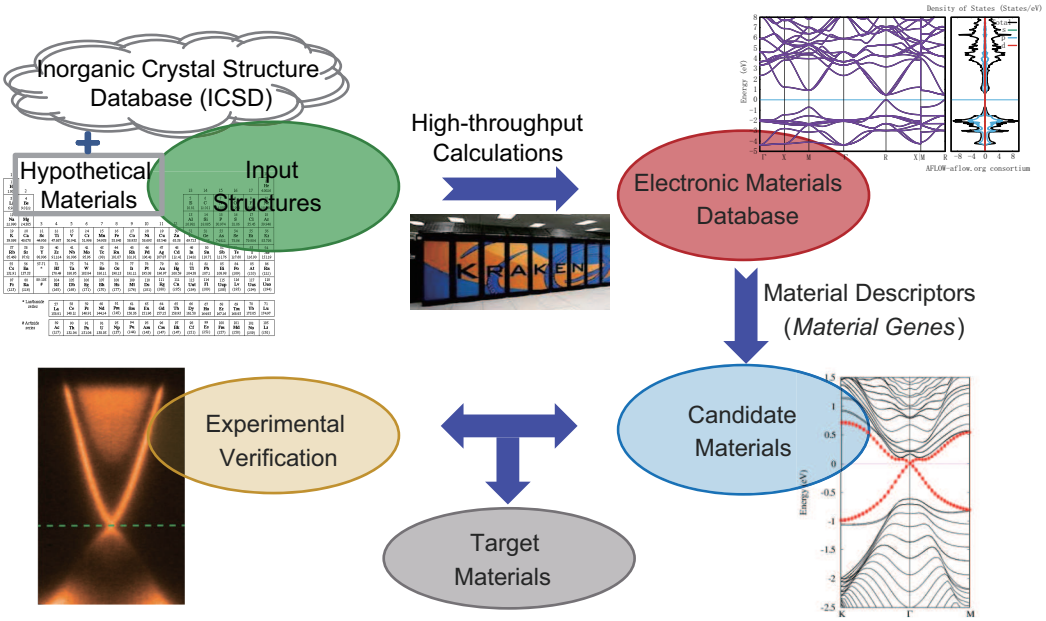


Fig. 1. (color online) Schematic workflow of high-throughput design of functional materials using materials genome approach.

2. Application examples

This section will discuss a few specific application examples of searching for functional materials using high-throughput materials design approach, with an emphasis on their electronic and magnetic properties for various types of applications in nanoelectronics.

2.1. Thermoelectric materials

Thermoelectric materials have received increasing attentions in the energy conversion field because of their ability to convert heat into electrical energy.[31] The energy conversion capability of a thermoelectric system can be estimated by a dimensionless figure of merit: $ZT = P \cdot T/\kappa$, where P, T, and κ are power factor, temperature, and thermal conductivity, respectively. The power factor P is defined as $P = \delta S^2$, in which δ is the electrical conductivity and S is the Seebeck coefficient. Accordingly, a high energy conversion capability requires a low thermal conductivity (κ) but a high electrical conductivity (δ). To achieve commercial applications of thermoelectric materials, a number of research efforts are being made to improve their efficiencies, and one way is to use nanostructured materials such as sintered powder nanocomposites.[32–35] Nevertheless, the traditional process of experimental materials design in laboratory is often a painstaking trial-and-error process and is bound by high time- and cost-consumption. In 2006, Madsen reported the first search model for new thermoelectric materials using automated band structure calculations for all the Sb-based compounds based on inorganic crystal structure database.[36] In this work, Madsen first extracted about 1630 Sb-containing compounds from the Inorganic Crystal Structure Database.[37] Next, the authors performed an initial screening and removed all the undesired compounds on the basis of criteria: i) disordered compounds; ii) compounds containing lanthanide atoms; iii) compounds that contain highly electronegative elements N, O, F, and Cl. This initial screening reduced the dataset to 570 compounds. Third, the authors carried out electronic structure calculations for all these 570 compounds and extracted their transport coefficients to select thermoelectric materials, in which air-sensitive compounds were not considered. In this work, one selected material, LiZnSb, was presented in detail.

In 2011, Wang *et al.* evaluated thermoelectric properties of sintered compounds using high-throughput first-principle calculations.[7] The fundamental reason to use sintered com-

pounds is that the wide existence of grain boundaries reduces the thermal conductivity (κ) by reducing the phonon mean free path (λ). In this study, the authors have calculated the thermoelectric properties of more than 2500 sintered compounds from the online quantum materials repository AFLOWLIB[20] and found 20 best candidate compounds with the highest power factor, see Fig. 2. To make the high-throughput calculations possible, several assumptions were made in this work: i) Diffusive electron scattering occurs at grain boundaries; ii) Grain sizes are smaller than the mean free path of the bulk compounds; iii) Constant-mean-free-path is equal to the grain size of the compound; and iv) The same grain size is used for all the compounds. This work also reveals some interesting correlations. That is, sintered thermoelectric compounds with large power factors tend to have large band gaps, heavy carrier effective masses, and many atoms per primitive cell.

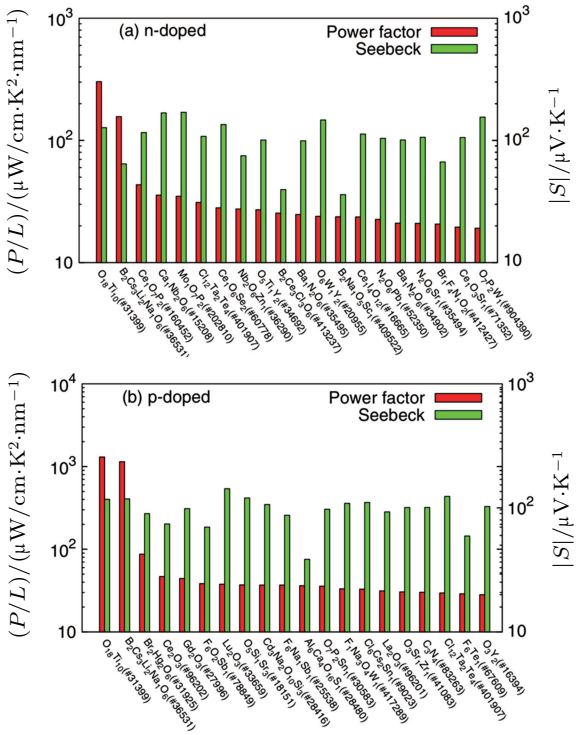


Fig. 2. (color online) Selected 20 best nanosintered thermoelectric materials from high-throughput screening. (a) n-doped and (b) p-doped compounds, along with the calculated averaged power factor (P) and Seebeck coefficient at optimal doping. Adapted from Ref. [7].

2.2. Topological insulator

Topological insulators (TIs) have great potential for nanoelectronics applications because of their novel materials properties; *i.e.*, coexistence of conducting surface states protected by time-reversal-symmetry and an insulating bulk gap.[38–41] To identify novel TIs, Lin *et al.* investigated ternary thermoelectric half-Heusler compounds.[42] Half-Heusler compounds have a cubic lattice structure with a space group $F\bar{4}3m$ and a chemical formula of XYZ, where X and Y are transition metal elements and Z is a p-block element. The authors found that a large majority of the known Heuslers such as TiNiSn and LuNiBi are topological trivial, while the distorted LnPtSb-type compounds such as LnPtBi or LnPdBi (Ln = f^n lanthanides) are topological nontrivial. Similarly, Chadov *et al.* found around 50 half-Heusler compounds with a band inversion similar to that of HgTe, and proposed that the topological states in these zero-gap semiconductors can be achieved by applying strain or designing an appropriate quantum-well structure.[43] These two works show that half-Heuslers can serve a new platform for the realization of multifunctional topological devices.

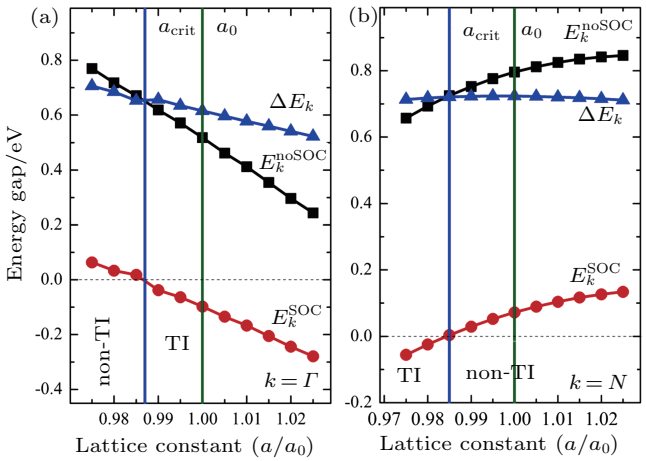


Fig. 3. (color online) Material descriptor for the high-throughput screening of topological insulators.[9] The extraction of the TI robustness descriptor from band energy versus strain variations in SOC and noSOC calculations is shown. E_k^{noSOC} are E_k^{SOC} are the band gaps (at k point) from SOC and noSOC calculations, and ΔE_k is defined as the energy difference between E_k^{noSOC} are E_k^{SOC}. (a) Bi_2Te_2S at $k=\Gamma$, and (b) PbTe at $k=N$ of the body-centered tetragonal Brillouin zone.[27] The vertical green line shows the relaxed lattice constant from DFT calculations, a_0, whereas the vertical blue line indicates the critical value, a_{crit}, for the TI to non-TI transition. Adapted from Ref. [9]

In a later study, instead of screening a special class of compounds, Yang *et al.* scanned the electronic band structures of more than 15000 inorganic crystal structures and discovered 28 TIs (some of them already known) on the basis of the quantum materials repository AFLOWLIB.[20] It is known that the fundamental mechanism of protected conducting surface states involves a band inversion caused by spin–orbit coupling (SOC).[9] Accordingly, the central idea of high-throughput screening for TIs is to identify the key physical parameters that are related to the band inversion. In this work, by analyzing the fundamentals of TIs, Yang *et al.* developed a group of materials descriptors and implemented them into an algorithm for the efficient high-throughput search of TIs. The material descriptors include: (i) a relatively small band gap (less than 1.5 eV) calculated from noSOC density functional theory (DFT) calculations, because that SOC is more likely to lead to an inverted band order in a compound with a smaller band gap; (ii) the compounds should contain heavy elements, the underlying mechanism here is that the heavy elements have strong SOC effects and thus larger ability to flip a band order; and (iii) a "TI robustness descriptor" that measures the robustness of the TI, see Fig. 3. The initial motivation of developing

TI robustness descriptor is to make effective use of the electronic structure data of more than 15000 inorganic compounds already available in the AFLOWLIB. Yang *et al.* studied band energy variations as the lattice strain using noSOC and SOC calculations, respectively; *i.e.*, E_k^{noSOC} and E_k^{SOC}. They found that the SOC band energy difference (E_k^{SOC}) could change sign at some time-reversal invariant k-points when the lattice parameter crosses a critical value. This is a direct signal of band inversion. More interestingly, the authors found that the energy difference between E_k^{noSOC} and E_k^{SOC} (ΔE_k) varies much less than E_k^{noSOC} and E_k^{SOC} (even a constant for some compounds such as PbTe) over a wide range of surface strain. The fundamental reason behind this phenomenon is that most of the SOC effects come from core electrons and will not be affected (or less affected) during the bond strain. On the basis of this finding, the authors approximated

$$\frac{\delta E_k^{\mathrm{SOC}}(a)}{\delta(a)} \approx \frac{\delta E_k^{\mathrm{SOC}}(a)}{\delta(a)|_{a_0}} \approx \frac{\delta E_k^{\mathrm{noSOC}}(a)}{\delta(a)|_{a_0}},$$

and introduced the variational "high-throughput TI robustness descriptor":

$$\hat{\chi}_{\mathrm{TI}} \equiv -\frac{E_k^{\mathrm{SOC}}(a_0)/a_0}{\delta E_k^{\mathrm{noSOC}}(a)/\delta(a)|_{a_0}}.$$

Its physical meaning is a "strain" that represents either the robustness or feasibility of the TI state. If a compound is already a TI under its equilibrium condition, then $\hat{\chi}_{\mathrm{TI}}$ indicates a minimum biaxial strain to lose the TI state. If a compound is not a TI under its equilibrium condition but might be made so by applying a specific strain, then $\hat{\chi}_{\mathrm{TI}}$ indicates a minimum biaxial strain to obtain the TI state. Using the TI robustness descriptor, the authors scanned the whole AFLOWLIB repository, and eventually identify 28 potential TIs in a high-throughput fashion. These compounds are classified into five groups according to prototype structures and constitutes, and for each compound, the possibility to lose or obtain a TI state from a biaxial strain was evaluated by the $\hat{\chi}_{\mathrm{TI}}$.

2.3. p-type transparent conducting oxides

Transparent conducting oxides (TCOs) have a wide range of applications in electronic devices, such as light-emitting diodes and solar cells. Compared to the currently commercialized n-type TCOs like Sn-doped In_2O_3 (ITO), F-doped SnO_2 (FTO),[44] p-type TCOs have a slow development progress because of great challenges in realizing a combination of high hole conductivity and optical transparency.[45] A fundamental reason of the difficulty of developing p-type TCOs with high mobility is because that 2p orbitals of oxygen anions in most oxides are rather localized and make the bands in the valence band very flat, thus leading to large hole effective masses.

In 2013, Hautier *et al.* reported a high-throughput computational search on binary and ternary oxides and identified several highly promising compounds as p-type host.[11] In this work, the authors have calculated the electronic band structures for 3052 oxides using the relaxed crystalline structures in the Materials Project Database.[30] The authors searched the database using effective mass (< 1.5) as a screening criterion and identified 20 oxides as TCO candidates. Figure 4 shows the effective mass versus the band gap for the 20 compounds. A good p-type TCO should have a large band gap and meanwhile a low hole effective mass, and thus the ideal TCOs should appear in the lower right corner. Next, the authors analyzed p-type dopability of the identified most promising candidates ($Sb_4Cl_2O_5$, K_2SnO_3, and $K_2Pb_2O_3$) by calculating their defect-formation energy as a function of the Fermi energy in oxidizing conditions. They found that $Sb_4Cl_2O_5$ is not likely to be p-type doped because oxygen and chlorine vacancies will form spontaneously and act as hole killers when the Fermi energy is close to valence band. In contrast, $K_2Sn_2O_3$ is one promising p-type TCO because oxygen vacancy will not compensate hole formation, and the presence of low-energy potassium vacancy can lead to an intrinsic p-type behavior. On the basis of the screening results, the authors further proposed two design principles for the p-type TCOs: i) hybridization of a d^{10} ($3d^{10}4s^2$ or $4d^{10}5s^2$) ion with oxygen; ii) presence of an anion (such as S^2, P^3, Cl, or Br) with more delocalized p-orbitals than oxygen. These design principles serve as useful guidance for future searches for high-mobility p-type TCOs.

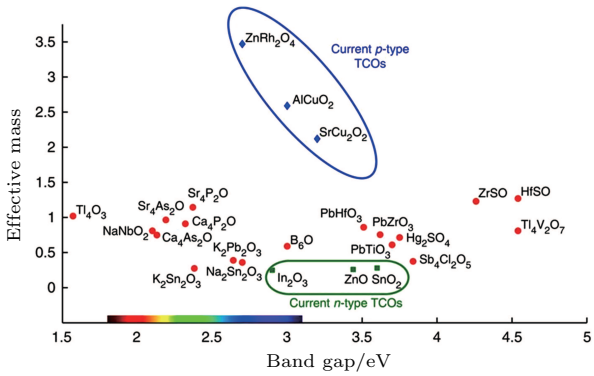


Fig. 4. (color online) Calculated effective mass versus band gap for the p-type TCO candidates. Adapted from Ref. [11].

2.4. Perovskite–oxide-based 2DEG

Two-dimensional electron gas (2DEG) at the interface of perovskite oxide heterostructures has recently attracted much interest because of its potential applications in the next-generation of nanoelectronics.[46–49] The 2DEG at polar/nonpolar $LaAlO_3/SrTiO_3$ heterointerface was firstly reported by Hwang *et al.* in 2004,[46] though the fundamental formation mechanism of the interfacial 2DEG has been on debate since its discovery.[49–57] In spite of the controversy, the polar nature of the overlayer in the perovskite oxide heterostructure still plays an important role in the formation of

the 2DEG in a direct or indirect way.[50,51,57] In fact, with motivations either to search for novel heterostructures with better electron transport property or to explore the roles of the polar perovskite oxides, a series of experimental efforts have been made to explore the interfacial electronic properties in the novel polar/nonpolar perovskite oxide heterostructures such as

$LaGaO_3/SrTiO_3$,[58,59] $NdGaO_3/SrTiO_3$,[57,59,60]
$NdAlO_3/SrTiO_3$,[57,59] $PrAlO_3/SrTiO_3$,[57,59]
and $DyScO_3/SrTiO_3$.[61]

From a computational perspective, by carefully studying the formation mechanisms of the 2DEG at the perovskite interfaces,[44,54,62–70] Yang *et al.* proposes a group of combinatorial descriptors including polar character, lattice mismatch, band gap, and band alignment between the perovskite oxide insulators and $SrTiO_3$ substrate for the high-throughput design of $SrTiO_3$-based 2DEG systems on the basis of perovskite oxide quantum database.[14] These materials descriptors are based on the following arguments:

(i) The polar catastrophe mechanism requires that the heterostructure consists of polar and non-polar perovskite oxides. Note that in this work, for convenience, by using the same naming convention for the polar/nonpolar $(LaO)^+/(TiO_2)^0$ interface in the $LaAlO_3/SrTiO_3$ heterostructure, the authors defined the charged layer of a perovskite oxide along [001] direction like $(LaO)^+$ as polar and the neutral layer like $(TiO_2)^0$ as nonpolar. Accordingly, the polar oxides have a formula like $A^{+1}B^{+5}O_3$ or $A^{+3}B^{+3}O_3$, and the nonpolar oxides have a formula like $A^{+2}B^{+4}O_3$.

(ii) The lattice constant of the perovskite oxide film should be close to that of the substrate ($SrTiO_3$) so that a close lattice mismatch between the epitaxially grown film and its substrate is helpful to minimize defects and to improve the electron mobility.

(iii) The nature of 2DEG requires that the perovskite heterostructure should be conductive along the in-plane direction but insulating in the direction perpendicular to the interfacial plane. This condition requires that the 2DEG must be formed at the interface between two wide-band-gap oxides.

(iv) According to the polar catastrophe mechanism, the 2DEG at the interface of $SrTiO_3$-based heterostructure was caused by the charge transfer from the polar film to the interfacial TiO_2 layer of the $SrTiO_3$ substrate. This requires that the conduction band bottom of $SrTiO_3$ should be lower than that of the electron-donor oxide, *i.e.*, polar oxide, so that the transferred electrons can be accumulated near the conduction band bottom, forming n-type interfacial conductivity.

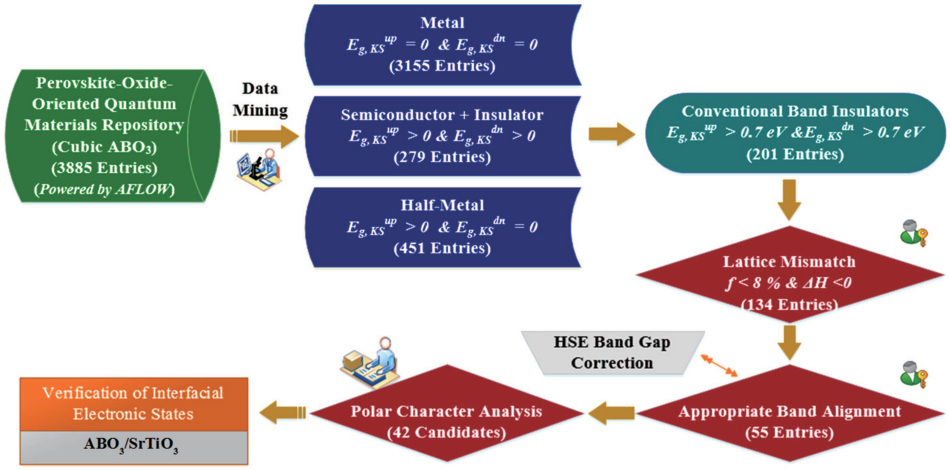


Fig. 5. (color online) High-throughput design of two-dimensional electron gas systems via polar catastrophe mechanism from large-scale first-principles calculations. Adapted from Ref. [14].

Equipped with these combinatorial descriptors, Yang *et al.* carried out a screening of all the polar perovskite compounds, uncovering 42 compounds of potential interests (some of them are already known), see the high-throughput screening procedure in Fig. 5. These candidate perovskite oxides include Ga-based, Al-based, Sc-based, In-based, and Ag-based compounds. By building heterostructure models using these candidate perovskite oxides, the authors further studied their interfacial electronic states, and predicted that Al-, Ga-, Sc-, and Ta-based compounds can form the 2DEG with the $SrTiO_3$ substrate, while In-based compounds cannot. This happens because the compressive-strain-induced polarization in the film neutralizes the polar catastrophe and prevent the formation of 2DEG. This work opens new avenues for the discovery of perovskite-oxide-based functional interface materials in a high-throughput fashion by defining materials descriptors solely based on the bulk materials properties on the basis of the perovskite-oriented quantum materials repository. In addition, as one reference for designing future functional materials, several gene elements can be identified: Al, Ga, and Sc, are *gene elements* for designing polar/nonpolar interface to produce 2DEG, while In could be one key *gene element* for designing piezoelectric materials.

In addition to the polar catastrophe, another possible mechanism to produce 2DEG is via polarization discontinuity.[71–73] The polarization has been found to lead

to 2DEG in traditional semiconductor materials such as ZnMgO/ZnO,[16,17] but few efforts have been made to explore possibility of producing 2DEG in the perovskite-oxide heterostructures through polarization discontinuity. In 2016, Chen *et al.* have observed a highly mobile 2DEG in the nonpolar/nonpolar $CaZrO_3/SrTiO_3$ perovskite oxide heterostructures.[74] Later first-principles computational study has confirmed that the 2DEG can be created by the lattice-mismatch-induced polarization discontinuity between the film and substrate.[64] Compared to the great success of producing 2DEG in the polar/nonpolar perovskite heterostructure systems,[14,57–61] the two-dimensional electron gas at the nonpolar/nonpolar perovskites interface remains rarely explored. Moreover, the polarization-discontinuity-induced 2DEG has a great advantage as compared to that induced by polar catastrophe. That is, the polarization in the film can be controlled by external electrical filed more efficiently, thus bringing enhanced functionalities. In addition, many perovskite oxides are equipped with ferroelectric/piezoelectric property, so to search for perovskite-based 2DEG systems via polarization discontinuity is of great fundamental and practical interest.

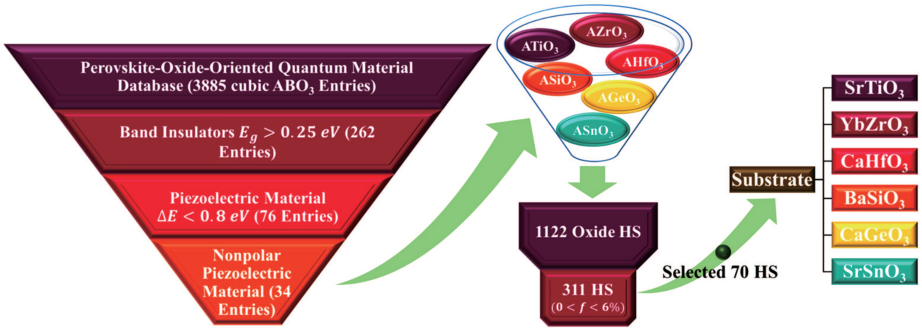


Fig. 6. (color online) High-throughput design of two-dimensional electron gas systems via polarization discontinuity from large-scale first-principles calculations. Adapted from Ref. [15].

In a recent study, Cheng *et al.* carried out a high-throughput design of perovskite-oxide-based 2DEG systems via polarization discontinuity mechanism using large-scale first-principles calculations, see the screening procedure in Fig. 6. In this work, by using a group of combinatorial materials descriptors including band gap, piezoelectric property, and nonpolar character, and high-throughput first-principles calculations, the authors have successfully identified 34 nonpolar piezoelectric perovskite oxides. These oxides can be divided into six groups based on their constituents: Ti-, Zr-, Hf-, Si-, Ge-, and Sn-based oxides. These results indicate that Ti, Zr, Hf, Si, Ge, and Sn, are the key *gene elements* for the piezoelectric oxide materials, which could be an important reference in future design of piezoelectric materials. It is noted that Armiento *et al.* has carried out an early work on the high-throughput screening of bulk perovskite alloys,[12] in which the discovered relevant compounds are also well consistent with our heterostructure calculations. Note that, in this work, only the nonpolar oxides are considered, and thus the trivalent elements such as In are not highlighted here. However, as discussed in the section of high-throughput of 2DEG via polar catastrophe mechanism, In and Ta are also effective *gene elements* for designing piezoelectric oxide materials. The authors have further built 1122 unique heterostructures by combining these 34 compounds. Next, by using a simple but critical material descriptor, lattice mismatch $0 < f < 6\%$, they were able to shrink the screening space of the heterostructures from 1122 to 311. It is worth mentioning that the lattice mismatch here is defined as $f = (a_f - a_s)/a_s$, and the positive f means that the film will experience a compressive strain from the substrate during the epitaxial film growth process, a necessary condition to produce polarization in the film. Cheng *et al.* further performed first-principles electronic structure calculations for 70 selected heterostructures and validated the formation of the 2DEG in most of the selected heterostructures. This work demonstrates an efficient way to accelerate the design of perovskite-based functional materials.

2.5. Halide perovskites

The organic-inorganic halide perovskites are an emerging class of semiconductor materials that have many promising optoelectronic applications.[75–81] In spite of their exceptional optoelectronic properties, particularly for photovoltaic applications, the organic-inorganic halide perovskites face major challenges because of their poor stability and presence of toxic lead, which limits their large-scale applications. To overcome these challenges, one solution is to search for novel hybrid materials with potentially superior properties beyond or like that of lead-based organohalide perovskites.[82,83]

In 2015, Filip *et al.* screened all the possible homovalent metal ions over the entire periodic table for replacing the lead in the perovskite halide structural configuration.[16] The screening process is shown in Fig. 7. In this work, as the first step, the authors selected 29 possible metal ions with +2 oxidation state and generated a total of 116 possible metal–halide combinations by excluding all rare-earth and radioactive metals. Next, the authors performed the structural relaxations from five ideal unit cells constructed with five types of metal-halide-metal bond angles, and the compound with the lowest energy configuration was selected for

further electronic structure calculations because structural optimizations of $CH_3NH_3PbI_3$ are sensitive to the starting point geometry.[84–86] The third step is to select semiconductor compounds with band gaps less than 3.5 eV, which reduces the total number of the compounds to 40. The authors next probe the dynamic stability of these newly formed compounds by displacing the atomic positions and lattice parameters of each unit cell randomly and relaxing the structures. The difference between the maximum and minimum metal-halide distance (Δ) within an octahedron of the lowest energy structures was calculated to evaluate the degree of distortion. By excluding all structures with $\Delta > 0.5$ Å, the total number of compounds was further reduced to 32. They next calculated the band gaps of remaining structures using DFT calculations including spin–orbit coupling and lowered the band gap threshold from 3.5 eV to 2 eV. After an analysis of the compounds that satisfy this criterion, the elements Ge^{2+}, Sn^{2+}, Mg^{2+}, V^{2+}, Mn^{2+}, Ni^{2+}, Cd^{2+}, Hg^{2+}, Ca^{2+}, and In^{2+} were considered as novel candidates for the replacement of Pb in metal-halide perovskites. Next, by considering band gap types (direct band gap only) and the criterion for reducing toxicity, Mg, Mn, V, and Ni were selected as potential alternatives to Pb in lead-halide perovskites. In particular, the magnesium-iodine perovskite was proposed to be an efficient electron conductor due to low electron effective masses.

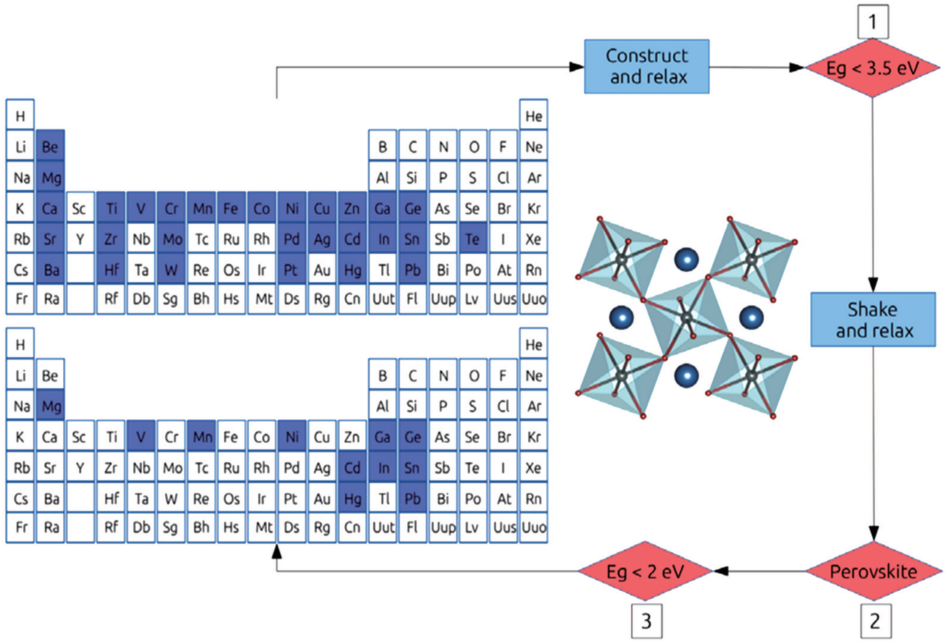


Fig. 7. (color online) Scheme of the computational screening process. The initial input elements and selected candidate elements are shown in the blue shading on the periodic table. The numbers 1, 2, and 3 mark the steps of the screening process, with the screening criteria. Adapted with permission from Ref. [16] Copyright (2015) American Chemical Society.

In a later research work, Jain *et al.* screened lead-free perovskite-like materials with compositions $A_2BB'X_6$, ABX_4, and $A_3B_2X_9$ for optoelectronic performance using high-throughput DFT calculations.[87] In this research, the monovalent A and B' cations were selected from Na, Ka, Rb, Cs, Cu, and Ag, and trivalent B cations were chosen from Ga, In, and Sb, with monovalent anions $X =$ Cl, Br, and I. The authors employed two material descriptors including formation energy and hybrid HSE06 band gaps and found 10 compounds with bandgaps in the range of 1.5 eV–2.5 eV from more than 480 compounds. To evaluate their optoelectronic applications, the authors further characterize effective masses, density of states, and absorption coefficients of all these selected compounds.

2.6. Light-emitting diodes

Phosphor-converted white light-emitting diodes (pc-LEDs) are emerging as an important light source for solid-state lighting because of their high efficiency and long lifetime.[88] To produce a white light LED with a better performance, an efficient red-emitting phosphor is desired[17] Wang *et al.* recently reported a high-throughput search for new narrow-band red-emitting phosphors by developing a group of quantitative descriptors for Eu^{2+}-activated emission on the basis of the electronic structure database Materials Project.[17] An ideal red-emitting phosphor host material should have properties such as high phase stability, red-orange-light emission, high thermal quenching resistance, and high photoluminescence quantum efficiency. In light of these material properties, the authors set a group of materials descriptors: i) $E_{\text{hull}} < 50$ meV, whereas E_{hull} is defined as the energy above the linear combination of stable phases in the phase diagram; ii) calculated band gap (E_g) at the Perdew–Burke–Ernzerhof (PBE) functional[89] level should be in the range from 2.42 eV to 3.58 eV, this criterion was chosen based on the analysis

of the band gap data of 10 well-known red-emitting phosphor hosts; iii) energy splitting, $\Delta E_S > 0.1$ eV, whereas ΔE_S is defined as the energy gap between the highest and the second highest Eu^{2+}4f band. This feature leads to a narrow-band emission; and iv) the authors calculated the Debye temperature (Θ_D) using quasi-harmonic approximation for estimating structural rigidity and band gap using screened hybrid Heyd–Scuseria–Ernzerhof (HSE) functional,[90] which is in much better agreement (within 0.3 eV) with the experimental values.

The flowchart of the high-throughput screening procedure for narrow-band red-emitting phosphor hosts is shown in Fig. 8. By using this procedure, the authors screened 2259 materials (including 203 ternary, 156 quaternary, and 1900 nitridosilicate and nitriodoaluminate quarternary structures), and eventually found a total of eight narrow-band-emitting phosphor hosts. Three of them have already been previously reported experimentally, and five of identified hosts are new.

A high-throughput virtual screening and experimental approach was also employed to design high-performance molecular organic light-emitting diodes.[18] In this work, the authors explored a search space of 1.6 million molecules and screened over 4×10^5 of them using time-dependent density functional theory calculations, and eventually identified nearly one thousand of molecules across the visible spectrum for light-emitting diode applications. The workflow of materials discovery for this work is shown in Fig. 9.

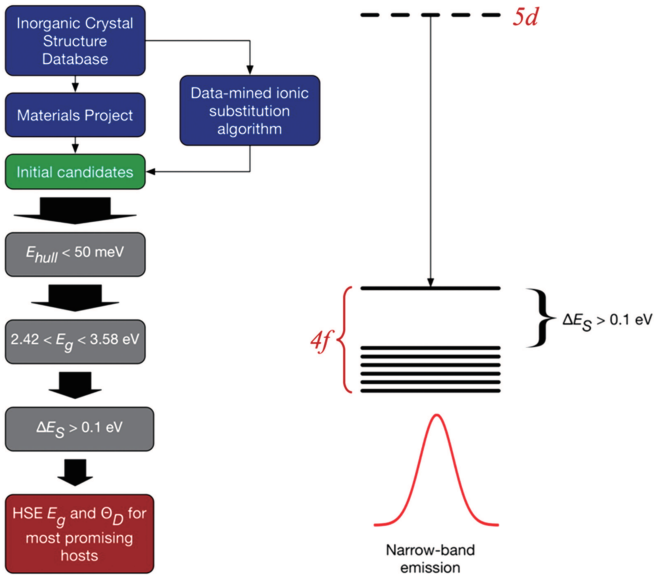


Fig. 8. (color online) (a) Flowchart of the high-throughput screening for narrow-band red-emitting phosphor hosts, and (b) scheme of the narrow-band electron emission of Eu^{2+}5d–4f. Adapted with permission from ref.[17] Copyright (2016) American Chemical Society.

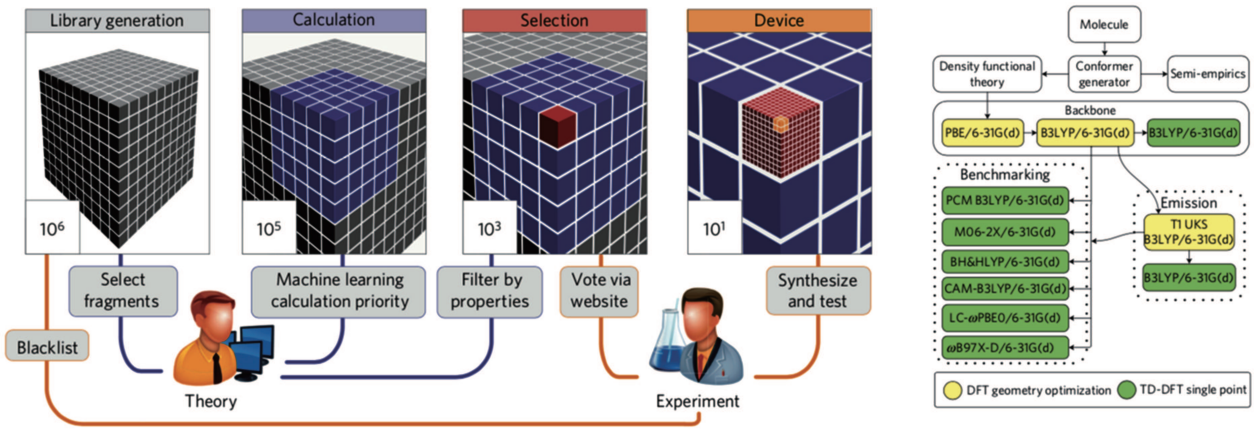


Fig. 9. (color online) Workflow of materials discovery. (a) As shown is the diagram of collaborative materials discovery approach, whereas the search space reduces by over five orders of magnitudes with the screening process. It involves different theoretical and computational approaches and experimental input and testing from left to right. (b) Dependency graph of the quantum chemistry calculations in this study. Adapted from Ref. [18].

This project involves several consequent steps. First, library generation of molecules. The authors produced more than 1.6 million candidates in their library by setting two major conditions: First, the molecules should have a low enough singlet-triplet energy gap. Second, molecular size should be less than 1100 g·mol^{-1}, limited by the requirement for vapor processing. Second, machine learning for prescreening. Owning to the large size of over 1.6 million candidate molecules, it is not possible or necessary to screen all of them with quantum simulation. To speed up the screening efficiency, by using results from previous calculations as training data, the authors employed machine learning approach to do a pre-screening, which is able to predict molecules with a high probability to have good outcomes and prioritize the molecules for next-step simulation. Third, quantum chemical calculations. The authors estimated the emission color via vertical absorption energies using time-dependent density functional theory (TD-DFT) calculations. Fourth, experimental calibrations. To get accurate excitation energy of charge-transfer excitations, the authors employed B3LYP functional by including an amount of exact Hartree–Fock exchange, since the standard TD-DFT calculations often underestimate the excitation energy because of the semi-local nature of the exchange–correlation functional.[91,92] The authors fur-

ther confirmed the accuracy of this approach by using a linear calibration scheme on 46 experimental data points from the past literature, and their calibration against 17 in-house measurements were verified. Finally, analysis and lead discovery. The authors' quantum calculations revealed thousands of candidate emitters with potentially high efficiency, with about 900 as extremely promising. At this step, it is possible to make human tractable decision to choose candidate materials, and the authors used a web-based selection process including data visualization and sorting interfaces to do this. Eventually, a small consensus set of molecules was selected for further experimental synthesis and characterization in devices.

2.7. Heusler-based magnets

Magnetic materials have played an important role in modern technologies from data storage to energy conversion.[19,93] Recently, Sanvito *et al.* demonstrated a high-throughput materials design approach to search for novel magnetic materials based on Heusler alloys (HAs).[19] In this study, on the basis of an extensive electronic structure library of HAs, the authors employed a series of material descriptors including magnetic order and materials stability, and eventually found about 20 novel magnetic materials. Figure 10 shows the high-throughput design procedure of Heusler-based magnets. In this work, the high-throughput materials design strategy consists of three main steps:

(I) The authors built an extensive electronic structure database that contains 236115 compounds based on the structure prototype of HAs. The authors choose to search for new magnets from HAs based on two reasons. First, many HAs contain magnetic ions and therefore it is likely to find candidate magnets from HAs. Second, HA is one large class of materials that can be formed from a wide variety of chemical compositions, which corresponds to a large number of prototypical compounds. These two features give authors a high chance to find new magnets.

(II) The authors performed a preliminary screening using two materials descriptors: formation enthalpy and magnetic moment. They analyzed the formation enthalpy for all the 236115 calculated compounds and set the first screening criterion of $\Delta H = H_{X_2YZ} - (2H_X + H_Y + H_Z) < 0$. This screening condition gives authors 35602 compounds, in which 6778 carry a magnetic moment. This number could be much larger than the actual number of stable magnetic HAs since this preliminary screening did not assess the stability of a given X_2YZ structure against decomposition over all the possible binary and ternary prototypes. After some preliminary analysis, the authors found that it is an extremely challenging task to map the stability of every calculated HA because it requires the calculation of approximately 15 million prototypes.

(III) The authors employed an alternative approach to screen the entire database after the preliminary screening. They narrowed down the entire database to intermetallic HAs consisting of elements of 3d, 4d, and 5d periods, with the number reduction from original 23615 to 36540. Next, on the basis of available energy data of corresponding binaries and known ternary compounds in the AFLOWLIB, the authors conducted a convex hull analysis, and found that, among the 36540 compounds, only 248 are thermodynamically stable.

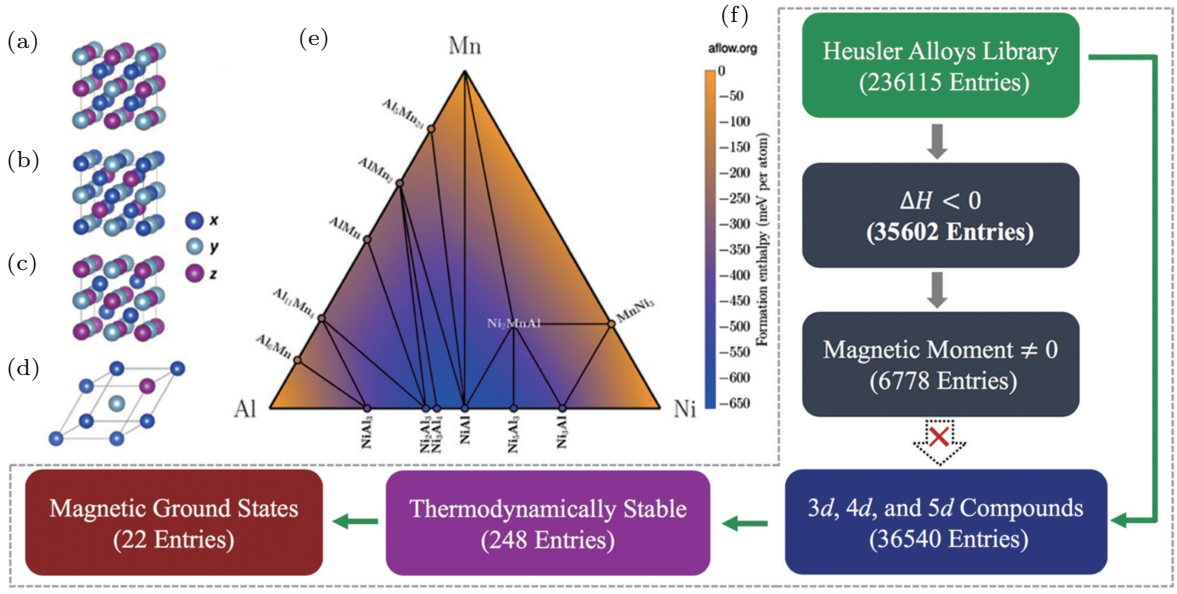


Fig. 10. (color online) High-throughput design of Heusler-based magnets. Structural models of (a) regular Heusler, (b) inverse Heusler, and (c) half Heusler. (d) Primitive unit cell to construct electronic structure database. (e) Convex hull diagram for Al–Mn–Ni. (f) High-throughput workflow for searching for Heusler-based magnets. Adapted and redesigned from Ref. [19]

(IV) Last, the authors found 22 compounds having a magnetic ground state among the 248 compounds from first-principles electronic structure calculations. Interestingly, 20 new magnetic HAs belong to the Co_2YZ, Mn_2YZ, and X_2MnZ classes, implying an important role of Co and Mn elements. Accordingly, Co and Mn can be considered as *gene elements* of magnets in future high-throughput search of new magnets.

In this work, the employed screening strategy is to first search for thermodynamically stable HAs instead of HAs with ferromagnetic ground state, which significantly improves the screening efficiency. In addition, to validate the computational predication, the authors attempted the synthesis of four HAs: two of them (Co_2MnTi and Mn_2PtPd) have been successfully synthesized, whereas the other two (Mn_2PtCo and Mn_2PtV) are decomposed into binary compounds. Notably, the authors found that the Co_2MnTi has a Curie temperature (T_C) as high as 938 K. This remarkable finding has proven that a new high-temperature ferromagnetic can be discovered using high-throughput approach.

3. Summary

In this review, we discussed the basic workflow of discovering and developing new materials using a high-throughput materials design approach (materials genome approach), with an emphasis on the rational design of a high-throughput screening procedure and the development of materials descriptors. Although this review concentrates on the electronic and magnetic properties of functional materials, similar workflows and screening procedures are expected to be applied in all other types of materials. Meanwhile, different computational levels and different material descriptors are usually required.

A successful high-throughput materials design requires a research platform that closely integrates three basic components: computations, experiments, and large-size digital materials data. It is now becoming relatively routine to produce large-size digital materials data using emerging high-throughput computational tools with the support of mature software development techniques. Compared to the construction of electronic materials database using high-throughput calculations, a more challenging task is to develop appropriate and effective materials descriptors (materials genes) to rapidly search for target materials with desired properties. In principle, the development of materials descriptors is based on a deep understanding of the fundamental physical and chemical properties of the target materials and the physical problems at hand. In other words, this strategy is a mechanism-driven materials discovery with the support of the large-size digital materials data. However, sometimes, due to great complexity of some physical problems and materials phenomena, state-of-the-art computational techniques may be unable to produce desired materials information and, therefore, the development of effective materials descriptors becomes an impossible mission. In this case, an alternative way to perform high-throughput search for target materials is to use machine-learning approach, in spite of its limited accuracy, which is also a rapidly expanded research area in the materials science. This research strategy, which purely relies on large-size materials data rather than extracting materials descriptors from the fundamental mechanism, is more like a data-driven materials discovery. Consequently, one may ask which strategy we should choose? My answer is that it depends on the specific physical problem and size of the data.

Acknowledgement

The author acknowledges support by National Science Foundation under award number ACI-1550404, American Chemical Society Petroleum Research Fund under the award number 55481-DNI6, Global Research Outreach (GRO) Program of Samsung Advanced Institute of Technology under the award number 20164974, and the Vannevar Bush Faculty Fellowship program sponsored by the Basic Research Office of the Assistant Secretary of Defense for Research and Engineering under the Office of Naval Research grant N00014-16-1-2569.

References

[1] 1979 *Materials Science in Energy Technology* (Elsevier Inc.)
[2] Kalil T and Wadia C 2011 *Tech. Rep. June* (National Science and Technology Council)
[3] Boussie T R, Diamond G M, Goh C, Hall K A, LaPointe A M, Leclerc M, Lund C, Murphy V, Shoemaker J A W, Tracht U, Turner H, Zhang J, Uno T, Rosen R K and Stevens J C 2003 *J. Am. Chem. Soc.* **125** 4306
[4] Greeley J, Jaramillo T F, Bonde J, Chorkendorff I and Norskov J K 2006 *Nat. Mater.* **5** 909
[5] Ceder G 2010 *MRS Bull.* **35** 693
[6] Derenzo S, Bizarri G, Borade R, Bourret-Courchesne E, Boutchko R, Canning A, Chaudhry A, Eagleman Y, Gundiah G, Hanrahan S, Janecek M and Weber M 2011 *Nucl. Instrum. Methods Phys. Res. Sect. A* **652** 247
[7] Wang S D, Wang Z, Setyawan W, Mingo N and Curtarolo S 2011 *Phys. Rev. X* **1** 021012
[8] Armiento R, Kozinsky B, Fornari M and Ceder G 2011 *Phys. Rev. B* **84** 014103
[9] Yang K, Setyawan W, Wang S, Buongiorno Nardelli M and Curtarolo S 2012 *Nat. Mater.* **11** 614
[10] Curtarolo S, Hart G L W, Nardelli M B, Mingo N, Sanvito S and Levy O 2013 *Nat. Mater.* **12** 191
[11] Hautier G, Miglio A, Ceder G, Rignanese G M and Gonze X 2013 *Nat. Commun.* **4** 2292
[12] Armiento R, Kozinsky B, Hautier G, Fornari M and Ceder G 2014 *Phys. Rev. B* **89** 134103
[13] Gautier R, Zhang X, Hu L, Yu L, Lin Y, Sunde T O, Chon D, Poeppelmeier K R and Zunger A 2015 *Nat. Chem.* **7** 308
[14] Yang K, Nazir S, Behtash M and Cheng J 2016 *Sci. Rep.* **6** 34667
[15] Cheng J and Yang K 2018 *J. Mater. Chem. C* **6** 6680
[16] Filip M R and Giustino F 2016 *J. Phys. Chem. C* **120** 166
[17] Wang Z, Chu I H, Zhou F and Ong S P 2016 *Chem. Mater.* **28** 4024
[18] Gomez-Bombarelli R, Aguilera-Iparraguirre J, Hirzel T D, Duvenaud D, Maclaurin D, Blood-Forsythe M A, Chae H S, Einzinger M, Ha D G, Wu T, Markopoulos G, Jeon S, Kang H, Miyazaki H, Numata M, Kim S, Huang W, Hong S I, Baldo M, Adams R P and Aspuru-Guzik A 2016 *Nat. Mater.* **15** 1120
[19] Sanvito S, Oses C, Xue J, Tiwari A, Zic M, Archer T, Tozman P, Venkatesan M, Coey M and Curtarolo S 2017 *Sci. Adv.* **3** e1602241
[20] Curtarolo S, Setyawan W, Wang S, Xue J, Yang K, Taylor R H, Nelson L J, Hart G L W, Sanvito S, Buongiorno-Nardelli M, Mingo N and Levy O 2012 *Comput. Mater. Sci.* **58** 227

[21] Saal J E, Kirklin S, Aykol M, Meredig B and Wolverton C 2013 *JOM* **65** 1501
[22] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G and Persson K A 2013 *APL Mater.* **1** 011002
[23] Landis D D, Hummelshoj J S, Nestorov S, Greeley J, Dulak M, Bligaard T, Norskov J K and Jacobsen K W 2012 *Comput. Sci. Eng.* **14** 51
[24] Ong S P, Richards W D, Jain A, Hautier G, Kocher M, Cholia S, Gunter D, Chevrier V L, Persson K A and Ceder G 2013 *Comput. Mater. Sci.* **68** 314
[25] Ong S P, Cholia S, Jain A, Brafman M, Gunter D, Ceder G and Persson K A 2015 *Comput. Mater. Sci.* **97** 209
[26] Hjorth Larsen A, Jorgen Mortensen J, Blomqvist J, Castelli I E, Christensen R, Dulak M, Friis J, Groves M N, Hammer B, Hargus C, Hermes E D, Jennings P C, Bjerre Jensen P, Kermode J, Kitchin J R, Leonhard Kolsbjerg E, Kubal J, Kaasbjerg K, Lysgaard S, Bergmann Maronsson J, Maxson T, Olsen T, Pastewka L, Peterson A, Rostgaard C, Schiotz J, Schutt O, Strange M, Thygesen K S, Vegge T, Vilhelmsen L, Walter M, Zeng Z and Jacobsen K W 2017 *J. Phys.: Condens. Matter* **29** 273002
[27] Setyawan W and Curtarolo S 2010 *Comput. Mater. Sci.* **49** 299
[28] Yang X, Wang Z, Zhao X, Song J, Zhang M and Liu H 2018 *Comput. Mater. Sci.* **146** 319
[29] Cheng J, Luo J and Yang K 2018 *Comput. Mater. Sci.* **155** 92
[30] Jain A, Hautier G, Moore C J, Ping Ong S, Fischer C C, Mueller T, Persson K A and Ceder G 2011 *Comput. Mater. Sci.* **50** 2295
[31] Bell L E 2008 *Science* **321** 1457
[32] Poudel B, Hao Q, Ma Y, Lan Y, Minnich A, Yu B, Yan X, Wang D, Muto A, Vashaee D, Chen X, Liu J, Dresselhaus M S, Chen G and Ren Z 2008 *Science* **320** 634
[33] Joshi G, Lee H, Lan Y, Wang X, Zhu G, Wang D, Gould R W, Cuff D C, Tang M Y, Dresselhaus M S, Chen G and Ren Z 2008 *Nano Lett.* **8** 4670
[34] Wang X W, Lee H, Lan Y C, Zhu G H, Joshi G, Wang D Z, Yang J, Muto A J, Tang M Y, Klatsky J, Song S, Dresselhaus M S, Chen G and Ren Z F 2008 *Appl. Phys. Lett.* **93** 193121
[35] Xie W, Tang X, Yan Y, Zhang Q and Tritt T M 2009 *Appl. Phys. Lett.* **94** 102111
[36] Madsen G K 2006 *J. Am. Chem. Soc.* **128** 12140
[37] Bergerhoff G and Brown I D 197 *International Union of Crystallography* 77
[38] Moore J E 2010 *Nature* **464** 194
[39] Moore J 2009 *Nat. Phys.* **5** 378
[40] Qi X L and Zhang S C 2011 *Rev. Mod. Phys.* **83** 1057
[41] Hasan M Z and Kane C L 2010 *Rev. Mod. Phys.* **82** 3045
[42] Lin H, Wray L A, Xia Y, Xu S, Jia S, Cava R J, Bansil A and Hasan M Z 2010 *Nat. Mater.* **9** 546
[43] Chadov S, Qi X, Kubler J, Fecher G H, Felser C and Zhang S C 2010 *Nat. Mater.* **9** 541
[44] Behtash M, Joo P H, Nazir S and Yang K 2015 *J. Appl. Phys.* **117** 175101
[45] Xu X, Bullock J, Schelhas L T, Stutz E Z, Fonseca J J, Hettick M, Pool V L, Tai K F, Toney M F, Fang X, Javey A, Wong L H and Ager J W 2016 *Nano Lett.* **16** 1925
[46] Ohtomo A and Hwang H Y 2004 *Nature* **427** 423
[47] Chakhalian J, Millis A J and Rondinelli J 2012 *Nat. Mater.* **11** 92
[48] Mannhart J and Schlom D G 2010 *Science* **327** 1607
[49] Schlom D G and Mannhart J 2011 *Nat. Mater.* **10** 168
[50] Liu Z Q, Li C J, Lü W M, Huang X H, Huang Z, Zeng S W, Qiu X P, Huang L S, Annadi A, Chen J S, Coey J M D, Venkatesan T and Ariando 2013 *Phys. Rev. X* **3** 021010
[51] Yu L and Zunger A 2014 *Nat. Commun.* **5** 5118
[52] Nakagawa N, Hwang H Y and Muller D A 2006 *Nat. Mater.* **5** 204
[53] Pentcheva R and Pickett W 2009 *Phys. Rev. Lett.* **102** 107602
[54] Behtash M, Nazir S, Wang Y and Yang K 2016 *Phys. Chem. Chem. Phys.* **18** 6831
[55] Willmott P R, Pauli S A, Herger R, Schlepütz C M, Martoccia D, Patterson B D, Delley B, Clarke R, Kumah D, Cionca C and Yacoby Y 2007 *Phys. Rev. Lett.* **99** 155502
[56] Eckstein J N 2007 *Nat. Mater.* **6** 473
[57] Annadi A, Putra A, Liu Z Q, Wang X, Gopinadhan K, Huang Z, Dhar S, Venkatesan T and Ariando 2012 *Phys. Rev. B* **86** 085450
[58] Perna P, Maccariello D, Radovic M, Scotti di Uccio U, Pallecchi I, Codda M, Marré D, Cantoni C, Gazquez J, Varela M, Pennycook S J and Granozio F M 2010 *Appl. Phys. Lett.* **97** 152111
[59] Li C, Liu Z, Lu W, Wang X R, Annadi A, Huang Z, Zeng S, Ariando and Venkatesan T 2015 *Sci. Rep.* **5** 13314
[60] Li C, Xu Q, Wen Z, Zhang S, Li A and Wu D 2013 *Appl. Phys. Lett.* **103** 201602
[61] Li D F, Wang Y and Dai J Y 2011 *Appl. Phys. Lett.* **98** 122108
[62] Wang Y, Tang W, Cheng J, Nazir S and Yang K 2016 *Phys. Chem. Chem. Phys.* **18** 31924
[63] Wang Y, Tang W, Cheng J, Behtash M and Yang K 2016 *ACS Appl. Mater. Interfaces* **8** 13659
[64] Nazir S, Cheng J and Yang K 2016 *ACS Appl. Mater. Interfaces* **8** 390
[65] Nazir S, Behtash M, Cheng J, Luo J and Yang K 2016 *Phys. Chem. Chem. Phys.* **18** 2379
[66] Nazir S, Bernal C and Yang K 2015 *ACS Appl. Mater. Interfaces* **7** 5305
[67] Nazir S, Behtash M and Yang K 2015 *J. Appl. Phys.* **117** 115305
[68] Nazir S, Behtash M and Yang K 2015 *RSC Adv.* **5** 15682
[69] Nazir S and Yang K 2014 *ACS Appl. Mater. Interfaces* **6** 22351
[70] Nazir S, Behtash M and Yang K 2014 *Appl. Phys. Lett.* **105** 141602
[71] Stemmer S and James Allen S 2014 *Ann. Rev. Mater. Res.* **44** 151
[72] Tsukazaki A, Ohtomo A, Kita T, Ohno Y, Ohno H and Kawasaki M 2007 *Science* **315** 1388
[73] Tampo H, Shibata H, Maejima K, Yamada A, Matsubara K, Fons P, Kashiwaya S, Niki S, Chiba Y, Wakamatsu T and Kanie H 2008 *Appl. Phys. Lett.* **93** 202104
[74] Chen Y, Trier F, Kasama T, Christensen D V, Bovet N, Balogh Z I, Li H, Thyden K T, Zhang W, Yazdi S, Norby P, Pryds N and Linderoth S 2015 *Nano Lett.* **15** 1849
[75] Kojima A, Teshima K, Shirai Y and Miyasaka T 2009 *J. Am. Chem. Soc.* **131** 6050
[76] Im J H, Lee C R, Lee J W, Park S W and Park N G 2011 *Nanoscale* **3** 4088
[77] Kim H S, Lee C R, Im J H, Lee K B, Moehl T, Marchioro A, Moon S J, Humphry-Baker R, Yum J H, Moser J E, Gratzel M and Park N G 2012 *Sci. Rep.* **2** 591
[78] Lee M M, Teuscher J, Miyasaka T, Murakami T N and Snaith H J 2012 *Science* **338** 643
[79] Zhou H, Chen Q, Li G, Luo S, Song T-b, Duan H S, Hong Z, You J, Liu Y and Yang Y 2014 *Science* **345** 542
[80] Livraghi S, Paganini M C, Giamello E, Selloni A, Di Valentin C and Pacchioni G 2006 *J. Am. Chem. Soc.* **128** 15666
[81] Li X, Bi D, Yi C, Décoppet J D, Luo J, Zakeeruddin S M, Hagfeldt A and Grätzel M 2016 *Science* aaf8060
[82] Zhao X G, Yang D, Sun Y, Li T, Zhang L, Yu L and Zunger A 2017 *J. Am. Chem. Soc.* **139** 6718
[83] Zhao X G, Yang J H, Fu Y, Yang D, Xu Q, Yu L, Wei S H and Zhang L 2017 *J. Am. Chem. Soc.* **139** 2630
[84] Motta C, El-Mellouhi F, Kais S, Tabet N, Alharbi F and Sanvito S 2015 *Nat. Commun.* **6** 7026
[85] Filippetti A, Delugas P and Mattoni A 2014 *J. Phys. Chem. C* **118** 24843
[86] Mosconi E, Amat A, Nazeeruddin M K, Grätzel M and De Angelis F 2013 *J. Phys. Chem. C* **117** 13902
[87] Jain A, Voznyy O and Sargent E H 2017 *J. Phys. Chem. C* **121** 7183
[88] Uheda K, Hirosaki N, Yamamoto Y, Naito A, Nakajima T and Yamamoto H 2006 *Electrochem. Solid-State Lett.* **9** H22
[89] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[90] Heyd J, Scuseria G E and Ernzerhof M 2003 *J. Chem. Phys.* **118** 8207
[91] Dreuw A and Head-Gordon M 2005 *Chem. Rev.* **105** 4009
[92] Gritsenko O and Baerends E J 2004 *J. Chem. Phys.* **121** 655
[93] Hirosawa S, Nishino M and Miyashita S 2017 *Adv. Nat. Sci: Nanosci. Nanotechnol.* **8** 013002