Title
Candidate ferroelectrics via ab initio high-throughput screening of polar materials

Journal
npj Computational Materials, 10(1)

ISSN
2057-3960

Authors
Ricci, Francesco
Reyes-Lillo, Sebastian E
Mack, Stephanie A
et al.

Publication Date
2024

DOI
10.1038/s41524-023-01193-3

npj computational materials

www.nature.com/npjcompumats

ARTICLE OPEN

Check for updates

Candidate ferroelectrics via ab initio high-throughput screening of polar materials

Francesco Ricci[1,2], Sebastian E. Reyes-Lillo[3], Stephanie A. Mack[1,2] and Jeffrey B. Neaton[1,2,4]✉

Ferroelectrics are a class of polar and switchable functional materials with diverse applications, from microelectronics to energy conversion. Computational searches for new ferroelectric materials have been constrained by accurate prediction of the polarization and switchability with electric field, properties that, in principle, require a comparison with a nonpolar phase whose atomic-scale unit cell is continuously deformable from the polar ground state. For most polar materials, such a higher-symmetry nonpolar phase does not exist or is unknown. Here, we introduce a general high-throughput workflow that screens polar materials as potential ferroelectrics. We demonstrate our workflow on 1978 polar structures in the Materials Project database, for which we automatically generate a nonpolar reference structure using pseudosymmetries, and then compute the polarization difference and energy barrier between polar and nonpolar phases, comparing the predicted values to known ferroelectrics. Focusing on a subset of 182 potential ferroelectrics, we implement a systematic ranking strategy that prioritizes candidates with large polarization and small polar-nonpolar energy differences. To assess stability and synthesizability, we combine information including the computed formation energy above the convex hull, the Inorganic Crystal Structure Database id number, a previously reported machine learning-based synthesizability score, and ab initio phonon band structures. To distinguish between previously reported ferroelectrics, materials known for alternative applications, and lesser-known materials, we combine this ranking with a survey of the existing literature on these candidates through Google Scholar and Scopus databases, revealing ~130 promising materials uninvestigated as ferroelectric. Our workflow and large-scale high-throughput screening lays the groundwork for the discovery of novel ferroelectrics, revealing numerous candidates materials for future experimental and theoretical endeavors.

npj Computational Materials (2024)10:15 ; https://doi.org/10.1038/s41524-023-01193-3

INTRODUCTION

Ferroelectrics are insulating materials with a nonzero electric polarization switchable by an applied electric field. Empirically, a ferroelectric is characterized by a measured polarization versus electric field hysteresis loop, with spontaneous polarization being half of the measured change in polarization at zero electric fields (Fig. 1a). Ferroelectric materials are critical for many technologies and have found applications in electronic devices such as memories, sensors, and capacitors among others[1]. Well-known ferroelectrics include those in the perovskite family, such as $BaTiO_3$, $PbTiO_3$, and $KNbO_3$[2]. Ferroelectric materials with magnetic order, or multiferroics, find applications in which the magnetic order can be altered by an external electric field[3,4]. $BiFeO_3$ is an example of a multiferroic where the electric polarization can be coupled to the magnetic order[5–7].

For more than three decades, ferroelectric materials have been studied with density functional theory (DFT) and the modern theory of polarization[8–14], where extensive progress has been made in understanding the physical and chemical mechanisms underlying the emergence of macroscopic polarization in bulk and finite materials[15,16]. DFT with standard exchange-correlation functionals has also been shown to predict polarization with good accuracy, often matching values measured by experiments within a small error[1].

A necessary condition for having a nonzero spontaneous polarization is that the material possess a polar space group. However, many polar crystals with a nonzero polarization are not switchable, and, therefore, not ferroelectric. Landau-Devonshire theory is often used to rationalize and predict the likelihood the polarization in a material is switchable. Within this framework, a key descriptor for switchability is the energy gained by symmetry-breaking atomic distortion taking a higher-symmetry nonpolar phase to a lower-symmetry polar crystal structure (Fig. 1b). Early studies by Abrahams[17–19] used this descriptor and symmetry criteria to search for switchable polar materials in crystallographic databases predicting candidate ferroelectrics, some of which, more recently, have been synthesized and confirmed[20–23]. Nowadays, Abrahams' analysis can be readily performed using the online software on the Bilbao Crystallographic Server (BCS), such as PSEUDO[24], which has been used in prior work to find a nonpolar reference structure starting from a polar one in certain families of compounds[25–28].

The advent of high-throughput DFT approaches and materials databases has enabled new search strategies and larger-scale high-throughput ferroelectric materials searches[29–34]. Smidt et al.[35] developed the first automated workflow for ferroelectric materials discovery and searched the Materials Project (MP) database for pairs of polar and nonpolar structures connected by a subgroup-group relationship. DFT values of the polarization of hundreds of polar compounds were automatically computed using the modern theory of polarization and an interpolation method for automatically selecting the branch of polarization[13,14], the latter also having been implemented in the Atomate software[36]. For known compounds, the computed polarization was compared with the experiments and found to be in agreement to within 10%, or better. Ultimately, ref. [35] identified 200 ferroelectrics, where 74 were known or previously proposed,

[1]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [2]Department of Physics, University of California, Berkeley, CA, USA. [3]Departamento de Ciencias Físicas, Universidad Andres Bello, Santiago 837-0136, Chile. [4]Kavli Energy NanoSciences Institute at Berkeley, Berkeley, CA, USA. ✉email: jbneaton@lbl.gov

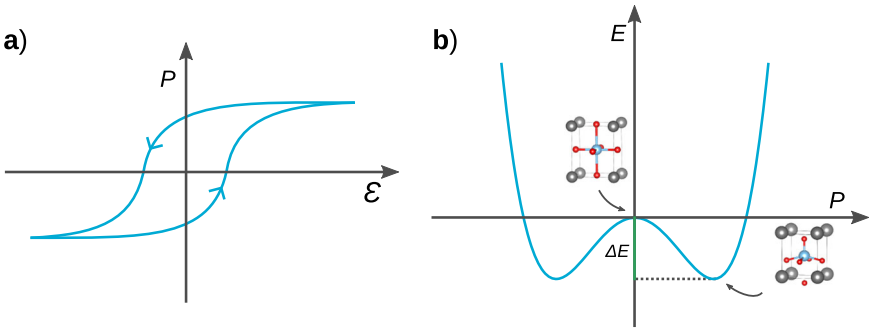

Fig. 1 Schematic representation of the key properties of a ferroelectric. a The measured polarization P as a function of the applied electric field ϵ and **b** the typical double-well profile of the energy E as function of the polarization P of a ferroelectric material. Highlighted in **b** are an atomic-scale schematic of the polar and nonpolar phases and their energy difference ΔE, used as a simple indicator of polarization switchability within the Landau-Devonshire framework.

while 124 were new candidates, some of which were later experimentally investigated[37,38].

Although ref. [35] represents the largest database of potential ferroelectric materials automatically generated via first-principles methods to date, it did not discriminate for synthesizability and stability. Further, it neglected polar compounds that did not have a nonpolar counterpart phase in the database. Since there are ~10,000 polar insulating materials in the MP database without a nonpolar reference phase, relaxing this constraint and generating a nonpolar reference for each polar material would increase the number of candidate ferroelectrics.

Here, we present a new workflow for ferroelectric materials discovery, and use it to compute the polarization of 447 polar materials in the MP database, predicting 182 potential candidate ferroelectric materials, 130 of which are previously uninvestigated. We make use of pseudosymmetries to generate a nonpolar reference structure starting from an initial polar structure, using the PSEUDO tool[24] of the BCS (https://www.cryst.ehu.es/). While pseudosymmetries have been used to find reference nonpolar structures of potential ferroelectrics in the past[39–41], they have yet to be applied in an automated fashion at the large scale we have here. In the present work, we automatically query and extract data from the PSEUDO web interface, enabling the automatic search of a nonpolar reference for any polar structure. In particular, with such a method, we generate a nonpolar reference and compute the polarization for a large set of insulating polar structures present in the MP database, which do not have a nonpolar reference in the MP database. We focus on materials that are non-magnetic in the MP database, and we consider only materials that have formation energy on the convex hull, increasing their relevance to the experiment. The results are publicly available through the MPContribs platform (https://contribs.materialsproject.org/). None of the polar compounds presented here were studied in ref. [35]. Further, for the subset of materials that are most promising as ferroelectrics, we develop a ranking that combines the computed polarization and polar-nonpolar energy difference (Fig. 1b) with an automated literature search and a machine learning-based prediction of synthesizability[42] to provide suggestions for new potential ferroelectrics for experimental investigation.

RESULTS AND DISCUSSION

Screening

We use the MP database for our high-throughput screening for ferroelectric materials. We consider materials in the database that have a polar space group, that are insulating with a DFT-PBE Kohn-Sham band gap higher than 0.1 eV, and that are non magnetic (as reported by the MP). We consider only materials that have energy above hull equal to zero eV per atom, that is, materials that are 'on the hull'. The initial set defined by these criteria contains 1978 materials. We note that many materials reported in the MP to be above but within 0.1 eV per atom of the hull are synthesizable[43,44] (e.g., $BiFeO_3$). The constraint of zero energy above hull can be easily lifted, doubling or more the number of potential candidates; however, as this also significantly increases the computational cost and analysis effort, we leave these systems for future work. Consistent with potential synthesizability, we focus only on binary, ternary, and quaternary compounds with less than 56 atoms per primitive cell. Also, we exclude materials containing rare-earth elements because such elements can possess partially filled f-shells, and strong spin-orbit coupling and can require a case-by-case treatment. Finally, we restrict our focus to polar materials that either do not have a nonpolar reference phase in the MP database or that have such a reference but were not available at the time the database in Ref. [35] was developed. None of the polar compounds presented here were studied in ref. [35].

For all polar materials studied, we automatically generate a nonpolar reference phase using the BCS PSEUDO tool. Among the higher-symmetry structures proposed by PSEUDO, we consider only the nonpolar ones, and we rank them according to their maximum atomic displacement (MAD) relative to the polar phase and select the nonpolar structure that has the smallest MAD. This approach automatically provides nonpolar structures most likely to have minimal energy differences with the polar phase. The cases where PSEUDO returns only polar structures, about 40% of the initial set, are discarded. A natural extension is to re-run the pseudosymmetry search on each of these newly identified polar structures and find the closest nonpolar with higher symmetry. This extension of our workflow would recover ~40% of the compounds, but we relegate it to future work. Finally, since the nonpolar structure in its conventional setting obtained from PSEUDO is already in the same setting as the polar, it can be directly used in the automated workflow. To reduce the computational time whenever possible, we reduce both structures to their primitive unit cell.

The screening of the MP database with the above-mentioned filters and the subsequent pseudosymmetry search of the nonpolar parent structures returns 618 pairs of structures for which we compute the polarization. Eighty of these pairs had a calculation issue at some point in the workflow. Among the remaining 538 completed workflows, 447 were completed successfully and the polarization is computed. In 91 cases, either the nonpolar structure (69 cases) or one of the other structures is metallic, and the polarization calculation was not complete. Figure 2a summarizes the screening process. In the Results section, the 447 compounds that completed the workflow are summarized, along with a brief comment on those with a nonpolar phase that is metallic within DFT-PBE. Fig 2b summarizes the classifications made for these remaining compounds. Compounds that are

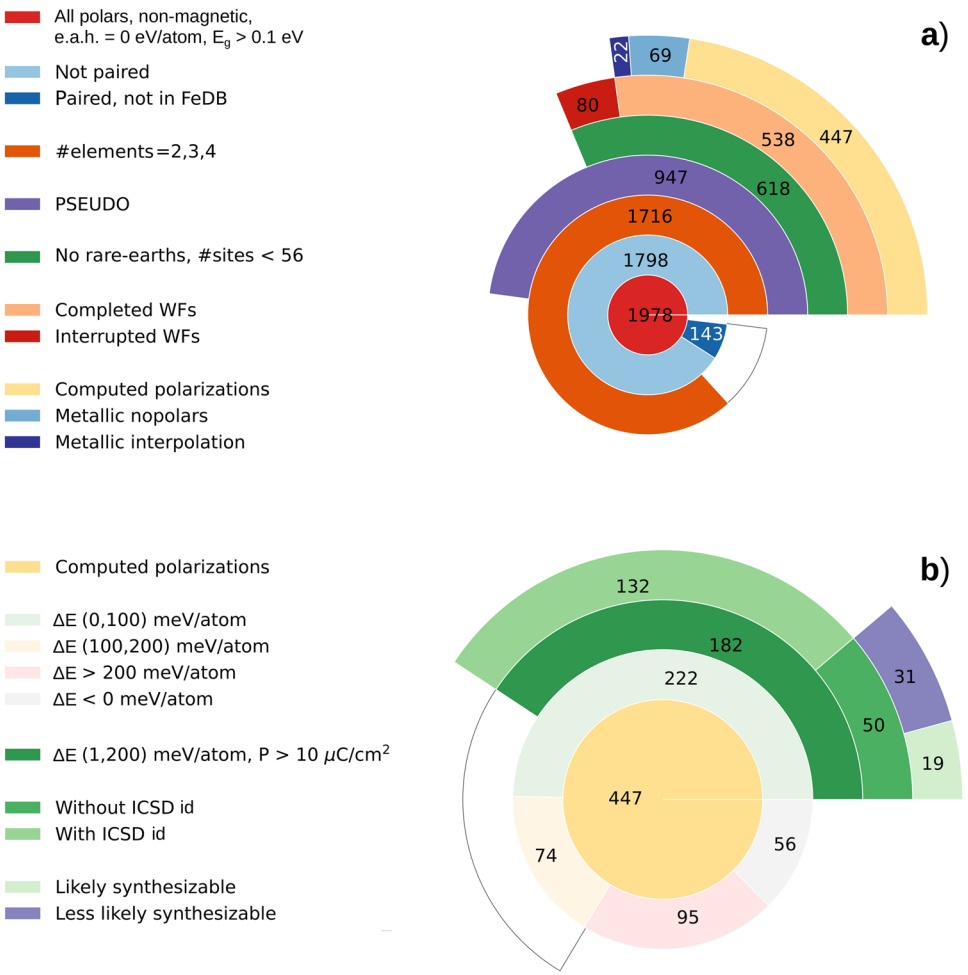

Fig. 2 Summary of the screening process and number of structures in output at each step. a Subsets of candidate ferroelectric materials found during our screening and their size. Starting from the initial entries in the MP, filters on the number and type of elements are applied; PSEUDO is used to generate the nonpolar reference; and an automated workflow (WF) is used to compute the polarization and the polar-nonpolar phase energy difference. **b** Subsequent classification based on the polar-nonpolar phase energy difference, polarization, ICSD ids, and synthesizability (as computed following prior work[42]). In the charts, the abbreviations e.a.h., E_g, FeDB, #elements, no rare-earths, #sites, P, and ΔE stand for the energy above hull, the electronic band gap, the database of ferroelectrics in ref. [35], the number of elements, rare-earth elements are excluded, the number of elements, the polarization, and the polar-nonpolar energy difference, respectively. Paired and not paired refer to polar-nonpolar pairs present or not in the FeDB.

simultaneously polar and magnetic, which require a careful assessment of the magnetic ground state, are not considered here and will be the subject of future work.

Main classification

We summarize the 447 compounds that emerge from our workflow in terms of their computed polarization P and the energy difference between polar and nonpolar phases ΔE, as shown in Fig. 3. The computed values of both quantities cover a large range. In particular, 50% of materials have ΔE in the range 0–100 meV per atom, and 70% possess P less than 50 μC cm^{-2}. Based on previously reported values of these quantities for known ferroelectric and polar materials, such as $BaTiO_3$, $PbTiO_3$, $LiNbO_3$, $BiFeO_3$, $YMnO_3$, HfO_2[35,45], $GaFeO_3$[46], AlN, $Al_{0.5}Sc_{0.5}N$[47], ZnO[48], and GaN (present work) computed with same exchange-correlation functional (except for $YMnO_3$ where LDA functional was used) and nonpolar reference as those used here, we classify our 447 candidate materials in three tiers based on ΔE (Fig. 1b), indicated with different colors in Fig. 3.

The first tier contains 222 materials with ΔE in the range 0–100 meV per atom (green region), similar to values reported for known ferroelectrics. The second tier contains 74 materials with ΔE in the range 100–200 meV per atom (orange region), similar to polar materials such as AlN. The pristine version of AlN is a polar material with a high P (~120 μC cm^{-2}), but not a ferroelectric, as its P cannot be switched by an electric field. However, as found both theoretically[47,49,50] and experimentally[50], doping AlN with Sc atoms preserves the P, reduces the ΔE, and makes switching possible. $Al_{x-1}Sc_xN$ has been synthesized[50,51], and it is promising for applications[52]. Proposed compounds with ΔE in the same range may exhibit switchability via doping or other alternative methods (e.g., strain, thin films). The third tier contains the remaining 95 compounds with ΔE above 200 meV per atom (red region). This may be an upper limit for conventional ferroelectricity; $GaFeO_3$ possesses the largest predicted ΔE that has also been experimentally demonstrated to be ferroelectric[46], to our knowledge. These polar compounds, like GaN, could also be interesting for other properties (e.g., dielectric, piezoelectric, photovoltaic, nonlinear optical), but may not be bulk ferroelectrics, at least in pristine crystalline form.

Subset of potential ferroelectrics

We now focus on a subset of the 182 identified candidate ferroelectric materials in tiers 1 and 2 with a computed ΔE

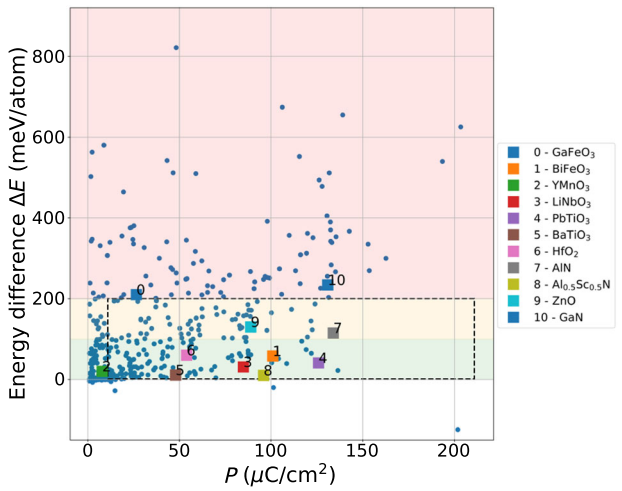

Fig. 3 Main classification of the potential ferroelectrics in the dataset. Computed polarization P (μC cm^{-2}) and the polar-nonpolar phase energy difference ΔE (meV per atom) of polar compounds in the present dataset. Three tiers are highlighted according to the energy difference: the green tier contains potential ferroelectrics that are similar to common ferroelectrics; the orange tier contains polar materials where the energy difference could be tuned, and ferroelectricity appears, as it happens for AlN with Sc doping; the red tier contains polar materials with an energy difference that may be too high to allow ferroelectricity. The dashed box encloses a subset of materials most relevant as ferroelectrics and considered in the ranking and literature search. The negative values of the energy difference represent cases where the nonpolar reference structure has lower energy than the polar phase. Values of the ΔE and P for $BaTiO_3$, $PbTiO_3$, $LiNbO_3$, $BiFeO_3$, $YMnO_3$, HfO_2[35,45], $GaFeO_3$[46], AlN, $Al_{0.5}Sc_{0.5}N$[47], ZnO[48], GaN (present work) are reported for comparison.

between 1 and 200 meV per atom and a predicted P larger than 10 μC cm^{-2}. This subset is highlighted by the dash line region in the Figures 3 and 2b.

We assess these candidate materials via three metrics. First, we rank promising ferroelectric materials identified here by combining their computed values of P and ΔE into a unitless effective score, F_{score}, defined as follows: $F_{score} = \overline{P} - 2 \cdot \Delta\overline{E}$, where $\overline{P}$ and $\Delta\overline{E}$ are P and ΔE normalized with respect to their respective maximum value. F_{score} is then normalized and scaled to take on values between zero and unity. F_{score} prioritizes candidates with high P and low ΔE. We note that our choice of ranking is not unique, and present data are publicly available; the materials computed here can be ranked differently depending on specific target properties.

Second, we use Scopus (https://www.scopus.com/home.uri) and Google Scholar (GS) databases to determine whether the compounds have been synthesized or studied previously by theory, as described in Methods. We count the total number of entries for each material's chemical formula in Scopus and GS databases, the sum of the number of abstracts containing relevant ferroelectric-related keywords, and the number of abstracts that refer to synthesis. These cumulative numbers indicate the extent of prior studies of the material both in general and in the context of ferroelectricity or piezoelectricity.

Third, we separate this subset of materials into two categories depending on whether they are in the Inorganic Crystal Structure Database (ICSD) database[53]. The first category contains the 132 compounds with an ICSD id, which we consider as increasing its likelihood to be synthesizable with conventional methods. The second category contains 50 materials which are not included in the ICSD database, indicating that some of them have not yet been synthesized. To assess the synthesizability of these 50 compounds, we adopt the machine learning-based approach presented in ref. [42], and briefly described in the Methods section, which provides an approximate measure of those materials more likely to be synthesizable. Using this model 19 compounds are predicted to be likely synthesizable and 31 compounds less likely synthesizable. This subsets are visually represented in Fig. 2b.

The 30 compounds with an ICSD id with the highest F_{score} are reported in Table 1. The reminder of materials in this category is provided in Supplementary Table 1. In Table 2 the same ranking is reported for compounds absent from the ICSD database but predicted likely synthesizable. Those predicted less-likely synthesizable are reported in Supplementary Table 2. In these tables, the LP column indicates the presence in the literature of each formula; the R column indicates the number of abstracts containing at least one of the keywords related to ferroelectricity; the S column reports the number of abstracts referring to a synthesis method. The details on how these three quantities are determined can be found in the Methods section.

Tables 1 and 2, together with Supplementary Tables 1 and 2 summarize promising ferroelectrics, and they provide information on which of these materials have already been synthesized. A higher number of entries in Scopus or GS databases (LP column) suggests that these compounds are relatively more prominent in the literature, and those materials that have a higher number of entries reporting synthesis (S column) are more likely to be readily synthesized. Cases where LP is 'Medium' and 'High' and R is equal to zero suggest that either they have been studied for applications other than ferroelectrics, or the investigated phase is different from the polar one. Overall, we find 53 compounds that are, for the most part, previously unknown, i.e., they have a total number of entries in the two databases lower than 10 ('Low' LP). Considering the results from Scopus, 51 have entries that can be related to ferroelectricity and piezoelectricity, 80 are known in the literature but have been studied for different properties and applications, and 119 have been featured in at least one paper mentioning synthesis methods. Therefore, our literature search suggests that ~130 materials are potential ferroelectrics which may be targeted for further theoretical and experimental effort.

Our literature search finds that the main applications explored for these materials include photovoltaic, photocatalytic, and nonlinear optical properties such as second harmonic generation (SHG); electrodes and electrolytes for batteries; and thermoelectricity. The study of polar and ferroelectric materials for photovoltaic applications is motivated by the generation of photocurrent in these materials[54]. More generally, nonlinear optical effects constitute an important application area[55], and $LiNbO_3$, for example, shows strong SHG and is also a ferroelectric[56,57].

In the following, we discuss some of the materials present in Tables 1 and 2 and recently reported in the literature. In Table 1, among those with an ICSD id and reported as relevant, we recognize $LiTaO_3$, GeTe, Bi_2WO_6, and $SrNb_2Bi_2O_9$ that have already been extensively studied as ferroelectric or piezoelectric materials; $Sr_2Nb_2O_7$, $SnPS_3$, $LiGaO_2$, Ga_2S_3, $SnPSe_3$, $KNbSi_2O_7$, B_4PbO_7, and Li_2GeO_3 are also known ferroelectrics but less extensively studied. In Ref. [58], the dielectric and piezoelectric properties are reported of vanadium-alloyed $Sr_2Nb_2O_7$, but the potential for ferroelectricity in this system has, to our knowledge, yet to be assessed. Jia et al.[59] performed a computational study of the NbX_2O family, finding materials with intrinsic ferroelectricity and antiferroelectricity, and proposed these materials as potential 2D ferroelectrics. Predicted value of P for NbX_2O (where X = Cl, I, Br) compounds in our database are in agreement with ref. [59]. Some of our A_2B_3 candidates, and in particular some in the III_2-VI_3 family, appear to be promising from prior theoretical and experimental studies[60,61] as 2D ferroelectrics[62–66]. Further, $CsGeX_3$ (with X = Cl, Br, I) compounds are part of another family recently synthesized where ferroelectricity is reported[67]. Our computed values of P, for these

Table 1. Top 30 polar compounds with experimentally known structures.

mp-id	Formula	ICSD id	E_g	P	ΔE	SG_p	SG_{np}	F_{score}	LP	R	S
mp-557441	KAu(IO3)4	417267	2.1	135.5	22	P1	P-1	1.0	Low	0	0
mp-28027	AsF3	35132	5.4	69.0	4	Pna2_1	Pnma	0.89	High	0	8
mp-23480	Bi2WO6	415669	1.9	53.4	9	Pca2_1	Pbca	0.82	High	58	570
mp-3666	LiTaO3	239373	3.8	57.8	13	R3c	R-3c	0.82	High	223	205
mp-938	GeTe	159907	0.8	66.7	23	R3m	Fm-3m	0.81	High	170	175
mp-23614	SrNb2Bi2O9	95919	2.7	57.7	15	Cmc2_1	Cmcm	0.81	Medium	40	56
mp-11674	Al2Se3	14373	2.2	54.7	17	Cc	C2/c	0.8	High	0	5
mp-1029378	Mg2SbN3	131077	1.4	103.1	52	Cmc2_1	Cmcm	0.8	Low	1	1
mp-1020106	LiB(SO4)2	425174	6.4	57.1	20	Pc	P2_1/c	0.79	Low	0	0
mp-1198721	Rb2Mo3SeO12	167530	2.5	68.4	32	P6_3	C222_1	0.78	Low	0	0
mp-539	Ga2S3	488	2.0	60.9	25	Cc	C2/c	0.78	High	6	78
mp-3870	Sr2Nb2O7	281135	2.8	38.9	9	Cmc2_1	Cmcm	0.78	High	23	34
mp-13923	SnPS3	25357	2.2	32.0	11	Pc	P2_1/c	0.76	High	19	14
mp-550070	NbBr2O	416669	0.9	23.4	5	C2	C2/m	0.76	Low	0	0
mp-27384	AgIO3	14100	2.5	19.1	2	Pca2_1	Pbca	0.75	High	0	18
mp-23064	Bi2MoO6	47139	2.1	78.0	46	Pca2_1	Pbca	0.75	High	22	312
mp-570370	SnPSe3	656173	1.4	24.7	8	Pc	P2_1/c	0.75	Medium	4	3
mp-9747	B4PbO7	203201	4.6	118.2	74	Pmn2_1	Pmmn	0.75	Low	1	4
mp-549720	NbI2O	36255	0.8	17.5	1	C2	C2/m	0.75	Low	0	0
mp-1190450	NaNbSe2O7	239096	2.9	14.0	3	Cmc2_1	Cmcm	0.74	Low	0	1
mp-558063	KCu(BiS2)2	91297	0.9	14.3	2	Cmc2_1	Cmcm	0.74	Low	0	0
mp-545359	KNa2CuO2	47105	1.6	22.8	8	I4mm	I4/mmm	0.74	Low	0	0
mp-5854	LiGaO2	93087	3.2	86.8	57	Pna2_1	Pnma	0.73	High	7	29
mp-1205304	TiI2O7	424632	2.6	12.9	3	Cmc2_1	Cmcm	0.73	Low	0	0
mp-560407	KNbSi2O7	72111	3.4	22.5	14	P4bm	P4/mbm	0.72	Medium	4	5
mp-6765	Li3AlGeO5	72098	4.0	82.6	58	Pna2_1	Pnma	0.72	Low	0	0
mp-18115	K2Sn2Hg3S8	281505	1.4	45.7	29	Aea2	Cmce	0.72	Low	0	0
mp-6057	TlSnPS4	68294	1.8	15.1	8	Pna2_1	Pnma	0.72	Low	0	0
mp-13979	Ca2Nb2O7	26010	3.0	22.6	15	P2_1	P2_1/c	0.72	High	4	12
mp-1068340	CsGeBr3	80318	2.2	20.0	14	R3m	Pm-3m	0.71	High	0	1

Compounds are ranked according to computed polarization P ($\mu C\,cm^{-2}$) and nonpolar-polar phase energy difference ΔE (meV per atom), as combined in the F_{score}. The mp-id, the formula, the band gap E_g (eV), the ICSD ID, and the space group (SG) label of the polar and nonpolar phases are also provided. The LP column indicates the presence of each formula in the literature from the automatic screening of Scopus and GS databases. Low, medium, and high mean number of entries in the range [0,10], (10,100), and ≥100, respectively. The R and S columns report the number of abstracts in Scopus containing at least one of the ferroelectric-related keywords and referring to a synthesis method, respectively.

compounds, are very close to those previously measured and computed.

Of the remaining materials with an ICSD id reported in Supplementary Table 1, we mention SrAlGeH, a member of the hydrogenated Zintl family AeTrTtH (Ae = Ca, Sr, and Ba; Tr = Al and Ga; Tt = Si, Ge, and Sn); some of these compounds have been synthesized and studied for their potential bulk photovoltaic effects and ferroelectricity[68]. In addition, $NaGaO_2$ has been reportedly synthesized[69,70] but has not yet been explored as ferroelectric. Interestingly, $CuGaO_2$ has been identified as promising ferroelectric[71] and was synthesized recently,[72], and its P measured[73]. This material is not in our screening because its energy above hull is ~0.1 eV per atom according to the MP. Furthermore, multiferroic $NaFeO_2$ has been reported as ferroelectric and a weak ferromagnet[74]. Finally, we mention $Ca_4YB_3O_{10}$, which has been investigated as a piezoelectric material[75] and has been synthesized[76]. This systems may not be a ferroelectric in its pristine form due to its high ΔE (low F_{score}).

Among the structures without an ICSD number predicted to be synthesizable in Table 2, $LiAlS_2$ is the only one reported as relevant. $LiAlS_2$ has been recently reported as part of a study of a larger family of $LiMX_2$ compounds computationally predicted as promising monolayer piezoelectric compounds[77]. Furthermore, its relative $LiAlTe_2$ has been studied as ferroelectric in both bulk and 2D layered forms, and both P and ΔE reported values are similar to our calculated values[78]. Many materials reported here are quaternaries and hence relatively less explored, especially as ferroelectrics, according to our literature search. Also, their ΔE values are relatively high, suggesting that while they may not be ferroelectric in the bulk, they may be tuned to enable ferroelectricity.

In Supplementary Table 2, we also tabulate candidate ferroelectrics predicted to be less-likely synthesizable. Among the materials with higher F_{score}, we identify some nitrides with stoichiometry XYN_2 and XYN_3, with large P and large ΔE values. These materials were also predicted in a prior high-throughput search[79] and recently added to the MP database. Some of these nitrides were reported in previous computational studies as promising materials for photovoltaic applications[80]. Notably, $TiZnN_2$ has been recently synthesized in its polar wurtzite-like phase[81], but its P has not been measured yet. This class of materials certainly deserves further investigation to assess if

Table 2. Theoretical polar structures with the synthesizability CL_{score} higher then 0.5, as defined in Methods.

mp-id	Formula	E_g	P	ΔE	SG_p	SG_{np}	F_{score}	LP	R	S
mp-775149	Ti(SO4)2	2.2	107.6	39	R3	R-3	0.86	High	0	19
mp-1025567	NbCl2O	0.9	26.9	5	C2	C2/m	0.76	Low	0	1
mp-1219595	RbHgSbTe3	0.6	14.6	3	Cmc2_1	Cmcm	0.74	Low	0	0
mp-1220802	NaHSeO4	3.4	14.9	9	P1	P-1	0.72	Medium	0	1
mp-1227524	BaTi(IO3)6	2.7	10.4	9	R3	R-3	0.7	Medium	0	0
mp-1224198	K4Ba2SnBi4	0.1	16.8	14	Cm	C2/m	0.7	Low	0	0
mp-1210797	Li2MgGeO4	3.8	86.6	65	Pmn2_1	Pmmn	0.7	Medium	0	0
mp-979137	SrGaSiH	0.4	39.6	51	P3m1	P-6m2	0.62	Low	0	0
mp-1213763	ClF5	3.1	31.2	62	Cmc2_1	Cmcm	0.56	High	0	2
mp-1224100	K4Ba2SnSb4	0.4	10.1	74	Cc	C2/c	0.45	Low	0	0
mp-1106183	LiAlS2	4.1	60.3	110	Pna2_1	Pnma	0.45	Medium	1	0
mp-34579	Na3HfF7	5.6	26.0	90	Cm	C2/m	0.44	Medium	0	0
mp-1212662	GaClO	2.7	50.0	113	Pca2_1	Pbcm	0.41	High	0	0
mp-1214215	Ca11AlSb9	0.9	17.6	103	Iba2	Ibam	0.36	Low	0	0
mp-1224063	K4Ba2SnAs4	0.5	12.3	99	Cc	C2/c	0.36	Low	0	0
mp-1223469	KLiZn3O4	1.5	58.2	152	Cm	P-1	0.29	Low	0	0
mp-1210804	Li2ZnGeS4	2.2	61.6	155	Pmn2_1	Pmmn	0.28	Medium	0	0
mp-1214910	Al2CdBr8	2.9	32.4	151	Pc	P2/c	0.22	Low	0	0
mp-1213982	Cd(GaCl4)2	3.8	36.2	180	Pc	P2/c	0.12	Low	0	0

These polar structures on the hull have yet to be reported experimentally. The ranking is performed according to computed polarization P (μC cm^{-2}) and nonpolar-polar phase energy difference ΔE, as combined in the F_{score}. The mp-id, the formula, the band gap E_g (eV), and the space group (SG) label of the polar and nopolar phases are also provided. The LP column indicates the presence of each formula in the literature from the automatic screening of Scopus and GS databases. Low, medium, and high mean number of entries in the range [0,10], (10,100), and ≥100, respectively. The R and S columns report the number of abstracts in Scopus containing at least one of the ferroelectric-related keywords and referring to a synthesis method, respectively.

ferroelectricity is possible. In general, although most of the materials in Supplementary Table 2 appear to be appealing as ferroelectric candidates, they may be somewhat more challenging to synthesize by conventional methods.

Overall, our large-scale high-throughput screening offers the opportunity to identify already known ferroelectrics and to highlight new potential ferroelectrics among materials that have been less investigated or overlooked. Materials present in our dataset, such as $CsGeX_3$, were very recently synthesized and reported to be ferroelectrics; NbX_2O and A_2B_3 have been the subject of recent theoretical investigations for ferroelectric properties, with computed P in agreement with values reported here, providing validation to our approach. Future studies building on our work could take multiple approaches. One could select candidates from our study that have been already reported in the literature but not studied as ferroelectrics. In this case, papers reporting synthesis techniques could be helpful as a starting point for realizing these materials experimentally. Or, one could select those that are relatively less known, in an attempt to identify a completely new ferroelectric material. We hope our database and workflow will inspire several such studies.

Dynamical stability of polar phases

The thermodynamic stability of the polar phases studied here is imposed during our screening by the use of the stringent criterion of working with materials with only zero energy above hull. At zero temperature, dynamical stability is conventionally evaluated by computing the phonon spectrum, and inspecting for imaginary frequencies. We make use of the phonon band structures provided by Togo et al.[82,83] (phonondb, available at https://mdr.nims.go.jp/collections/8g84ms862), obtained via finite differences approach and within DFT, to roughly access the dynamical stability of the polar materials in our subset of potential ferroelectric compounds. Among the 182 materials in this group, the phonondb database contains the phonon band structure for a subset of 94 compounds. Inspection of the phonondb suggests that out of these 94 cases, 23 are susceptible to instability at 0K due to the presence of imaginary frequencies at **q**-points beyond the Γ point, principally zone boundary instabilities. Notably, of the 23 compounds with potential 0K instabilities, 22 are associated with an ICSD id, suggesting plausible stability at higher temperatures. Hence, while assessing phonons at 0K serves as a reliable metric for ascertaining dynamical stability, they are not necessarily indicative of a material's stability and synthesizability at different temperatures, pressures, strain, etc. For example, the tetragonal phase of $BaTiO_3$, which is stable and ferroelectric at room temperature[2,84,85], exhibits imaginary frequencies both at zone center and zone boundary points in its computed 0K phonon band structure (https://next-gen.materialsproject.org/materials/mp-5986).

Role of the nonpolar structure and energy profiles

In this section, we comment on the role of the nonpolar phases in the context of defining a material as ferroelectric. Using the modern theory of polarization[8], the polarization of a polar material is defined as the difference between the polarization in a polar and nonpolar phase along an adiabatic path. Therefore, a reference nonpolar structure is often required to have an accurate value of polarization for comparison with the experiment. In some cases, e.g. in perovskites, the nonpolar phase can have a physical meaning as a high-temperature phase, and the difference in energy between polar and nonpolar phases provides a measure of the stability of the polar phase with temperature, as well as being a measure of the ability to switch the polarization direction via an electric field (Fig. 1b).

While the nonpolar structure that we obtain using pseudosymmetries is valid as a reference to compute the polarization, because it is on an adiabatic path, there is no guarantee that this

nonpolar structure is relevant to switching under an applied electric field. For instance, the polarization could rotate under applied field, instead of maintaining its direction and reducing its intensity, with the rotated phase still possessing a polar space group[86]. There also may be another nonpolar phase, one not related to the polar one by group-subgroup relationships, that could compete in energy and manifests during the switching. At microscopic level, switching is generally accepted to occur in many ferroelectrics by domain wall motion[87]. Overall, this means that the intermediate phases occurring during switching, their energetic, and the corresponding electric field required for switching, could differ from the ones we report here. The nonpolar structures used in our study, while strictly a means for computing the polarizations and energy difference, are closest to the polar phase in terms of atomic displacements, and, thereby, the most relevant and simplest for assessing the stability of the polar material and its potential ferroelectricity. Therefore, we use it to classify the materials in three tiers, separating those cases that are likely to be ferroelectric from others which are most likely not, despite the caveats mentioned above. Finally, we note that a quantitative estimate for the coercive field could be an alternative descriptor for switchability, but its accurate prediction, which depends on the microscopic nature of the switching pathway, is currently beyond the scope of high-throughput ab initio investigations.

We now comment on the energy profiles obtained by our interpolation method. The majority of these energy profiles show a classic double-well shape (Fig. 1b). However, in 35 cases the energy profile deviates from this standard shape and indicates that one of the intermediate interpolated structures has an energy higher than the nonpolar reference. This would suggest that the relevant energy barrier could be actually higher than the simple energy difference between the polar and nonpolar phase. However, the linear interpolation that we perform to generate the intermediate structures is fictitious, and the actual structures could find different ways to rearrange the atoms to minimize the energy. In the context of two-dimensional ferroelectric materials[62], the nudged elastic band approach[88] is used to recover the usual double-well profile in such a case. The energy profiles generated automatically from linear interpolation of the endpoint structures do not always provide reliable information, and the actual path can require more advanced methods to be determined accurately. Also, in some case, the atomic distortions along this fictitious path can lead to metallic structures. This was reported in ref. [35] and can be linked to our 22 cases (Fig. 2a) where a metallic interpolated structure is found.

Other sets of materials

We note that 75 compounds have an energy difference close to zero or negative (<1 meV per atom). In these compounds the nonpolar phase we create is actually more stable than, or competing with, the polar one. In general, these cases constitute a discovery of another phase, not yet reported in the MP database, that expands knowledge of the energy landscape around the known polar phase. Cases with energy differences of ~1 meV per atom should be treated with care and are at the accuracy limit of our calculations; moreover, the synthesis of these materials could likely lead to the competing or more stable nonpolar phase instead, as reported recently for $BiInO_3$ in ref. [37]

For 69 compounds, as reported in Supplementary Table 3, the polarization could not be computed due to the fact that the nonpolar structure turns out to be metallic (i.e., the computed DFT band gap is zero) within DFT-PBE. In these cases, there may be another nonpolar structure that is semiconducting, or the metallic nonpolar phase could be an artifact of the exchange-correlation functional we use. Although we do not investigate these cases further in the present work, we cannot rule these systems out as potential ferroelectric or piezoelectric materials. An example is $LaWN_3$, which has been recently synthesized in its polar phase[89] and shows piezoelectric properties; its nonpolar phase is known to be metallic[89,90]. We note that this specific compound is not among our current candidates because we exclude materials with rare-earth elements, but it underscores that the materials in this group should be considered for further investigation in future work.

METHODS

Pseudosymmetries

The present workflow implements a large-scale use of the BCS's PSEUDO tool via an automated interface to generate a nonpolar parent structure from an initial polar structure that is used as a reference to compute the polarization. PSEUDO is a tool that can be used via a web interface on the BCS website. This method requires an input structure in the conventional cell setting and a tolerance value representing the maximum atomic displacement in Angstroms allowed for the pseudosymmetry search. A default value of 2 Å is used in our search. Starting with a polar structure, PSEUDO returns a list of possible parent structures of higher symmetry respecting a) the group-supergroup relationship, (b) the lattice distortion default values, (c) the fact that the maximum atomic displacement is lower than the chosen threshold, and (d) the consistency of Wyckoff splitting for each species[24]. Therefore, PSEUDO does not simply rely on decreasing the tolerance for the symmetry identification of the input structure, as discussed in previous works[33]. PSEUDO instead represents a more general, rigorous, and effective approach for detecting high symmetry parents of an initial structure. We note that the structures proposed by PSEUDO have the same cell parameters of the input polar structure, and therefore, a relaxation is needed to access its total energy; we further note that their space groups are not necessarily only nonpolar. Then, a nonpolar structure has to be chosen, downloaded in cif format, and used for different types of investigations[24]. The whole procedure to obtain a nonpolar structure is typically done manually by the user through the PSEUDO web interface, which limits the use of this tool at a large scale. To overcome this limitation, we implement a python interface that can automatically interact with the PSEUDO website, retrieve the proposed structures as pymatgen[91] structure objects, and rank them by their maximum atomic displacement and space group. We note that the application of this workflow to polar phases of well-known ferroelectrics $BaTiO_3$ and $PbTiO_3$ leads to the cubic perovskite nonpolar reference structure. Other cases where the nonpolar phase found by our procedure matches the one reported in the literature are mentioned in the 'Subset of potential ferroelectrics' section. Also, we highlight that our code can be used to automatically build the complete graph of all possible parent structures starting from an input structure, which can be used to study competing phases and the energy landscape around a certain structure of interest (not necessarily polar).

Computational details

The polarization is computed via DFT and the Berry phase approach[8–14], as implemented in Vienna Ab initio Software Package (VASP) version 6.3[92–94]. We use the generalized gradient approximation (GGA) functional of Perdew, Burke, and Ernzerhof (PBE)[95] and the projected augmented wave (PAW) pseudopotentials[96]. The interpolation of the polarization branch is implemented in the workflow described in ref. [35] and in the Atomate code[36]. We refer to these papers for more details on the workflow and the DFT parameters, as both have been used here in the same way. We reiterate that the values of lattice parameters, total energy, and polarization depend on the choice of the exchange-correlation functional[45]; this must be kept in mind when

comparing these values for the materials reported in the present work with those taken from other works in the literature.

Polarization jumping detection

In ref. [35], cases such as CrO_3, where the polarization values jumped from one branch to another nearby, resulting in an incorrect prediction of the polarization, were marked with a warning. In the present workflow, we improve the detection of a polarization jumping branch, and for the problematic cases, their polarization is recomputed using double the number of interpolated structures. Briefly, 34 cases out of 447 are recomputed in this way and recovered. Only seven cases still present a change of branch even with a doubled number of interpolation points, although they are less interesting for ferroelectric applications due to their very high ΔE. More details of the improved detection method are included in the Supplementary Methods.

Literature search

The large-scale screening of the literature performed here has been implemented by searching two known databases, Scopus (https://www.scopus.com/home.uri) and GS. The first database is searched with the MatScholar web tool (https://matscholar.com/), which provides access, via a website and a dedicated API, to abstracts, titles, and authors of scientific papers published before 2018 and reported in Scopus. MatScholar uses an unsupervised machine learning technique based on natural language processing trained on the Scopus database of scientific paper abstracts[97,98]. This allows the automatic extraction of materials properties, applications, phase labels, synthesis and characterization methods, enabling the user to obtain information about materials from the literature. We searched this database for the formula both in the standard reduced format used in the MP and in alphabetical format. For each formula, we check how many entries contain relevant context-related tags (automatic extracted by the Matscholar algorithm) such as 'ferroelectric', 'spontaneous polarization', 'piezoelectric', 'remnant polarization', 'synthesis_methods', to estimate if that formula has been previously associated to ferroelectricity and if a synthesis technique is reported. Furthermore, to include more updated contributions, we screen for formulas and relevant context-related text strings (equivalents to the above mentioned tags) in the abstracts present in Scopus since 2018 by a simpler text-matching method. An extract of the output of this screening is shown in Supplementary Tables 4 and 5 for the top most known formulas with a number of abstracts containing that formula higher than 10. The sum of entries containing a ferroelectric-related or 'synthesis_method' keyword, in both Scopus datasets (pre- and post-2018), are collected in the R and S columns, respectively, of Tables 1 and 2 and Supplementary Tables 1 and 2. Further, since most of the materials are not found in the Scopus database, we search through the GS database, which, in principle, allows the screening of the whole literature, for the chemical formula, only in the standard reduced format used in the MP, and record the total number of entries. An extract of the output of this screening is shown in the Supplementary Table 6 for the top 30 most known formulas. Finally, to further understand which properties have been studied and which applications were proposed for these materials, we also searched each formula on GS, looked at the title and text snippet associated with each entry of the first page, and manually extracted the main material properties, applications, and the research topics.

Our approach presents some potential pitfalls, and, in particular, the number of false positives can be high. This can be due to the fact that (a) a keyword being in the abstract does not guarantee that it actually refers to the considered material, and (b) the phase of the material studied in the paper can be different from the polar one that we consider here. However, we consider the false positives a minor issue. In fact, on the one hand, those cases where many abstracts are found, will certainly contain false positives, but it is very unlikely that they are all false positives. Assuming a fixed percentage of false positives (as a systematic error), we can confidently assume that these materials are the most reported in the literature among all. On the other hand, those cases with few entries can still contain false positives, which can be easily checked manually, but the fact that these are not known in the literature is captured. Lastly, we could be missing papers where the chemical formula is not expressed in a conventional way. An example is $Ca_4YB_3O_{10}$, which is known in the literature with the formula $YCa_4O(BO_3)_3$. In order to take into account these limitations, we summarize the overall presence of the materials in the literature by summing the number of entries per each formula in the three different databases and by grouping the number of entries in three ranges, namely [0,10], [10,100], >100, and we refer to them with the labels 'Low', 'Medium', 'High', respectively.

Synthesizability score

The synthesizability of predicted compounds, i.e., those without an entry in the ICSD databases, has been assessed by a machine learning technique that has been recently used to predict the synthesizability of perovskites[99] and MAX phases[100]. In particular, it has also been applied to the entire MP database to provide a score of synthesizability for all the materials[42], which is particularly useful for those structures that are only theoretical structures. This technique leverages the fact that the majority of the materials in the MP database have an experimental crystal structure reported in the ICSD database of experimental crystal structures[53]. This can be exploited by ML to predict if a crystal structure is synthesizable measuring its difference with those structures that are known experimentally and returning a probability score called crystal-likeness score (CL_{score}). This unitless score ranges between 0 and 1, with 0.5 taken as a threshold to distinguish likely synthesizable materials (above 0.5) from those less likely (below 0.5). The model has been trained on the MP structures with an ICSD id and shows an accuracy higher than 80%[42]. Then, it has been applied to predict the synthesizability of the theoretical structures in the MP database. The score for all the structures of the MP database is publicly available, and we use it to assess the synthesizability of the theoretical structures in the present dataset.

CODE AVAILABILITY

Code for workflow generation and supporting analysis is available in the atomate code repository (atomate.org) and in the pymatgen code repository (pymatgen.org), respectively.

DATA AVAILABILITY

The data of this study are available in the manuscript and in its Supplementary Information in the form of Tables. The data will also be available on the Materials Project website (materialsproject.org) and through its API.

Received: 4 October 2023; Accepted: 18 December 2023;
Published online: 20 January 2024

REFERENCES

1. Rabe, K. M., Ahn, C. H. & Triscone, J.-M. (eds.) Physics of ferroelectrics: a modern perspective. *In*: Topics in applied physics, Vol. 105 (Springer, 2007).
2. Rabe, K. M. & Ghosez, P. First-principles studies of ferroelectric oxides. *In*: Physics of ferroelectrics: a modern perspective, topics in applied physics, 117–174 (Springer, 2007).
3. Khomskii, D. Classifying multiferroics: mechanisms and effects. *Physics* **2**, 20 (2009).

4. Spaldin, N. A. Advances in magnetoelectric multiferroics. *Nat. Mater.* **18**, 10 (2019).
5. Wang, J. et al. Epitaxial BiFeO3 multiferroic thin film heterostructures. *Science* **299**, 1719–1722 (2003).
6. Heron, J. T. et al. Deterministic switching of ferromagnetism at room temperature using an electric field. *Nature* **516**, 370–373 (2014).
7. Scott, J. F. Room-temperature multiferroic magnetoelectrics. *NPG Asia Mater.* **5**, e72–e72 (2013).
8. Resta, R. & Vanderbilt, D. Theory of polarization: a modern approach. *In:* Physics of ferroelectrics: a modern perspective, topics in applied physics, 31–68 (Springer, 2007).
9. Resta, R. Theory of the electric polarization in crystals. *Ferroelectrics* **136**, 51–55 (1992).
10. Resta, R. Macroscopic polarization in crystalline dielectrics: the geometric phase approach. *Rev. Mod. Phys.* **66**, 899–915 (1994).
11. King-Smith, R. D. & Vanderbilt, D. Theory of polarization of crystalline solids. *Phys. Rev. B* **47**, 1651–1654 (1993).
12. Vanderbilt, D. & King-Smith, R. D. Electric polarization as a bulk quantity and its relation to surface charge. *Phys. Rev. B* **48**, 4442–4455 (1993).
13. Spaldin, N. A. A beginner's guide to the modern theory of polarization. *J. Solid State Chem.* **195**, 2–10 (2012).
14. Vanderbilt, D. Berry phases in electronic structure theory: electric polarization, orbital magnetization and topological insulators (Cambridge University Press, 2018).
15. Ghosez, P., Michenaud, J.-P. & Gonze, X. Dynamical atomic charges: the case of ABO3 compounds. *Phys. Rev. B* **58**, 6224–6240 (1998).
16. Junquera, J. & Ghosez, P. Critical thickness for ferroelectricity in perovskite ultrathin films. *Nature* **422**, 506–509 (2003).
17. Abrahams, S. C. Systematic prediction of new inorganic ferroelectrics in point group 4. *Acta Crystallogr. Sect. B Struct. Sci.* **55**, 494–506 (1999).
18. Abrahams, S. C. Structurally based prediction of ferroelectricity in inorganic materials with point group 6mm. *Acta Crystallogr. Sect. B: Struct. Sci.* **44**, 585–595 (1988).
19. Abrahams, S. C. Structurally based predictions of ferroelectricity in seven inorganic materials with space group Pba2 and two experimental confirmations. *Acta Crystallogr. Sect. B: Struct. Sci.* **45**, 228–232 (1989).
20. Cha, J.-W. & Kim, J.-N. The study of ferroelectricity and phase transition in Li2B4O7single crystals. *Ferroelectrics* **197**, 93–96 (1997).
21. Foster, M. C., Nielson, R. M. & Abrahams, S. C. Sr2SbMnO6: a new semiconductor ferroelectric. *J. Appl. Phys.* **82**, 3076–3080 (1997).
22. Foster, M. C., Arbogast, D. J., Photinos, P., Nielson, R. M. & Abrahams, S. C. K2(NbO)2Si4O12: a new ferroelectric. *J. Appl. Crystallogr.* **32**, 421–425 (1999).
23. Christie, R. J., Wu, P. K., Photinos, P. & Abrahams, S. C. Phase transitions and ferroelectricity in $NaSb_3F_{10}$. *J. Appl. Crystallogr.* **42**, 58–62 (2009).
24. Capillas, C. et al. A new computer tool at the bilbao crystallographic server to detect and characterize pseudosymmetry. *Z. Für. Kristallogr.* **226**, 186–196 (2011).
25. Capillas, C., Aroyo, M. I. & Perez-Mato, J. M. Search for new $Pna2_1$ ferroelectrics. *Ferroelectrics* **301**, 203–206 (2004).
26. Kroumova, E., Aroyo, M. I., Pérez-Mato, J. M., Igartua, J. M. & Ivantchev, S. Systematic search of displacive ferroelectrics. *Ferroelectrics* **241**, 295–302 (2000).
27. Igartua, J. M., Aroyo, M. I., Kroumova, E. & Perez-Mato, J. M. Search for *Pnma* materials with high-temperature structural phase transitions. *Acta Crystallogr. Sect. B Struct. Sci.* **55**, 177–185 (1999).
28. Igartua, J. M., Aroyo, M. I. & Perez-Mato, J. M. Systematic search of materials with high-temperature structural phase transitions: application to space group $P\,2_12_12_1$. *Phys. Rev. B* **54**, 12744–12752 (1996).
29. Khan, A. C., Cook, A. S., Leginze, J. A. & Bennett, J. W. Developing new antiferroelectric and ferroelectric oxides and chalcogenides within the A2BX3 family. *J. Mater. Res.* **37**, 346–359 (2022).
30. Bennett, J. W. & Rabe, K. M. Integration of first-principles methods and crystallographic database searches for new ferroelectrics: Strategies and explorations. *J. Solid State Chem.* **195**, 21–31 (2012).
31. Bennett, J. W., Garrity, K. F., Rabe, K. M. & Vanderbilt, D. Hexagonal ABC semiconductors as ferroelectrics. *Phys. Rev. Lett.* **109**, 167602 (2012).
32. Bennett, J. W. Discovery and design of functional materials: integration of database searching and first principles calculations. *Phys. Procedia* **34**, 14–23 (2012).
33. Garrity, K. F. High-throughput first-principles search for new ferroelectrics. *Phys. Rev. B* **97**, 024115 (2018).
34. Markov, M. et al. Ferroelectricity and multiferroicity in anti–Ruddlesden–Popper structures. *Proc. Natl. Acad. Sci.* **118**, e2026020118 (2021).
35. Smidt, T. E., Mack, S. A., Reyes-Lillo, S. E., Jain, A. & Neaton, J. B. An automatically curated first-principles database of ferroelectrics. *Sci. Data* **7**, 72 (2020).
36. Mathew, K. et al. Atomate: a high-level interface to generate, execute, and analyze computational materials science workflows. *Comput. Mater. Sci.* **139**, 140–152 (2017).
37. Acharya, M. et al. Searching for new ferroelectric materials using high-throughput databases: an experimental perspective on $BiAlO_3$ and $BiInO_3$. *Chem. Mater.* **32**, 7274–7283 (2020).
38. Acharya, M. et al. Exploring the $Pb_{1-x}Sr_xHfO_3$ system and potential for high capacitive energy storage density and efficiency. *Adv. Mater.* **34**, 2105967 (2022).
39. López-Pérez, J. & Íñiguez, J. Ab Initio study of proper topological ferroelectricity in layered perovskite $La_2Ti_2O_7$. *Phys. Rev. B* **84**, 075121 (2011).
40. Chou, S.-C., Greiner, S., Magdysyuk, O. V., Dinnebier, R. E. & Schleid, T. Theoretical and experimental analysis of structural phase transitions for $ScF[SeO_3]$ and $YF[SeO_3]$. *Z. Für. Anorg. Allg. Chem.* **640**, 3203–3211 (2014).
41. Frey, R. et al. Accelerated search for new ferroelectric materials. *Phys. Rev. Res.* **5**, 023122 (2023).
42. Jang, J., Gu, G. H., Noh, J., Kim, J. & Jung, Y. Structure-based synthesizability prediction of crystals using partially supervised learning. *J. Am. Chem. Soc.* **142**, 18836–18843 (2020).
43. Bartel, C. J. Review of computational approaches to predict the thermodynamic stability of inorganic solids. *J. Mater. Sci.* **57**, 10475–10498 (2022).
44. Sun, W. et al. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).
45. Zhang, Y., Sun, J., Perdew, J. P. & Wu, X. Comparative first-principles studies of prototypical ferroelectric materials by LDA, GGA, and SCAN meta-GGA. *Phys. Rev. B* **96**, 035143 (2017).
46. Song, S. et al. Ferroelectric polarization switching with a remarkably high activation energy in orthorhombic $GaFeO_3$ thin films. *NPG Asia Mater.* **8**, e242–e242 (2016).
47. Ye, K. H., Han, G., Yeu, I. W., Hwang, C. S. & Choi, J.-H. Atomistic understanding of the ferroelectric properties of a wurtzite-structure $(AlN)_n/(ScN)_m$ superlattice. *Phys. Status Solidi (RRL) - Rapid Res. Lett.* **15**, 2100009 (2021).
48. Konishi, A. et al. Mechanism of polarization switching in wurtzite-structured zinc oxide thin films. *Appl. Phys. Lett.* **109**, 102903 (2016).
49. Zhang, S., Holec, D., Fu, W. Y., Humphreys, C. J. & Moram, M. A. Tunable optoelectronic and ferroelectric properties in Sc-based III-nitrides. *J. Appl. Phys.* **114**, 133510 (2013).
50. Fichtner, S., Wolff, N., Lofink, F., Kienle, L. & Wagner, B. AlScN: A III-V semiconductor based ferroelectric. *J. Appl. Phys.* **125**, 114103 (2019).
51. Yasuoka, S. et al. Effects of deposition conditions on the ferroelectric properties of $(Al_{1-x}Sc_x)N$ thin films. *J. Appl. Phys.* **128**, 114103 (2020).
52. Mikolajick, T. et al. Next generation ferroelectric materials for semiconductor process integration and their applications. *J. Appl. Phys.* **129**, 100901 (2021).
53. Hellenbrandt, M. The inorganic crystal structure database (ICSD)—present and future. *Crystallogr. Rev.* **10**, 17–22 (2004).
54. Han, X., Ji, Y. & Yang, Y. Ferroelectric photovoltaic materials and devices. *Adv. Funct. Mater.* **32**, 2109625 (2022).
55. Boyd, R. W. Chapter 1 - The nonlinear optical susceptibility. In: Boyd, R. W. (ed.) *Nonlinear Optics (Third Edition)*, 1–67 (Academic Press, 2008).
56. Fiebig, M., Pavlov, V. V. & Pisarev, R. V. Second-harmonic generation as a tool for studying electronic and magnetic structures of crystals: review. *JOSA B* **22**, 96–118 (2005).
57. Volk, T. & Wöhlecke, M. Lithium niobate: defects, photorefraction and ferroelectric switching. Vol. 115, *Springer Series in Materials Science* (Springer, 2009).
58. Ling, H., Acharya, M., Martin, L. W. & Persson, K. A. Theory-guided exploration of the $Sr_2Nb_2O_7$ system for increased dielectric and piezoelectric properties and synthesis of vanadium-alloyed $Sr_2Nb_2O_7$. *Chem. Mater.* **34**, 8536–8543 (2022).
59. Jia, Y., Zhao, M., Gou, G., Zeng, X. C. & Li, J. Niobium oxide dihalides $NbOX_2$: A new family of two-dimensional van der Waals layered materials with intrinsic ferroelectricity and antiferroelectricity. *Nanoscale Horiz.* **4**, 1113–1123 (2019).
60. Ding, W. et al. Prediction of intrinsic two-dimensional ferroelectrics in In_2Se_3 and other III_2-VI_3 van der Waals materials. *Nat. Commun.* **8**, 14956 (2017).
61. Xue, W. et al. Discovery of robust ferroelectricity in 2D defective semiconductor α-Ga_2Se_3. *Small* **18**, 2105599 (2022).
62. Kruse, M. et al. Two-dimensional ferroelectrics from high throughput computational screening. *npj Comput. Mater.* **9**, 1–11 (2023).
63. Qi, L., Ruan, S. & Zeng, Y.-J. Review on recent developments in 2D ferroelectrics: theories and applications. *Adv. Mater.* **33**, 2005098 (2021).
64. Tsymbal, E. Y. Two-dimensional ferroelectricity by design. *Science* **372**, 1389–1390 (2021).
65. Wu, M. & Li, J. Sliding ferroelectricity in 2D van der Waals materials: related physics and future opportunities. *Proc. Natl Acad. Sci.* **118**, e2115703118 (2021).
66. Lipatov, A. et al. Direct observation of ferroelectricity in two-dimensional MoS_2. *npj 2D Mater. Appl.* **6**, 1–9 (2022).
67. Zhang, Y. et al. Ferroelectricity in a semiconducting all-inorganic halide perovskite. *Sci. Adv.* **8**, eabj5881 (2022).

68. A. Brehm, J. Predicted bulk photovoltaic effect in hydrogenated Zintl compounds. *J. Mater. Chem. C.* **6**, 1470–1475 (2018).
69. Suzuki, I., Kakinuma, A., Ueda, M. & Omata, T. Flux growth of β-$NaGaO_2$ single crystals. *J. Cryst. Growth* **504**, 26–30 (2018).
70. Suzuki, I., Suzuki, S., Watanabe, T., Kita, M. & Omata, T. Growth of β-$NaGaO_2$ thin films using ultrasonic spray pyrolysis. *J. Asian Ceram. Soc.* **10**, 520–529 (2022).
71. Song, S. et al. β-$CuGaO_2$ as a strong candidate material for efficient ferroelectric photovoltaics. *Chem. Mater.* **29**, 7596–7603 (2017).
72. Omata, T. et al. Wurtzite $CuGaO_2$: a new direct and narrow band gap oxide semiconductor applicable as a solar cell absorber. *J. Am. Chem. Soc.* **136**, 3378–3381 (2014).
73. Yao, M.-C. et al. β-$CuGaO_2$: A ferroelectric semiconductor with narrow band gap as degradation catalyst for wastewater environmental remediation. *Rare Met.* **41**, 972–981 (2022).
74. Viret, M. et al. β-$NaFeO_2$, a new room-temperature multiferroic material. *Mater. Res. Bull.* **47**, 2294–2298 (2012).
75. Kim, K., Zhang, S., Huang, W., Yu, F. & Jiang, X. $YCa_4O(BO_3)_3$ (YCOB) high temperature vibration sensor. *J. Appl. Phys.* **109**, 126103 (2011).
76. Ye, Q. & Chai, B. H. T. Crystal growth of $YCa_4O(BO_3)_3$ and its orientation. *J. Cryst. Growth* **197**, 228–235 (1999).
77. Lv, Q. et al. Large in-plane and out-of-plane piezoelectricity in 2D γ-$LiMX_2$ (M=Al, Ga and In; X=S, Se and Te) monolayers. *Mater. Sci. Semicond. Process.* **154**, 107222 (2023).
78. Liu, Z., Sun, Y., Singh, D. J. & Zhang, L. Switchable out-of-plane polarization in 2D $LiAlTe_2$. *Adv. Electron. Mater.* **5**, 1900089 (2019).
79. Sun, W. et al. A map of the inorganic ternary metal nitrides. *Nat. Mater.* **18**, 732–739 (2019).
80. Adamski, N. L., Wickramaratne, D. & de Walle, C. G. V. Band alignments and polarization properties of the Zn-IV-nitrides. *J. Mater. Chem. C.* **8**, 7890–7898 (2020).
81. Greenaway, A. L. et al. Zinc titanium nitride semiconductor toward durable photoelectrochemical applications. *J. Am. Chem. Soc.* **144**, 13673–13687 (2022).
82. Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).
83. Togo, A. First-principles phonon calculations with phonopy and Phono3py. *J. Phys. Soc. Jpn.* **92**, 012001 (2023).
84. Evans, H. T. An X-ray diffraction study of tetragonal barium titanate. *Acta Crystallogr.* **14**, 1019–1026 (1961).
85. Kwei, G. H., Lawson, A. C., Billinge, S. J. L. & Cheong, S. W. Structures of the ferroelectric phases of barium titanate. *J. Phys. Chem.* **97**, 2368–2377 (1993).
86. Fu, H. & Cohen, R. E. Polarization rotation mechanism for ultrahigh electromechanical response in single-crystal piezoelectrics. *Nature* **403**, 281–283 (2000).
87. Lines, M. E. & Glass, A. M. Principles and applications of ferroelectrics and related materials (Oxford University Press, 2001).
88. Henkelman, G. & Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **113**, 9978–9985 (2000).
89. Talley, K. R., Perkins, C. L., Diercks, D. R., Brennecka, G. L. & Zakutayev, A. Synthesis of $LaWN_3$ nitride perovskite with polar symmetry. *Science* **374**, 1488–1491 (2021).
90. Fang, Y.-W. et al. Lattice dynamics and ferroelectric properties of the nitride perovskite $LaWN_3$. *Phys. Rev. B* **95**, 014111 (2017).
91. Ong, S. P. et al. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
92. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
93. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
94. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
95. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
96. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
97. Weston, L. et al. Named entity recognition and normalization applied to large-scale information extraction from the materials science literature. *J. Chem. Inf. Model.* **59**, 3692–3702 (2019).
98. Tshitoyan, V. et al. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **571**, 95–98 (2019).
99. Gu, G. H., Jang, J., Noh, J., Walsh, A. & Jung, Y. Perovskite synthesizability using graph neural networks. *npj Comput. Mater.* **8**, 1–8 (2022).
100. Frey, N. C. et al. Prediction of synthesis of 2D metal carbides and nitrides (MXenes) and their precursors with positive and unlabeled machine learning. *ACS Nano* **13**, 3031–3041 (2019).

ACKNOWLEDGEMENTS

This work was primarily funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DE-AC02-05-CH11231 (Materials Project program KC23MP) for the development of functional materials. Portions of the analysis were funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award DE-SC0020383. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05-CH11231 using NERSC award BES-matgen. S.E.R.-L. acknowledges support from ANID Fondecyt Regular grant number 1220986.

AUTHOR CONTRIBUTIONS

S.A.M., S.E.R.-L., F.R. and J.B.N. conceived the study, and S.A.M. performed preliminary calculations. F.R. designed and implemented the workflow, performed the calculations, analyzed the data with input from S.E.R.-L. and J.B.N., and drafted the manuscript. S.E.R.-L. and J.B.N. worked with F.R. to revise and finalize the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-023-01193-3.

Correspondence and requests for materials should be addressed to Jeffrey B. Neaton.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.