npj computational materials

www.nature.com/npjcompumats

ARTICLE OPEN

Check for updates

Designing high-T_C superconductors with BCS-inspired screening, density functional theory, and deep-learning

Kamal Choudhary[1,2,3]✉ and Kevin Garrity[1,2]

We develop a multi-step workflow for the discovery of conventional superconductors, starting with a Bardeen–Cooper–Schrieffer inspired pre-screening of 1736 materials with high Debye temperature and electronic density of states. Next, we perform electron-phonon coupling calculations for 1058 of them to establish a large and systematic database of BCS superconducting properties. Using the McMillan-Allen-Dynes formula, we identify 105 dynamically stable materials with transition temperatures, $T_C \geq 5$ K. Additionally, we analyze trends in our dataset and individual materials including MoN, VC, VTe, KB_6, Ru_3NbC, V_3Pt, ScN, LaN_2, RuO_2, and TaC. We demonstrate that deep-learning(DL) models can predict superconductor properties faster than direct first-principles computations. Notably, we find that by predicting the Eliashberg function as an intermediate quantity, we can improve model performance versus a direct DL prediction of T_C. We apply the trained models on the crystallographic open database and pre-screen candidates for further DFT calculations.

npj Computational Materials (2022)8:244 ; https://doi.org/10.1038/s41524-022-00933-1

INTRODUCTION

Since the discovery of superconductivity in 1911 by Onnes[1], the identification of novel superconducting materials, especially those with high transition temperatures (T_C), has been an active area of research in condensed matter physics[2,3]. The highest temperature conventional superconductor in ambient conditions, MgB_2 ($T_C = 39$ K), was discovered relatively recently[4], and progress in high pressure hydrogen-based superconductors[5] and topological superconductors[6] further motivates the search for additional ambient or near-ambient condition superconductors with high T_C that may be suitable for scientific or industrial applications.

There have been several previous efforts to computationally and/or experimentally discover superconducting materials falling into certain materials classes, such as transition metals[7], A15, B1[8,9], AB_2 compounds[4,10,11], cuprates[12], iron-based compounds[13], hydrides[14–16], and many other material classes[2,4,17–24]. However, a more systematic data-driven search should help expedite the discovery of potentially high-T_C superconductors. Moreover, machine learning has become popular in the search for superconductors. There have been several reports of machine-learning applications for finding superconductors[18,25–27], but thus far they are mostly based on chemical formulas, and lack detailed atomic structure information that can be critical for superconducting behavior.

Two key ingredients required to computationally identify Bardeen–Cooper–Schrieffer (BCS) conventional superconductors[28,29] with high-T_C are: 1) a robust computational workflow, and 2) a database of curated materials with prior knowledge such as elastic constants and electronic density of states. Using density functional theory perturbation theory (DFT-PT), electron–phonon coupling (EPC) can be calculated and used to predict T_C with reasonable accuracy for arbitrary materials[29,30]. However, the computational expense of these calculations is very high when compared to a single DFT self-consistent energy calculation, especially when fully converging the relevant k-point and q-point Brillioun zone samplings for electrons and phonons. Hence, a fast and reliable computational workflow for identifying BCS type conventional superconductors that balances computational cost, accuracy, and scope is needed.

In this work, we develop such a computational approach to discover BCS superconductors, combining several methods at various levels of computational expense and accuracy. We start with a BCS-inspired pre-screening based for materials with high Debye temperature (θ_D) and high electron density of states (DOS) at Fermi-level ($N(0)$), using the existing JARVIS-DFT database[31]. We then develop and apply a DFT-PT workflow to compute T_C using electron–phonon coupling and the McMillan–Allen–Dynes formula[32], with initially low convergence settings. We benchmark the DFT workflow on known superconductors and apply it to materials from our prescreening step. For the best candidates, we perform additional convergence tests to validate our predictions. In addition, we use the dataset to develop deep-learning models using the atomistic line-graph graph neural network (ALIGNN)[33,34] to predict the Debye temperature, electronic DOS, T_C, and electron–phonon coupling parameters for arbitrary crystal structures.

We utilize the publicly available JARVIS[31] infrastructure to achieve the goals mentioned above. JARVIS (Joint Automated Repository for Various Integrated Simulations, https://jarvis.nist.gov/) is a collection of databases and tools to automate materials design using classical force-field, density functional theory, machine learning calculations and experiments. In particular, we obtain elastic tensor and DOS data from JARVIS-DFT database, establish the DFT workflow with JARVIS-Tools package and train the deep-learning model using ALIGNN. JARVIS-DFT is a density functional theory based database of 55,645 materials with several material properties such as formation energy, bandgap with different levels of theories[35], solar-cell efficiency[36], topological spin–orbit coupling spillage[37–39], elastic tensors[40], dielectric tensors, piezoelectric tensors, infrared and Raman spectrum[41], electric field gradients[42], exfolation energies[43], etc. with stringent DFT-convergence setup[44]. Including T_C in

[1]Material Measurement Laboratory, Gaithersburg, MD 20899, USA. [2]National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. [3]Theiss Research, La Jolla, CA 92037, USA. ✉email: kamal.choudhary@nist.gov

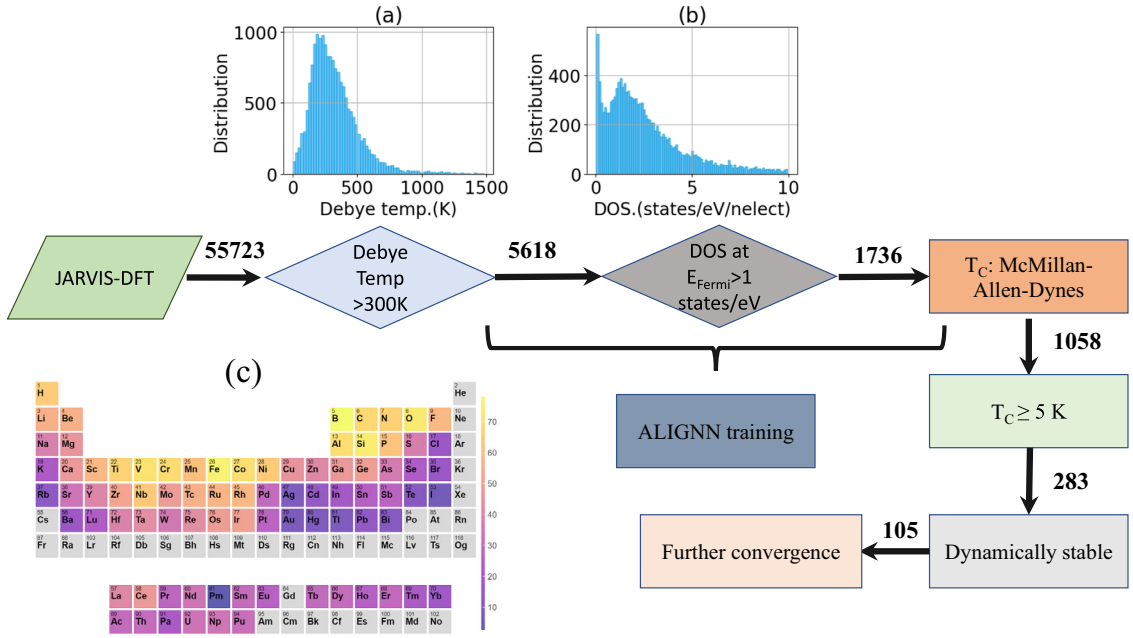

Fig. 1 Schematic showing the steps involved in identifying high-T_C superconductors. a statistical distribution of Debye temperature (K) and **b** statistical distribution of electronic density of states (states/eV/total number of electrons) at Fermi level from the JARVIS-DFT database, **c** probability that compounds containing a given element have $\theta_D > 300$ K. The flow chart shows the application of BCS-inspired screening, density functional theory calculations, and deep-learning training.

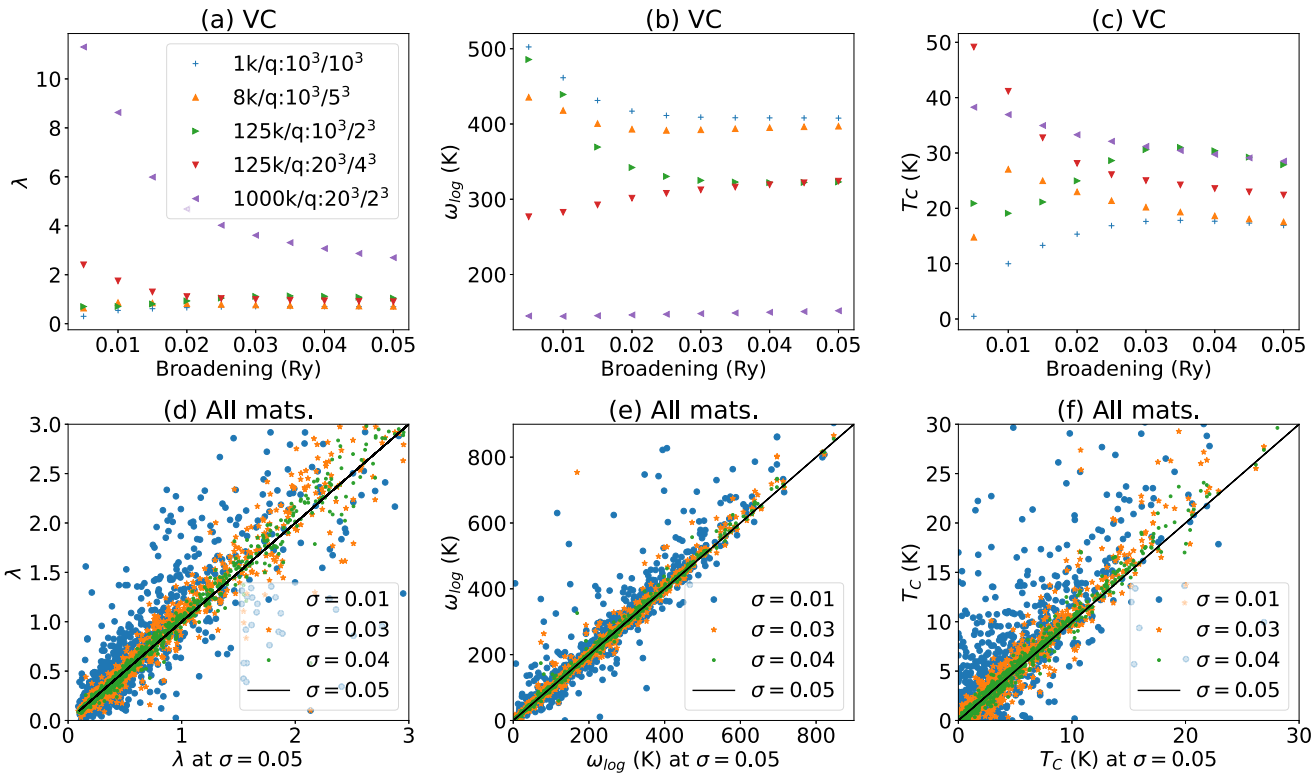

Fig. 2 Effect of different k-point and q-point selection on EPC parameters for VC. a λ, **b** ω_{log}, **c** T_C. Broadening dependent scaled **d** λ, **e** ω_{log} and **f** T_C for the all the systems in the database. The scaling was done with respect to the corresponding values at broadening of 0.05 Ry. The legend in **a** corresponds to the number of k-points and q-points, e.g., $10 \times 10 \times 10$ k-points and $10 \times 10 \times 10$ q-points (denoted by black star) is same as 1 k-point per q-point; and $20 \times 20 \times 20$ k-point and $2 \times 2 \times 2$ q-point correspond to 1000 k-point per q-point (denoted by left purple triangles).

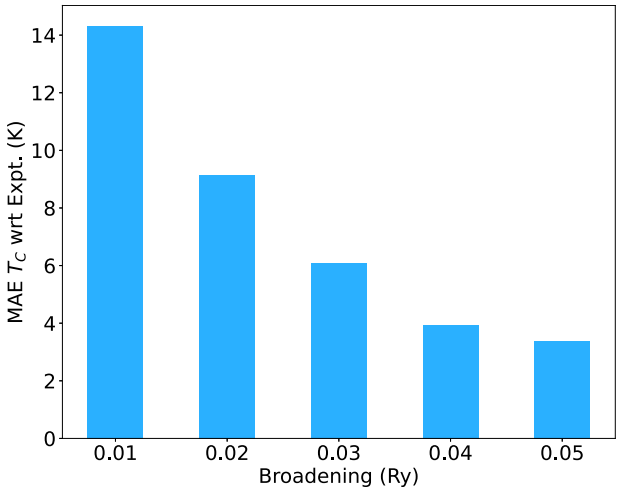

Fig. 3 Broadening convergence with respect to experimental T_C data for 14 superconductors.

JARVIS-DFT would greatly enrich the applicability of the database and aid in guided materials design. In this work, we add the T_C data in JARVIS-DFT.

RESULTS AND DISCUSSION

First principles workflow

A flow chart for designing conventional superconductors using BCS inspired screening (iBCS), JARVIS-Tools based DFT-Screening (J-Scr), and ALIGNN in an integrated way is shown in Fig. 1. The combination of computationally expensive DFT-based screening with empirical rules-based screening and deep learning methods allows us to apply our combined workflow to a wider group of materials than a brute-force computational approach.

First, we screen likely superconductors using a criteria inspired by the BCS equation for T_C (Eq. (1)). According to this equation, high-T_C materials usually have high Debye temperature (θ_D), and high electronic DOS (states/eV/total number of electrons) at the Fermi level ($N(0)$). The JARVIS-DFT database provides DFT-based electronic DOS at the Fermi level[35] and the Debye temperature, as derived from the elastic tensor database[40].

Currently, there are electronic DOS database available for 55,723 materials and elastic tensors for 17,419 materials (using the v08.18.2021 version of JARVIS-DFT database at the time of writing). We analyze the statistical distribution of θ_D and $N(0)$ in Fig. 1a, b, respectively. We observe that the Debye temperature can range up to 1500 K with high peak near 200 K (Fig. 1a). The DOS at the Fermi level can range up to 10 states/eV/Nelect with peak value around 1 (Fig. 1b). Out of 17,419 materials, we find 5618 of them with θ_D greater than 300 K. Furthermore, selecting materials with electronic DOS at the Fermi level greater than 1 states/eV/Nelect, we find 1736 materials. In Fig. 1c, we present the probability that compounds containing a given element have $\theta_D > 300$ K. We observe that light elements that form strong covalent bonds, including the 1*s*, 2*p*, and 3*p* elements, as well as many lighter 3*d* transition metals, tend to have high θ_D, while denser transition metals and elements that form weaker bonds tend to have low θ_D. There are multiple methods to calculate Debye temperature[45,46]. In this work, we primarily use the elastic tensor data, as it was available before our phonon calculations and thus available for pre-screening, but we compare to an approach based on the phonon DOS in Supplementary Figure 1.

Performing fully converged electron–phonon coupling calculations using DFT-PT is very computationally expensive, as the calculations in general require both high density *k*-point grids to sample to electronic states and high *q*-point sampling for the phonons[47], and the number of modes to calculate at each q-point increases with the number of atoms in the primitive cell. Therefore, we perform a large number of convergence checks to understand the minimal set of convergence criteria that provide useful information in identifying possible high-T_C conventional superconductors (see Supplementary Table 1, Supplementary Table 2, and Supplementary Table 3). We show the convergence of λ, ω_{log}, and T_C with respect to q-points in Supplementary Fig. 2, Supplementary Fig. 3, and Supplementary Fig. 4, respectively. The comparison for the same for PBE[48] vs PBEsol are shown in Supplementary Fig. 5, Supplementary Fig. 6, Supplementary Fig. 7, and Supplementary Fig. 8, respectively.

In Fig. 2, we show the effect of broadening parameter on EPC and T_C values. As an example of convergence results, in Fig. 2a, b, and c, we show convergence with several k-point and q-point settings for the VC compound. From Fig. 2a and b, it is clear that denser k-points are necessary for reasonable λ and ω_{log} values. We find that larger broadenings are necessary for lower q-point grids. Similar behavior was also found by Shipley et al. for rare-earth hydride system[14].

Table 1. Comparison of JARVIS-Tools screening workflow based (J-Scr) critical transition temperatures in K with experimental data (Exp)[7,55,56,106–111], SCDFT[53], and LM methods[15,112].

Form.	Spg	JID	Exp	SCDFT	LM	J-Scr
Al	225	816	1.2	1.55	0.3	1.6
Ta	229	1014	4.5	5.17	2.45	7.6
Pb	225	961	7.2	6.06	4.95	5.4
Nb	229	934	9.3	10.29	7.0	10.7
ZrN	225	19679	10.0	11.6	6.12	10.0
V_3Si	223	14960	17	18.1	13.1	17.6
MgB_2	191	1151	39	35.4	20.04	33.0
V	229	14837	5.3			18.3
Nb_3Si	223	15938	18.0			16.5
NbO	221	14492	1.38			3.6
NbC	225	19889	12			17.1
NbN	221	36335	16.0			17.6
YB_6	221	20620	7.2			5.1
Nb_3Al	223	11981	16.8			9.0
YH_{10}(250 GPa)	225		260			213.5
LaH_{10}(250 GPa)	225		211			190.0

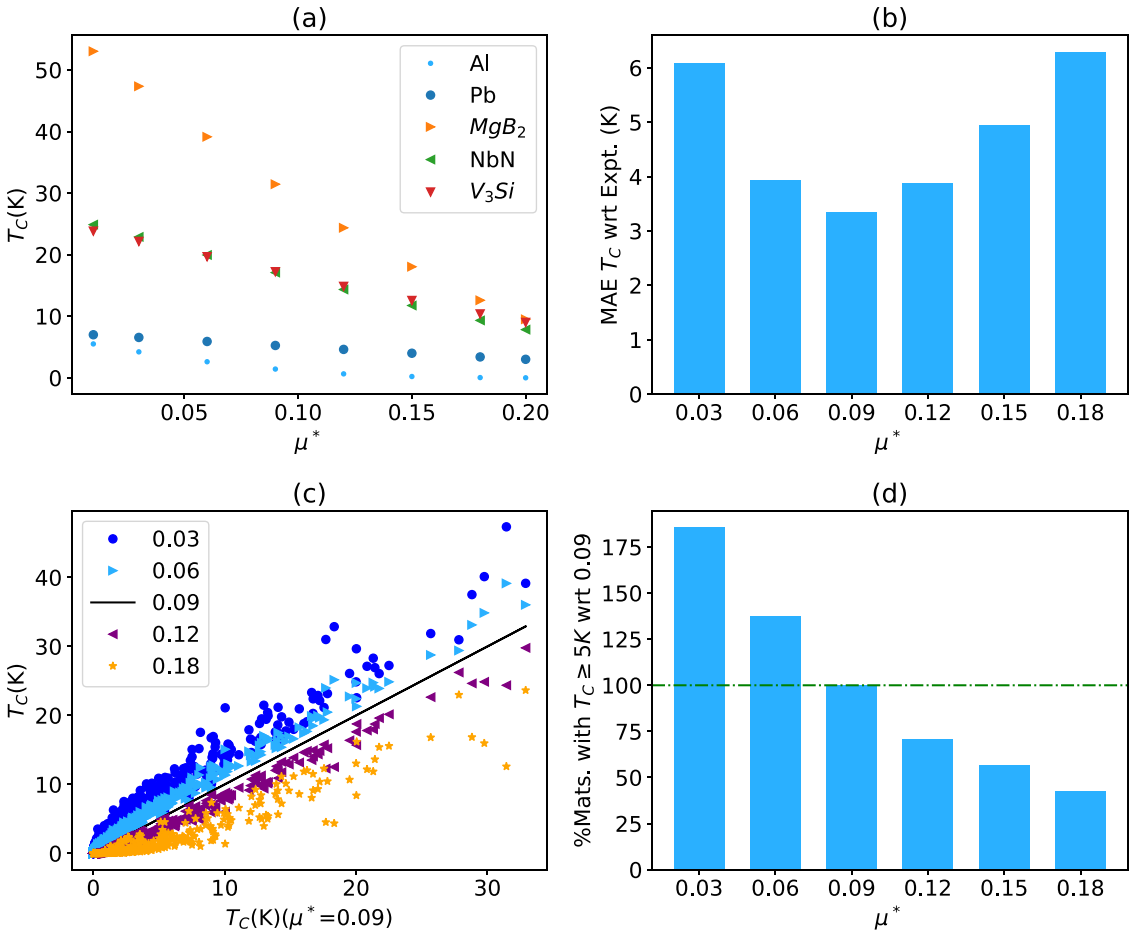

Fig. 4 Effect of μ^* on transition temperatures. a Change in T_C as we increase μ^* for a few representative superconductors, **b** change in mean absolute error (MAE) with respect to experimental data with varying μ^* suggesting an optimal value of 0.09, **c** change in T_c with respect to μ^* 0.09 for the entire dataset, **d** change in percenatge materials with T_C greater equal to 5 K. Here μ^* 0.09 is shown as 100%. As we decrease μ^*, we add materials and vice-versa.

In Fig. 2a, we find that though for $20\times20\times20$ k-points and $2\times2\times2$ q-points case, λ and ω_{log} values (panel a and panel b) are different from other selections, the T_C value is still closer to $10\times10x\times10$ and $2\times2\times2$ k-points and q-points (panel c) respectively. We find that T_C for $10\times10\times10$ k-points and $10\times10\times10$ q-points (black plus sign) are similar to that of $10\times10\times10$ k-points and $5\times5\times5$ q-points (orange upper triangle), suggesting that increasing q-points do not affect much for the $10\times10\times10$ k-points in terms of predicting T_C. Similarly, for $10\times10\times10$ k-points, $2\times2\times2$ q-points (right green triangle); $20\times20\times20$ k-points, $2\times2\times2$ q-points (left purple triangle), the T_C values are closer suggesting in this case increasing k-points do not change the T_C much for $20\times20\times20$ k-points. Interestingly, for the $20\times20\times20$ k-points and $4\times4\times4$ q-points case, the T_C lies in-between the above two extremes of T_C. We note that though ω_{log} seem to converge and overlap for right green and downward red triangles in Fig. 2b, the λ value is still slightly different (as from Fig. 2a) leading to large difference in T_C values indicating that it harder to converge λ that ω_{log}.

Next, we compare the effect of broadening parameters on the EPC parameters and T_C for the entire database in Fig. 2d, e, and f. We find relatively consistent results for broadenings of 0.03–0.05 with our relatively sparse q-point grids, while smaller broadenings lead to erratic unconverged results. In Fig. 3 we compare the effects of broadening on our prediction of T_C for 14 known superconductor (see Table 1), and we find that higher broadening values give better agreement.

We investigate the effect of μ^* parameter[29,47,49–52] in Eq. (7) on T_C for several materials in Fig. 4. In Fig. 4a, we show the effect of different μ^* on five well-known conventional superconductors (Al, Pb, MgB_2, NbN, V_3Si). We find that as we increase the μ^* values, the T_C decreases, as expected. The rate of decay somewhat is dependent on the material, but the materials remain in largely the same order as μ^* is varied. Furthermore, we compute the mean absolute error in predicting T_C of 14 known superconductors in Table 1 using different μ^* values. We find that minimum MAE was obtained for $\mu^* = 0.09$. Similarly, we compare the T_C for the entire database for different μ^* in Fig. 4c. We compare these values to $\mu^* = 0.09$ which is shown as a solid black line. We find that increasing μ^* decreases the T_C similar to Fig. 4a. Now, we quantify, how many materials in our dataset have $T_C \geq 5$ K as we vary μ^*. Relative to $\mu^* = 0.09$, (denoted by dotted green line), we find that we gain 85.5% and 37.4% materials for μ^* 0.03 and 0.06, respectively. As we increase the μ^* to 0.12, 0.15, and 0.18 we lose 29%, 43.5% and 57.3% materials, respectively.

Based on these tests, we find that q-point grids as small as $2\times2\times2$, combined with k-point grids similar to the typical grids used for self-consistent DFT total energy calculations are already useful in identifying candidate materials. Therefore, we adopt the strategy of performing an initial screening step with low convergence parameters applied to a high number of materials, which will be further refined later. During this step, we use the same k-points grid found during the JARVIS-DFT total energy

Table 2. JARVIS-Tools screening workflow based T_C for some of the potential candidate superconductors (T_C), their chemical formula (Form.), spacegroup number (Spg), JARVIS-DFT ID (JID), Inorganic Crystal Structure Database ID (ICSD)[61] wherever available, JARVIS-DFT based formation energy (E_{form} (eV/atom)) and energy above convex hull (E_{hull} (eV)). For the complete list, refer to the supplementary information.

Form.	Spg	JID	ICSD	E_{form}	E_{hull}	T_C(K)
MoN	187	16897	187185	−0.47	0.09	33.4
CaB_2	191	36379	237011	−0.25	0.09	31.0
ZrN	194	13861	161885	−1.76	0.18	30.0
VC	225	19657	619079	−0.48	0.06	28.1
V_2CN	123	105356	–	−0.82	0.11	26.2
Mn	225	25344	41509	0.08	0.08	23.0
NbFeB	187	4546	–	−0.15	0.39	22.1
$NbVC_2$	5	102190	–	−0.46	0.08	21.9
ScN	225	15086	290470	−2.15	0.0	20.8
LaN_2	2	118592		−1.05	0.0	20.4
VRu	221	19694	106010	−0.22	0.01	20.3
$TiReN_3$	161	36745	–	−0.68	0.10	20.0
B_2CN	51	91700	183794	−0.53	0.19	19.4
KB_6	221	20067	98987	−0.09	0.0	19.0
$ZrMoC_2$	166	99893	–	−0.49	0.08	17.9
TaB_2	191	20082	30420	−0.60	0.0	17.2
NbS	194	18923	44992	−0.98	0.05	17.0
$TaVC_2$	166	101106	–	−0.54	0.05	16.3
TaC	187	36405		−0.24	0.40	16.1
MgBH	11	120827	–	−0.03	0.11	15.5
CoN	216	14724	236792	−0.02	0.0	15.0
$NbRu_3C$	221	8528	77216	−0.02	0.19	15.0
TiB_4Mo	191	105371	–	−0.67	0.06	14.8
Al_4CO	6	53086	–	−0.60	0.41	14.4
ScC	194	37100	–	−0.02	0.20	14.3
TaS	187	4699	52114	−0.74	0.11	14.2
VOs	221	122961		−0.26	0.0	14.1
RuO_2	136	19852	236962	−1.23	0.0	14.0
NbVCN	160	99944	–	−0.78	0.15	13.6
NbB_2	191	14726	614908	−0.65	0.0	13.8
$CaRu_2N_2$	164	108175	–	−0.32	0.10	13.3
ZrNbCN	160	102642	–	−1.18	0.06	13.1
V_3Al	221	36199		−0.15	0.14	13.0
TiO	221	50092		−2.65	0.08	12.6
$LiBe_2Ru$	123	71476		−0.14	0.08	12.5
V_2TcRu	225	41603		−0.3	0.0	12.3
$ScSi_2$	191	4406	651822	−0.31	0.23	12.2
VTe	221	122995		0.22	0.50	11.3
AsP_2	191	120382		0.45	0.45	11.2
Cr_3GaN	221	15120	626036	−0.30	0.04	10.9
CrH	225	18397	626118	0.06	0.06	10.7
$TiReTc_2$	225	39667		−0.27	0.0	10.7
ZrSc	187	100313		−0.002	0.01	10.3
V_3Pt	223	20434	649824	−0.43	0.0	10.3
La_3SnC	221	101642	–	−0.60	0.0	10.2
NbS	187	20636	645320	−1.0	0.07	10.1
FeN	216	15718	236789	−0.30	0.0	10.0
InP	225	17716	188692	−0.01	0.20	10.5
V_3Ni	223	15777	647028	−0.16	0.0	10.0
$NbRu_3$	221	16588	105224	−0.12	0.0	10.0
MnN	216	78284	236788	−0.44	0.0	10.0
$CaZrO_2$	99	114723		−2.76	0.30	10.0
ScSe	225	19974	44972	−1.68	0.0	9.1

convergence, as well as a q-point grid of at least $2 \times 2 \times 2$ with a broadening of 0.05 Rydberg (≈0.68 eV).

We compare the present DFT-screening workflow (J-Scr) with other methods such as superconducting density functional theory (SCDFT)[53], Lüders-Marques (LM)[50,54] and experiments[7,55,56] in Table 1. Note that we ignore spin–orbit coupling, spin-polarization, and spin-fluctuation contributions[57–60] during the calculations because of the additional computational cost, but these effects should be considered in follow-up investigations. Importantly, we find that for the top seven systems shown in Table 1, the T_C from the McMillan-Allen-Dynes formula, using $\mu^* = 0.09$, from our J-Scr workflow has an excellent agreement with respect to both SCDFT and the experiments, justifying our approximations. We find that the mean absolute error (MAE) for J-Scr and experiment for all the top 14 near ambient condition superconductors is 3.3 K. Now, comparing the top 7 materials in Table 1 with respect to SCDFT and LM, the MAE of experiment vs J-Scr, experiment vs SCDFT, experiment vs LM, SCDFT vs J-Scr are: 1.9 K, 1.4 K, 4.9 K, and 1.1 K, respectively.

This suggests that there is a close agreement between the computational methods. We also evaluate the workflow for the well-known high-pressure hydrides[14–16] such as LaH_{10} and YH_{10} at 250 GPa as shown in Table 1. We find the MAE to be 25 K for these two systems, which is reasonable for high-throughput type screening[14–16]. Therefore, the workflow can be applied to study high-pressure superconductivity as well. However, for the present work, we analyze materials without external pressure.

Next, we apply the J-Scr workflow on materials with number of atoms less than equal to 5, $\theta_D > 300$K and $N(0) > 1$ as discussed in the previous section. As of now, we have applied J-Scr to 1058 materials. Out of these, only 626 of them are dynamically stable (i.e., no imaginary phonon modes) and 105 of them have $T_C > 5$ K, which is very promising. Table 2 provides the chemical formula, spacegroup, JARVIS-DFT ID, Inorganic Crystal Structure Database (ICSD) ID[61], and the OptB88vdW-based formation energy from JARVIS-DFT and predicted T_C for some of the best candidates for experimental synthesis. The JARVIS-IDs can be used to obtain more detailed information, e.g., https://www.ctcms.nist.gov/knc6/static/JARVIS-DFT/JVASP-16897.xml for JARVIS-ID: 16897

We find that several potential candidates are based on nitrides, borides and Vanadium containing compounds. All of these systems have high symmetry, with spacegroup numbers ranging from 99 to 225, although this may be an artifact of pre-selecting systems with few atoms. We visualize some of the crystal structures in Supplementary Fig. 9. Many of these systems have hexagonal and cubic symmetry. The JARVIS-DFT identifiers for corresponding systems can be used to obtain further properties. Importantly, most of these systems have negative formation energies and energy above convex hull less than 0.5 eV/atom. We also show Fermi-surfaces of such example compounds in Fig. S9. The shape of the Fermi-surface depicts the electron motion inside a material[62]. In earlier works[63] related to MgB_2 type compounds, the shape of the Fermi surface helped understand the mechanism of superconductivity in such compounds.

As shown in Table 2, some of our highest T_C candidate superconductors include MoN, VC, Mn, MnN, LaN_2, KB_6, TaC, etc.

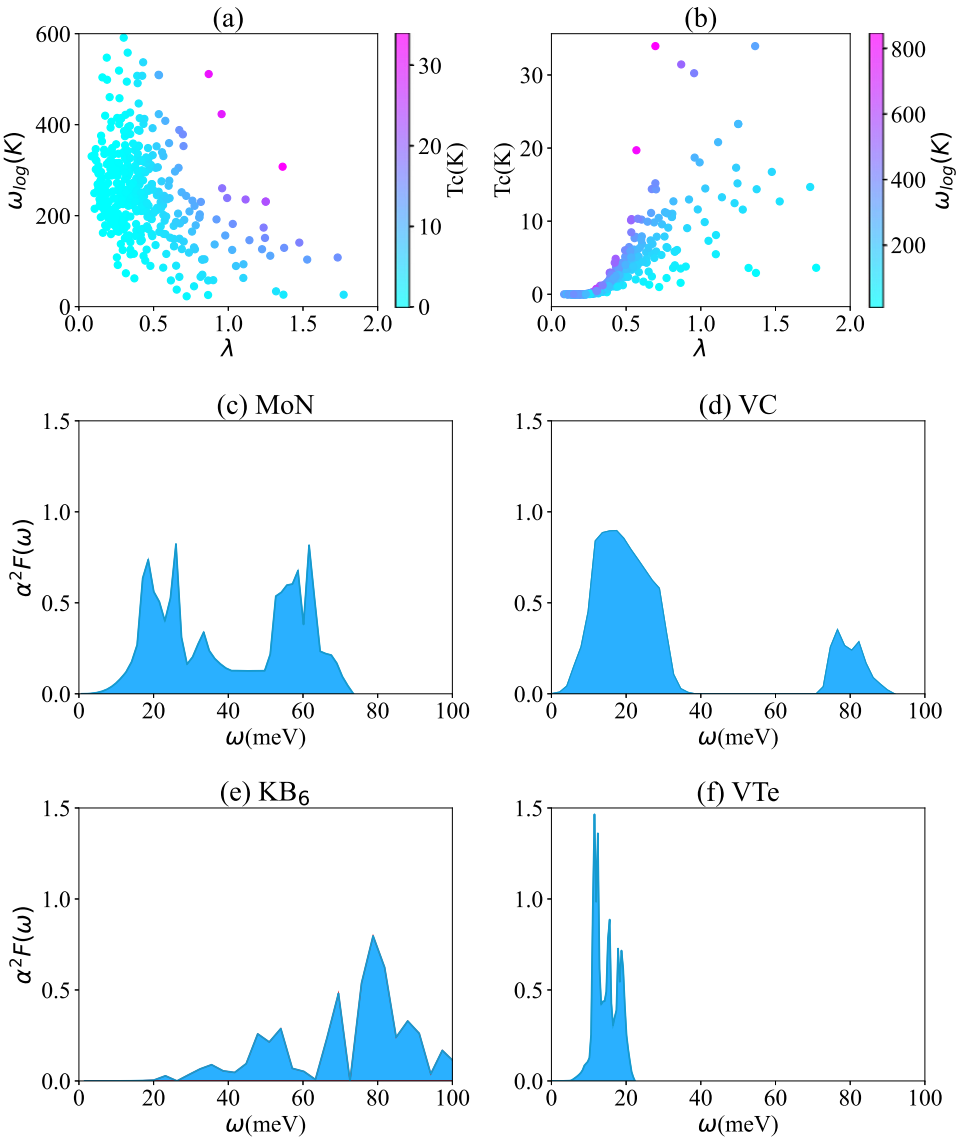

Fig. 5 Relation between electron–phonon coupling parameters and EPC function of some of the potential candidate superconductors. **a** ω_{log} vs λ, **b** T_C vs λ, **c** MoN, **d** VC, **e** KB_6, and **f** VTe.

MoN in a rock-salt structure has been previously been reported with a T_C as high as 30 K[64,65]; however, here we propose a hexagonal form which has not been observed experimentally. Similarly, ZrN in rock-salt form is known to have T_C of 10 K; however, here we propose a hexagonal form of ZrN to have a T_C around 30 K. CaB_2 was previously theoretically predicted be a high-T_C superconductor[66]. Previous reports of synthesizing VC in 1:1 ratio has been found to be challenging by previous researchers[67]. Lanthanum nitride[68] in 1:1 ratio compound has been found to have superconducting properties. Lanthanum nitride in 1:2 ratio is not characterized for the superconducting properties to the best of our knowledge. Vanadium Ruthenium alloy with a different ratio than 1:1 has been reported earlier as superconducting with T_C close to 5 K[69]. KB_6 has been proposed to be a high-T_C superconductor as well[70]. Tantalum carbide has been shown to be a potential topological superconductor with a T_C of 10.3 K[71]. Li et al.[72] recently reported the CoN system as a potential superconductor with T_C of 16 K. Nb-Ru-C alloy system has been synthesized experimentally before[73], but no superconducting data is available to the best of our knowledge. Ta-S system with a 1:2 ratio is a superconductor[74] with T_C of 5.4 K, however, a 1:1 ratio material i.e., Tantalum rich system has not been characterized for superconducting properties yet. Nb-B system has been reported earlier with a T_C of around 9.8 K[75]. V-Al alloy system can attain T_C of around 11.15 K as reported by Kodess et al.[76]. The Sc-Si system has been realized experimentally before but their superconducting properties has not been reported[77]. 2D B_2O system was recently proposed with a T_C of 10.3 K[78]. Cr_3GaN belongs to the cubic anti-perovskite family of superconductors and has been predicted to be superconductor by Tutuncu et al.[79]. V_3Pt belongs to the A15 family and has been investigated for superconducting properties earlier[80]. Fe-N system with 1:2 ratio was shown with a T_C of 8 K[81].

Next, we show the relationship between EPC parameters for the dynamically stable compounds and the Eliashberg spectral functions of some of the candidate materials in Fig. 5. In Fig. 5a, we find that λ and ω_{log} to have an inverse relationship while in Fig. 5 we see that λ and T_C tend to follow a linear relationship, both of which are typical for BCS superconductors. For high-T_C, high ω_{log} as well as high λ is favorable as evident from the colormap. The EPC Eliashberg spectral function expresses the electron–phonon interaction in the form of a spectral density. The weighted area

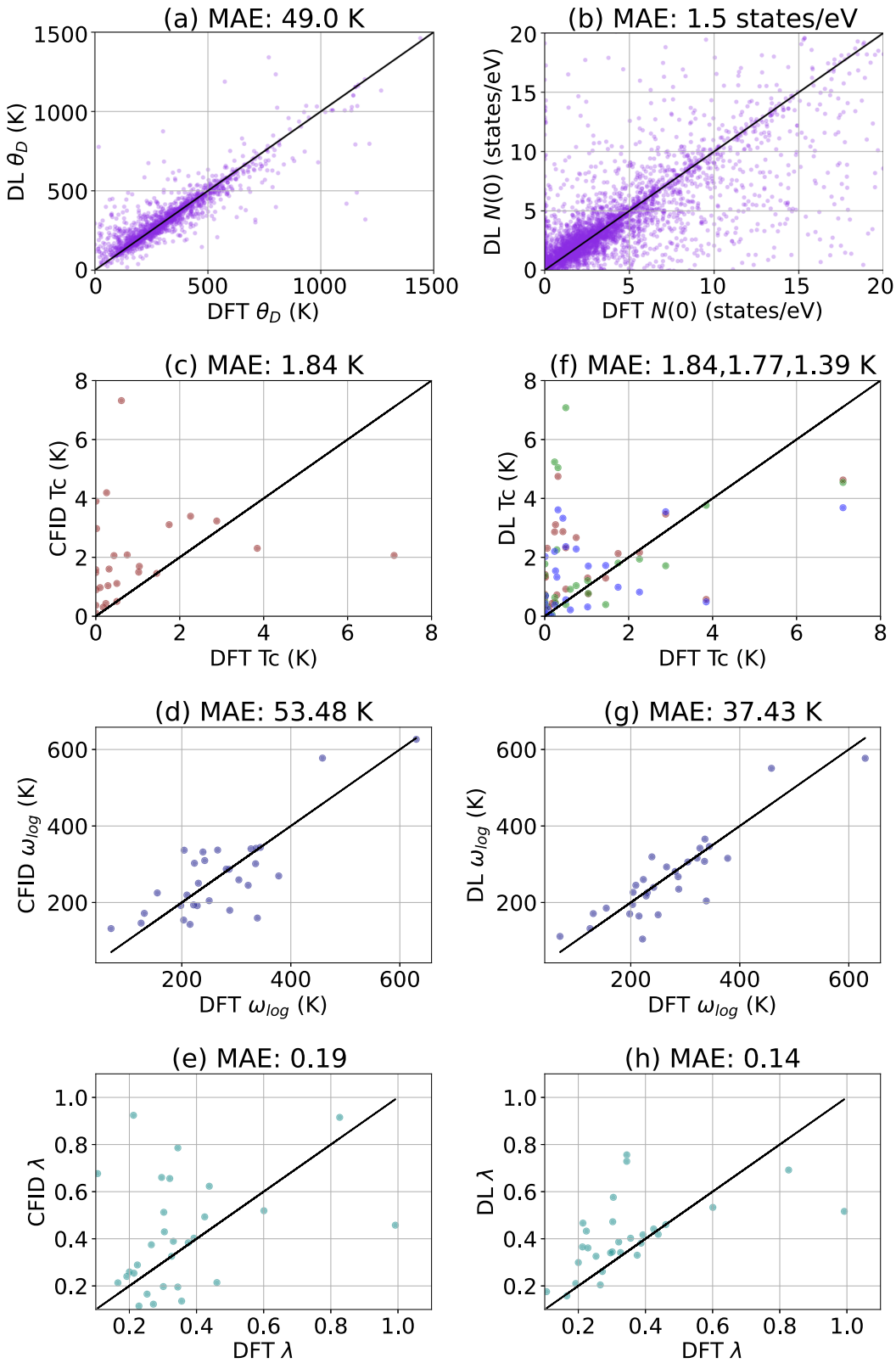

Fig. 6 Machine learning model performance. Atomistic line graph neural network based deep-learning (DL) regression model performance on 5% test set for **a** Debye temperature and **b** DOS. Classical force-field descriptor (CFID) (**c**, **d**, **e**) and DL (**f**, **g**, **h**) based regression model performance on 5% test set for DFT calculated transition temperature (T_C), EPC parameter ω_{log}, and EPC parameter λ. In **f**, we show performance with direct T_C prediction (red color), T_C prediction with direct prediction of wlog and lambda and then using Eq. (7) (green color) and T_C prediction with Eliashberg function and then using Eqs. (5–8) (black color).

Table 3. DL model performance for direct property predictions. We compare the mean absolute error (MAE) of a baseline model, the MAE for the DL model on the 5% test data, and the root mean squared error (RMSE) for the 5% test data. The baseline model always predicts the mean of the target values in the training dataset.

Property	BaselineMAE	TestMAE	TestRMSE
θ_D (K)	145.5	49.0	124.62
$N(0)$ (states/eV/Nelect)	3.62	1.5	4.40
λ	0.20	0.14	0.24
ω_{log} (K)	76.0	37.43	50.73
T_C (K)	2.72	1.84	3.31

under the EPC function determines the λ as well as ω_{log} parameters.

Machine learning

Next, we develop deep-learning models to accelerate both our initial BCS-inspired screening and our calculation of the electron–phonon coupling parameters.

While the BCS prescreening step is much less expensive than the full EPC calculation, it still requires the DOS and θ_D, which require substantial computation. Therefore, we develop regression models to predict these properties directly from an arbitrary crystal structure, using the large datasets available in the JARVIS-DFT database. There are multiple methods to establish the structure-property relation for crystal structures, but in this work we use atomistic line graph neural network (ALIGNN) model[33], which has been shown to outperform many well-known benchmarks for solids and explicitly capture chemical and many-body physical interactions in materials. Our results of these models on 5% held test sets are shown in Fig. 6 and Table 3. The baseline model MAE was computed by using mean of the target values in the training dataset and using it as predictions for all the materials in the 5% test data. We observe that the mean absolute error for the Debye temperature is 49 K while that for DOS is 1.5 states/eV/Nelect. The baseline model MAE for the Debye temperature and DOS are 145.5 K and 3.62 states/eV/Nelect respectively. Using these trained models, we applied it to Crystallography Open Database (COD)[82] with 431,778 materials and with number of atoms less than 100, and pre-screened 8293 materials with high Debye temperature and DOS values.

Next, we develop machine-learning models to directly predict EPC parameters and T_C using our database of 626 dynamically stable data-points from the j-Scr developed above. We use two methods: hand-craft descriptors (Classical force-field descriptor (CFID)[83]) and a deep-learning approach (ALIGNN)[33]. In particular, we train models for the McMillan–Allen–Dynes transition temperature (T_C) and the electron phonon coupling parameters- ω_{log} and λ. We note that ML models usually require large datasets with sample sizes in the range of thousands, but we show preliminary but useful results already with our current dataset, which will continue to grow.

Classical force-field descriptor (CFID)[83] based performances are shown in Fig. 6c, d, and e for T_C, ω_{log}, and λ respectively. Similarly, ALIGNN based performances are shown in Fig. 6f, g, and h. For a perfect agreement, the data-points should lie on the $y = x$ line. We observe that the MAE using the CFID approach for T_C, ω_{log} and λ are 1.84 K, 53.48 K and 0.19. Similarly, the MAEs for ALIGNN approach are 1.84 K, 37.43 K and 0.14, respectively. ALIGNN outperforms CFID in predicting ω_{log}, but the other performances are similar. Moreover, we notice that it is significantly easier to learn ω_{log} than T_C and λ as evident from the parity plots and Table 3. In particular, the model for λ is only slightly better than the baseline model. Using the ALIGNN models for ω_{log} and λ, and Eq. (7), we predict the T_C with an MAE of 1.77 K. We apply the ALIGNN TC predictor model on the COD database and find that 39,595 candidate materials can have $T_C \geq 10$ K and 2161 of them have $T_C \geq 15$ K.

As predicting λ directly using ML models is evidently challenging, we attempt an alternate method to directly predict Eliashberg function using the ALIGNN model. We choose an energy range of 0 to 100 meV with 1 meV binsize and predict the Eliashberg functions. We show the DFT and ALIGNN based Eliashberg functions for samples in the test set in Fig. 7. We find that the ALIGNN model does an good job of capturing most of the peaks. We calculate the T_C using the ALIGNN based Eliashberg function predictions, and find the MAE is 1.39 K, which improves on our direct prediction method above by 24%. Interestingly, we have observed a similar behavior for predicting phonon properties using phonon density of states as an intermediate quantity in ref. [84]. This suggests that learning more fundamental and information-rich quantities such as Eliashberg functions can be useful for ML approaches with limited data, as compared to more direct predictions of integrated quantities. Unfortunately, a drawback of deep-learning models is that it is difficult to extract physical insight from their internal parameters, but we hope to investigate these ideas further in future works.

The above ML-exercise indicates that while ω_{log} is relatively straightforward to predict using the crystal structure, λ and consequently T_C are much more challenging. We note that deep learning models usually require larger datasets of at least ten thousand entires, which remains well beyond the data available in this work. In addition to larger datasets, these results should motivate future works on designing better descriptors and models.

In summary, we have developed a combined high-throughput DFT and ML approach to study conventional superconductors, finding over 100 candidate materials with predicted $T_C \geq 5$K. We provide data from this work as well as machine learning models to help accelerate the discovery of superconductors. Because we employ a high-throughput approach to screen large databases, we employ several assumptions, but we perform significant benchmark testing as well as detailed convergence checks on particular materials found to be promising in our initial screening to verify our results. We have made our datasets and tools publicly available to enhance the reproducibility and transparency of our work. We believe that our work can be of great help to guide future computational as well as experimental efforts to discover and characterize BCS-superconductors.

METHODS

In this section, we discuss details of various computational methods.

BCS inspired screening

According to BCS-theory[85], the attractive electron-electron interaction mediated by phonons gives rise to Cooper pairs, i.e., bound states are formed by two electrons with opposite spins and momenta. BCS-theory provides the relation between the superconducting transition temperature (T_C), Debye temperature (θ_D), electronic DOS at Fermi level $N(0)$, and electron–phonon interaction (V) as follows:

$$T_c = 1.14\theta_D \exp\left(-\frac{1}{N(0)V}\right) \quad (1)$$

θ_D is defined as[45]:

$$\theta_D = \frac{h}{k_B}\left[\frac{3nN_a\rho}{4\pi M}\right]^{\frac{1}{3}} v_m \quad (2)$$

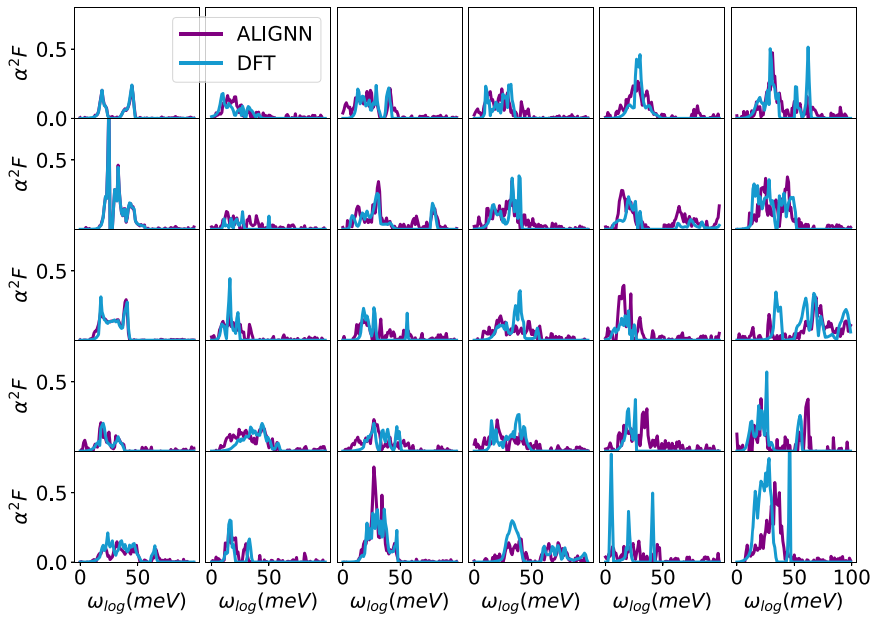

Fig. 7 Prediction of Eliashberg function with ALIGNN for the 5% test set. ALIGNN can capture peak-positions and heights reasonably well.

where h is Planck's constant, k_B is the Boltzmann constant, n is the number of atoms per formula unit, N_A is Avogadro constant, ρ is the crystal structure's density, M is the molar mass, and v_m is the average sound velocity obtained from the elastic tensor[45].

We use the θ_D obtained from finite-difference calculations of elastic tensors[40], and $N(0)$ is obtained from the DOS at the Fermi-level as available in the JARVIS-DFT database. Currently, there are electronic DOS database available for 55,723 materials and elastic tensors for 17,419 materials (v08.18.2021) and the database is continuously expanding. We normalize the electronic density of states by the total number valence electrons in the DFT calculation. JARVIS-DFT is primarily based on Vienna Ab initio Simulation Package (VASP)[86,87]. software and OptB88vdW[88] functional but also contains data with different functionals and methods. In JARVIS-DFT k-points are converged with respect to total energy[44]. These converged k-points are also used in subsequent electron–phonon calculations using Quantum Espresso package[89,90] (see below).

Density functional theory (DFT) calculations

The superconducting properties of conventional superconductors depend on EPC. There are several methods to perform EPC calculations such as interpolated/Gaussian broadening method[47], tetrahedron method[91] and other Wannier-based electron-phonon method[92], etc. In this work, we primarily report the interpolated method, but we also compare with the tetrahedron method. We perform EPC calculations using DFT-PT[93,94] with the Quantum Espresso software package[89] and the GBRV[95] pseudopotentials. We report results with the PBEsol[96] functional in the main text, and we compare to the PBE[48] functional in the Supplementary Table 2. We begin with structures from the JARVIS-DFT database, and perform full relaxation using our Quantum Espresso settings. The EPC parameter is derived from spectral function $\alpha^2F(\omega)$ which is calculated as follows:

$$\alpha^2F(\omega) = \frac{1}{2\pi N(\epsilon_F)} \sum_{qj} \frac{\gamma_{qj}}{\omega_{qj}} \delta(\omega - \omega_{qj}) w(q) \tag{3}$$

where ω_{qj} is the mode frequency, $N(\epsilon_F)$ is the DOS at the Fermi level ϵ_F, δ is the Dirac-delta function, $w(q)$ is the weight of the q point, γ_{qj} is the linewidth of a phonon mode j at wave vector q and is given by:

$$\gamma_{qj} = 2\pi\omega_{qj} \sum_{nm} \int \frac{d^3k}{\Omega_{BZ}} |g^j_{kn,k+qm}|^2 \delta(\epsilon_{kn} - \epsilon_F) \delta(\epsilon_{k+qm} - \epsilon_F) \tag{4}$$

Here, the integral is over the first Brillouin zone, ϵ_{kn} and ϵ_{k+qm} are the DFT eigenvalues with wavevector k and $k+q$ within the nth and mth bands respectively, $g^j_{kn,k+qm}$ is the electron–phonon matrix element. γ_{qj} is related to the mode EPC parameter λ_{qj} by:

$$\lambda_{qj} = \frac{\gamma_{qj}}{\pi h N(\epsilon_F) \omega^2_{qj}} \tag{5}$$

Now, the EPC parameter is given by:

$$\lambda = 2 \int \frac{\alpha^2F(\omega)}{\omega} d\omega = \sum_{qj} \lambda_{qj} w(q) \tag{6}$$

with $w(q)$ as the weight of a q point.

The superconducting transition temperature, T_C can then be approximated using McMillan–Allen–Dynes[32] equation as follows:

$$T_c = \frac{\omega_{log}}{1.2} \exp\left[-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}\right] \tag{7}$$

where

$$\omega_{log} = \exp\left[\frac{\int d\omega \frac{\alpha^2F(\omega)}{\omega} \ln \omega}{\int d\omega \frac{\alpha^2F(\omega)}{\omega}}\right] \tag{8}$$

We note that several variants of McMillan–Allen–Dynes formula have been proposed previously to better account for high-λ materials[55,97–99], however, in this work, we opt for the simpler original McMillan–Allen–Dynes formula as our screening tool. Furthermore, we note that materials with narrow electronic bandwidth may not be well-reproduced by the formula. In Eq. (7), the parameter μ^* is the effective Coulomb potential parameter. While there are ways to calculate this parameter from first principles[49], this parameter generally varies over a relatively small range (such as 0.09–0.18). Following previous works[29,47,49–52], we take a fixed value (here $\mu^* = 0.09$) when reporting our main results, and we present a discussion on the effect of varying μ^* in the later section.

Deep-learning training

Deep-learning is one of the fastest developing methods especially for materials science applications[34]. Graph neural networks (GNN) are based on deep-learning framework because they can work with unstructured and non-Euclidean data i.e., non-grid data such as atomic structures. In this work, we use the recently developed atomistic line graph neural network (ALIGNN)[33], which is publicly available at https://github.com/usnistgov/alignn. ALIGNN has been used to train fast and accurate models for more than 60 properties of solids and molecules with high accuracy. We use ALIGNN to train models for the Debye temperature as well as the electronic density of states at the Fermi-level using large datasets from the JARVIS-DFT database. We also train models to predict the DFT-based transition temperatures (T_C) and the underlying electron-phonon coupling parameters ω_{log} and λ, using the smaller datasets computed in this work. We use both ALIGNN and the hand-crafted (Classical force-field descriptor (CFID)[83])-based approaches for comparison. CFID approach utilizes the LightGBM package[100].

In ALIGNN, a crystal structure is represented as a graph using atomic elements as nodes and atomic bonds as edges. Each node in the atomistic graph is assigned 9 input node features based on its atomic species: electronegativity, group number, covalent radius, valence electrons, first ionization energy, electron affinity, block and atomic volume. The inter-atomic bond distances are used as edge features with radial basis function up to 8 Å cut-off. We use a periodic 12-nearest-neighbor (N) graph construction. This atomistic graph is then used for constructing the corresponding line graph using interatomic bond-distances as nodes and bond-angles as edge features. ALIGNN uses edge-gated graph convolution for updating nodes as well as edge features using a propagation function (f) for layer (l), atom features (h), and node (i), details of which can be found in ref. [33]:

$$h_i^{(l+1)} = f(h_i^l \{h_j^l\}_i) \quad (9)$$

Unlike many other conventional GNNs ([101–103]), ALIGNN uses bond-distances as well as bond-angles to distinguish atomic structures. One ALIGNN layer composes an edge-gated graph convolution on the bond graph with an edge-gated graph convolution on the line graph. The line graph convolution produces bond messages that are propagated to the atomistic graph, which further updates the bond features in combination with atom features. The ALIGNN model is implemented in PyTorch[104] and deep graph library (DGL)[105]. The hyperparameters for ALIGNN are kept same as the original paper. We use the available 17419 Debye temperature data and 55645 electronic DOS at Fermi level data for model training. In addition, we use the 626 electron-phonon coupling parameters for dynamically stable materials, their transition temperatures and Eliashberg functions. For the Debye temperature and electronic density of states at the Fermi-level, we use a batch size of 64 for 500 epochs and 80:10:10 training-validation-testing data split, while for predicting the EPC parameters and Eliashberg functions, we use a batch size of 16, 90:5:5 split and training for 300 epochs. The test set was never used during the training procedure. We train use Tesla V100 SXM2 32 gigabyte Graphics processing unit (GPU), with 8 Intel Xeon E5-2698 v4 CPU cores for concurrently fetching and preprocessing batches of data during training. Please note commercial software and hardware are identified to specify procedures. Such identification does not imply recommendation by National Institute of Standards and Technology.

DATA AVAILABILITY

The dataset is available at the Figshare repository: https://doi.org/10.6084/m9.figshare.21370572.

CODE AVAILABILITY

Software packages mentioned in the article can be found at https://github.com/usnistgov/jarvis and https://github.com/usnistgov/alignn.

Received: 28 June 2022; Accepted: 4 November 2022;
Published online: 22 November 2022

REFERENCES

1. Kamerlingh Onnes, H. The resistance of pure mercury at helium temperatures. Commun. Phys. Lab. Univ. Leiden, b 120 (1911) .
2. Poole, C. P., Farach, H. A. & Creswick, R. J. Superconductivity (Academic press, 2013).
3. Rogalla, H. & Kes, P.H. 100 years of superconductivity (Taylor & Francis, 2011).
4. Nagamatsu, J., Nakagawa, N., Muranaka, T., Zenitani, Y. & Akimitsu, J. Superconductivity at 39 k in magnesium diboride. *Nature* **410**, 63–64 (2001).
5. Pickard, C. J., Errea, I. & Eremets, M. I. Superconducting hydrides under pressure. *Annu. Rev. Condens. Matter Phys.* **11**, 57–76 (2020).
6. Sato, M. & Ando, Y. Topological superconductors: a review. *Rep. Prog. Phys.* **80**, 076501 (2017).
7. Roberts, B. W. Survey of superconductive materials and critical evaluation of selected properties. *J. Phys. Chem. Ref. Data* **5**, 581–822 (1976).
8. Kihlstrom, K., Simon, R. & Wolf, S. Tunneling α2f (ω) on high tc a15 and b1 compounds. *Physica B+ C* **135**, 198–202 (1985).
9. Stewart, G. R. Superconductivity in the a15 structure. *Phys. C: Supercond. Appl* **514**, 28–35 (2015).
10. Ivanovskii, A. Band structure and properties of superconducting mgb2 and related compounds (a review). *Phys. Solid State* **45**, 1829–1859 (2003).
11. Buzea, C. & Yamashita, T. Review of the superconducting properties of mgb2. *Supercond. Sci. Technol.* **14**, R115 (2001).
12. Plakida, N. High-temperature cuprate superconductors: experiment, theory, and applications. Vol. 166 (Springer Science & Business Media, 2010).
13. Hosono, H. & Kuroki, K. Iron-based superconductors: current status of materials and pairing mechanism. *Phys. C: Supercond. Appl* **514**, 399–422 (2015).
14. Shipley, A. M., Hutcheon, M. J., Johnson, M. S., Needs, R. J. & Pickard, C. J. Stability and superconductivity of lanthanum and yttrium decahydrides. *Phys. Rev. B* **101**, 224511 (2020).
15. Liu, H., Naumov, I. I., Hoffmann, R., Ashcroft, N. & Hemley, R. J. Potential high-tc superconducting lanthanum and yttrium hydrides at high pressure. *Proc. Natl Acad. Sci.* **114**, 6990–6995 (2017).
16. Drozdov, A., Eremets, M., Troyan, I., Ksenofontov, V. & Shylin, S. I. Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system. *Nature* **525**, 73–76 (2015).
17. Shipley, A. M., Hutcheon, M. J., Needs, R. J. & Pickard, C. J. High-throughput discovery of high-temperature conventional superconductors. *Phys. Rev. B* **104**, 054501 (2021).
18. Yuan, J., Stanev, V., Gao, C., Takeuchi, I. & Jin, K. Recent advances in high-throughput superconductivity research. *Supercond. Sci. Technol.* **32**, 123001 (2019).
19. Rodriguez, C. et al. Optical near-zone-center phonons and their interaction with electrons in yba 2 cu 3 o 7: results of the local-density approximation. *Phys. Rev. B* **42**, 2692 (1990).
20. Subedi, A., Ortenzi, L. & Boeri, L. Electron-phonon superconductivity in a pt 3 p (a= sr, ca, la) compounds: from weak to strong coupling. *Phys. Rev. B* **87**, 144504 (2013).
21. Duan, D., Yu, H., Xie, H. & Cui, T. Ab initio approach and its impact on superconductivity. *J. Supercond. Nov. Magn.* **32**, 53–60 (2019).
22. Kolmogorov, A. et al. New superconducting and semiconducting fe-b compounds predicted with an ab initio evolutionary search. *Phys. Rev. Lett.* **105**, 217003 (2010).
23. Gao, G. et al. High-pressure crystal structures and superconductivity of stannane (snh4). *Proc. Natl Acad. Sci.* **107**, 1317–1320 (2010).
24. Duan, D. et al. Pressure-induced metallization of dense (h2s) 2h2 with high-tc superconductivity. *Sci. Rep.* **4**, 1–6 (2014).
25. Hutcheon, M. J., Shipley, A. M. & Needs, R. J. Predicting novel superconducting hydrides using machine learning approaches. *Phys. Rev. B* **101**, 144505 (2020).
26. Stanev, V. et al. Machine learning modeling of superconducting critical temperature. *NPJ Comput. Mater.* **4**, 1–14 (2018).

27. Xie, S. et al. Machine learning of superconducting critical temperature from eliashberg theory. *NPJ Comput. Mater.* **8**, 1–8 (2022).
28. Cooper, L. N. & Feldman, D. BCS: 50 years (World scientific, 2010).
29. Giustino, F. Electron-phonon interactions from first principles. *Rev. Mod. Phys.* **89**, 015003 (2017).
30. Kawamura, M., Hizume, Y. & Ozaki, T. Benchmark of density functional theory for superconductors in elemental materials. *Phys. Rev. B* **101**, 134511 (2020).
31. Choudhary, K. et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *NPJ Comput. Mater.* **6**, 1–13 (2020).
32. McMillan, W. Transition temperature of strong-coupled superconductors. *Phys. Rev.* **167**, 331 (1968).
33. Choudhary, K. & DeCost, B. Atomistic line graph neural network for improved materials property predictions. *NPJ Comput. Mater.* **7**, 1–8 (2021).
34. Choudhary, K. et al. Recent advances and applications of deep learning methods in materials science. *NPJ Comput. Mater.* **8**, 1–26 (2022).
35. Choudhary, K. et al. Computational screening of high-performance optoelectronic materials using optb88vdw and tb-mbj formalisms. *Sci. Data* **5**, 1–12 (2018).
36. Choudhary, K. et al. Accelerated discovery of efficient solar cell materials using quantum and machine-learning methods. *Chem. Mater.* **31**, 5900–5908 (2019).
37. Choudhary, K., Garrity, K. F., Ghimire, N. J., Anand, N. & Tavazza, F. High-throughput search for magnetic topological materials using spin-orbit spillage, machine learning, and experiments. *Phys. Rev. B* **103**, 155131 (2021).
38. Choudhary, K., Garrity, K. F. & Tavazza, F. High-throughput discovery of topologically non-trivial materials using spin-orbit spillage. *Sci. Rep.* **9**, 1–8 (2019).
39. Choudhary, K., Garrity, K. F., Jiang, J., Pachter, R. & Tavazza, F. Computational search for magnetic and non-magnetic 2d topological materials using unified spin–orbit spillage screening. *NPJ Comput. Mater.* **6**, 1–8 (2020).
40. Choudhary, K., Cheon, G., Reed, E. & Tavazza, F. Elastic properties of bulk and low-dimensional materials using van der waals density functional. *Phys. Rev. B* **98**, 014107 (2018).
41. Choudhary, K. et al. High-throughput density functional perturbation theory and machine learning predictions of infrared, piezoelectric, and dielectric responses. *NPJ Comput. Mater.* **6**, 1–13 (2020).
42. Choudhary, K., Ansari, J. N., Mazin, I. I. & Sauer, K. L. Density functional theory-based electric field gradient database. *Sci. Data* **7**, 1–10 (2020).
43. Choudhary, K., Kalish, I., Beams, R. & Tavazza, F. High-throughput identification and characterization of two-dimensional materials using density functional theory. *Sci. Rep.* **7**, 1–16 (2017).
44. Choudhary, K. & Tavazza, F. Convergence and machine learning predictions of monkhorst-pack k-points and plane-wave cut-off in high-throughput dft calculations. *Comput. Mater. Sci.* **161**, 300–308 (2019).
45. Anderson, O. L. A simplified method for calculating the debye temperature from elastic constants. *JJ Phys. Chem. Solids* **24**, 909–917 (1963).
46. Garrity, K. F. First-principles search for n-type oxide, nitride, and sulfide thermoelectrics. *Phys. Rev. B* **94**, 045122 (2016).
47. Wierzbowska, M., de Gironcoli, S. & Giannozzi, P. Origins of low-and high-pressure discontinuities of *t_{c}* in niobium. Preprint at https://arxiv.org/abs/cond-mat/0504077v2 (2005).
48. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
49. Lee, K.-H., Chang, K.-J. & Cohen, M. L. First-principles calculations of the coulomb pseudopotential μ^*: application to al. *Phys. Rev. B* **52**, 1425 (1995).
50. Marques, M. et al. Ab initio theory of superconductivity. II. Application to elemental metals. *Phys. Rev. B* **72**, 024546 (2005).
51. Morel, P. & Anderson, P. Calculation of the superconducting state parameters with retarded electron-phonon interaction. *Phys. Rev.* **125**, 1263 (1962).
52. Zarifi, N., Bi, T., Liu, H. & Zurek, E. Crystal structures and properties of iron hydrides at high pressure. *J. Phys. Chem. C.* **122**, 24262–24269 (2018).
53. Sanna, A., Pellegrini, C. & Gross, E. Combining eliashberg theory with density functional theory for the accurate prediction of superconducting transition temperatures and gap functions. *Phys. Rev. Lett.* **125**, 057001 (2020).
54. Lüders, M. et al. Ab initio theory of superconductivity. I. Density functional formalism and approximate functionals. *Phys. Rev. B* **72**, 024545 (2005).
55. Allen, P. B. & Dynes, R. Transition temperature of strong-coupled superconductors reanalyzed. *Phys. Rev. B* **12**, 905 (1975).
56. Webb, G., Marsiglio, F. & Hirsch, J. Superconductivity in the elements, alloys and simple compounds. *Phys. C: Supercond. Appl.* **514**, 17–27 (2015).
57. Heid, R., Bohnen, K.-P., Sklyadneva, I. Y. & Chulkov, E. Effect of spin-orbit coupling on the electron-phonon interaction of the superconductors pb and tl. *Phys. Rev. B* **81**, 174527 (2010).
58. Gibson, G. & Meservey, R. Evidence for spin fluctuations in vanadium from a tunneling study of fermi-liquid effects. *Phys. Rev. B* **40**, 8705 (1989).
59. Mandal, S., Cohen, R. E. & Haule, K. Strong pressure-dependent electron-phonon coupling in fese. *Phys. Rev. B* **89**, 220502 (2014).
60. Gerber, S. et al. Femtosecond electron-phonon lock-in by photoemission and x-ray free-electron laser. *Science* **357**, 71–75 (2017).
61. Belsky, A., Hellenbrandt, M., Karen, V. L. & Luksch, P. New developments in the inorganic crystal structure database (ICSD): accessibility in support of materials research and design. *Acta Crystallogr. B Struct. Sci.* **58**, 364–369 (2002).
62. Yamase, H., Sakurai, Y., Fujita, M., Wakimoto, S. & Yamada, K. Fermi surface in la-based cuprate superconductors from compton scattering imaging. *Nat. Commun.* **12**, 1–7 (2021).
63. An, Y. et al. Superconductivity and topological properties of mgb 2-type diborides from first principles. *Phys. Rev. B* **104**, 134510 (2021).
64. Inumaru, K., Nishikawa, T., Nakamura, K. & Yamanaka, S. High-pressure synthesis of superconducting molybdenum nitride δ-mon by in situ nitridation. *Chem. Mater.* **20**, 4756–4761 (2008).
65. Wang, S. et al. The hardest superconducting metal nitride. *Sci. Rep.* **5**, 1–8 (2015).
66. Choi, H. J., Louie, S. G. & Cohen, M. L. Prediction of superconducting properties of cab 2 using anisotropic eliashberg theory. *Phys. Rev. B* **80**, 064503 (2009).
67. Zbasnik, J. & Toth, L. Electronic structure of vanadium carbide. *Phys. Rev. B* **8**, 452 (1973).
68. Vaitheeswaran, G., Kanchana, V. & Rajagopalan, M. Structural phase stability and superconductivity of lan. *Solid State Commun.* **124**, 97–102 (2002).
69. Chu, C., Bucher, E., Cooper, A. & Maita, J. Electronic transitions in near-equiatomic vanadium-ruthenium alloys. *Phys. Rev. B* **4**, 320 (1971).
70. Katsura, Y. et al. On the possibility of mgb2-like superconductivity in potassium hexaboride. *Phys. C: Supercond. Appl.* **470**, S633–S634 (2010).
71. Yan, D. Y. et al. Superconductivity in centrosymmetric topological superconductor candidate tac. *Supercond. Sci. Technol.* **34**, 035025 (2021).
72. Li, Q. & Chen, L. Superconducting atmospheric structure and pressure-induced novel phases of cobalt mononitride. *Comput. Mater. Sci.* **174**, 109464 (2020).
73. Hong, D., Zeng, W., Liu, F.-S., Tang, B. & Liu, Q.-J. Structural, electronic, elastic and mechanical properties of nbc-based compounds: first-principles calculations. *Phys. B Condens. Matter* **558**, 100–108 (2019).
74. Sernetz, F., Lerf, A. & Schöllhorn, R. Superconductivity in hydrated ternary tantalum sulfides. *Mater. Res. Bull.* **9**, 1597–1602 (1974).
75. Ren, Z.-A., Kuroiwa, S., Tomita, Y. & Akimitsu, J. Structural phase evolution and superconductivity in the non-stoichiometric intermetallic compound niobium diboride. *Phys. C: Supercond.* **468**, 411–416 (2008).
76. Kodess, B. N. Superconductivity of some new w compounds. *Phys. Status Solidi (a)* **4**, K109–K112 (1971).
77. Gokhale, A. & Abbaschian, G. The sc- si (scandium-silicon) system. *Bull. Alloy Phase Diagr.* **7**, 333–336 (1986).
78. Yan, L. et al. Theoretical dissection of superconductivity in two-dimensional honeycomb borophene oxide b2o crystal with a high stability. *NPJ Comput. Mater.* **6**, 1–7 (2020).
79. Tutuncu, H. M. & Srivastava, G. Theoretical examination of superconductivity in the cubic antiperovskite cr3gan under pressure. *J. Appl. Phys.* **114**, 053905 (2013).
80. Ramakrishnan, S., Nigam, A. & Chandra, G. Resistivity and magnetoresistance studies on superconducting A15 V_3 Ga, V_3 Au, and V_3 Pt compounds. *Phys. Rev. B* **34**, 6166 (1986).
81. Chen, Y., Cai, X., Wang, H., Wang, H. & Wang, H. Novel triadius-like N4 specie of iron nitride compounds under high pressure. *Sci. Rep.* **8**, 1–7 (2018).
82. Gražulis, S. et al. Crystallography open database–an open-access collection of crystal structures. *J. Appl. Crystallogr.* **42**, 726–729 (2009).
83. Choudhary, K., DeCost, B. & Tavazza, F. Machine learning with force-field-inspired descriptors for materials: Fast screening and mapping energy landscape. *Phys. Rev. Mater.* **2**, 083801 (2018).
84. Gurunathan, R., Choudhary, K. & Tavazza, F. Rapid prediction of phonon structure and properties using an atomistic line graph neural network (alignn). Preprint at https://arxiv.org/abs/2207.12510 (2022).
85. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. Theory of superconductivity. *Phys. Rev.* **108**, 1175 (1957).
86. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
87. 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).
88. Klimeš, J., Bowler, D. R. & Michaelides, A. Chemical accuracy for the van der waals density functional. *J. Phys. Condens. Matter* **22**, 022201 (2009).
89. Giannozzi, P. et al. Quantum espresso: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
90. Giannozzi, P. et al. Quantum espresso toward the exascale. *J. Chem. Phys.* **152**, 154105 (2020).

91. Kawamura, M., Gohda, Y. & Tsuneyuki, S. Improved tetrahedron method for the brillouin-zone integration applicable to response functions. *Phys. Rev. B* **89**, 094515 (2014).
92. Poncé, S., Margine, E. R., Verdi, C. & Giustino, F. Epw: Electron–phonon coupling, transport and superconducting properties using maximally localized wannier functions. *Comput. Phys. Commun.* **209**, 116–133 (2016).
93. Baroni, S., Giannozzi, P. & Testa, A. Green's-function approach to linear response in solids. *Phys. Rev. Lett.* **58**, 1861 (1987).
94. Gonze, X. Perturbation expansion of variational principles at arbitrary order. *Phys. Rev. A* **52**, 1086 (1995).
95. Garrity, K. F., Bennett, J. W., Rabe, K. M. & Vanderbilt, D. Pseudopotentials for high-throughput dft calculations. *Comput. Mater. Sci.* **81**, 446–452 (2014).
96. Perdew, J. P. et al. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
97. Allen, P. B. & Mitrović, B. Theory of superconducting tc. *Solid State Phys.* **37**, 1–92 (1983).
98. Margine, E. R. & Giustino, F. Anisotropic migdal-eliashberg theory using wannier functions. *Phys. Rev. B* **87**, 024505 (2013).
99. Marsiglio, F., Schossmann, M. & Carbotte, J. Iterative analytic continuation of the electron self-energy to the real axis. *Phys. Rev. B* **37**, 4965 (1988).
100. Ke, G. et al. Lightgbm: a highly efficient gradient boosting decision tree. Adv. Neural Inform. Process. Syst. 30 (2017).
101. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
102. Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A. & Müller, K.-R. Schnet—a deep learning architecture for molecules and materials. *J. Chem. Phys.* **148**, 241722 (2018).
103. Chen, C., Ye, W., Zuo, Y., Zheng, C. & Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chem. Mater.* **31**, 3564–3572 (2019).
104. Paszke, A. et al. Pytorch: an imperative style, high-performance deep learning library. *Adv. Neural Inform. Process Syst.* **32**, 8026–8037 (2019).
105. Wang, M. et al. Deep graph library: towards efficient and scalable deep learning on graphs. Preprint at https://arxiv.org/abs/1909.01315v2 (2019).
106. Drzazga, E., Domagalska, I., Jarosik, M., Szczeniak, R. & Kalaga, J. Characteristics of superconducting state in vanadium: the eliashberg equations and semi-analytical formulas. *J. Supercond. Nov. Magn.* **31**, 1029–1034 (2018).
107. Dew-Hughes, D. & Linse, V. Formation of superconducting nb3si by explosive compression. *J. Appl. Phys.* **50**, 3500–3504 (1979).
108. Tsindlekht, M. et al. Linear and nonlinear low-frequency electrodynamics of surface superconducting states in an yttrium hexaboride single crystal. *Phys. Rev. B* **78**, 024522 (2008).
109. Matthias, B. T., Geballe, T. H. & Compton, V. B. Superconductivity. *Rev. Mod. Phys.* **35**, 1–22 (1963).
110. Hulm, J., Jones, C., Hein, R. & Gibson, J. Superconductivity in the tio and nbo systems. *J. Low. Temp. Phys.* **7**, 291–307 (1972).
111. Pechen, E. et al. Tunneling and critical-magnetic-field study of superconducting nbc thin films. *Phys. C: Supercond.* **235**, 2511–2512 (1994).
112. Kruglov, I. A. et al. Superconductivity of lah 10 and lah 16 polyhydrides. *Phys. Rev. B* **101**, 024508 (2020).

ACKNOWLEDGEMENTS

K.C. and K.F.G. thank the National Institute of Standards and Technology for funding, computational, and data-management resources. K.C. thanks the computational support from XSEDE computational resources under allocation number TG-DMR 190095. Contributions from K.C. were supported by the financial assistance award 70NANB19H117 from the U.S. Department of Commerce, National Institute of Standards and Technology.

AUTHOR CONTRIBUTIONS

K.C. and K.F.G. jointly developed the workflow. K.C. carried out the high-throughput DFT calculations and developed machine learning workflow. K.C. and K.F.G. jointly analyzed the DFT data and contributed in writing 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/s41524-022-00933-1.

Correspondence and requests for materials should be addressed to Kamal Choudhary.

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.