Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Universal Graph Deep Learning Interatomic Potential for the Periodic Table

Chi Chen (chc273@eng.ucsd.edu)

Department of Nanoengineering, University of California San Diego, 9500 Gilman Dr, Mail Code 0448, La Jolla, CA 92093-0448, USA https://orcid.org/0000-0001-8008-7043

Shyue Ong

University of California San Diego https://orcid.org/0000-0001-5726-2587

Article

Keywords:

Posted Date: April 4th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1348213/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Nature Computational Science on November 28th, 2022. See the published version at https://doi.org/10.1038/s43588-022-00349-3.

A Universal Graph Deep Learning Interatomic Potential for the Periodic Table

Chi Chen* and Shyue Ping Ong*

Department of NanoEngineering, University of California San Diego, CA, USA

E-mail: chc273@eng.ucsd.edu; ongsp@eng.ucsd.edu

Abstract

Interatomic potentials (IAPs), which describe the potential energy surface of a collection of atoms, are a fundamental input for atomistic simulations. However, existing IAPs are either fitted to narrow chemistries or too inaccurate for general applications. Here, we report a universal IAP for materials based on graph neural networks with three-body interactions (M3GNet). The M3GNet IAP was trained on the massive database of structural relaxations performed by the Materials Project over the past 10 years and has broad applications in structural relaxation, dynamic simulations and property prediction of materials across diverse chemical spaces. About 1.8 million potentially stable materials were identified from a screening of 31 million hypothetical crystal structures, demonstrating a machine learning-accelerated pathway to the discovery of synthesizable materials with exceptional properties.

Atomistic simulations are the bedrock of *in silico* materials design. The first step in most computational studies of materials is obtaining an equilibrium structure, which involves navigating the potential energy surface (PES) across all independent lattice and atomic degrees of freedom in search of a minimum. Atomistic simulations are also used to probe the dynamical evolution of materials systems and to obtain thermodynamic averages and kinetic properties (e.g., diffusion constants). While electronic structure methods such as density functional theory (DFT) provide the most accurate description of the PES, they are computationally expensive and scale poorly with system size.

For large-scale materials studies, efficient, linear-scaling interatomic potentials (IAPs) that describe the PES in terms of many-body interactions between atoms are often necessary. However, most IAPs today are custom-fitted for a very narrow range of chemistries, often for a single element or up to no more than 4-5 elements. The most popular "general purpose" IAPs are the AMBER family of force fields[1–3] and the Universal Force Field (UFF).[4] However, both were formulated primarily for molecular/organic systems and have

limited support and accuracy in modeling crystal structures. More recently, machine learning (ML) of the PES has emerged as a particularly promising approach to IAP development.[5–9] These so-called ML-IAPs typically express the PES as a function of local environment descriptors such as the interatomic distances and angles, or atomic densities, and have been demonstrated to significantly outperform classical IAPs across a broad range of chemistries.[10] Message-passing and graph deep learning models[11–13] have also been shown to yield highly accurate predictions of energies and/or forces of molecules as well as a limited number of crystals, such as $Li_7P_3S_{11}$[14] and Li_xSi_y[15] for lithium-ion batteries. Nevertheless, no work has demonstrated a universally applicable IAP across the periodic table and for all crystal types.

In the past decade, the advent of efficient and reliable electronic structure codes[16] with high-throughput automation frameworks[17–20] have led to the development of large federated databases of computed materials data, such as the Materials Project,[21] AFLOW,[22] Open Quantum Mechanical Database (OQMD),[23] NOMAD,[24] etc. Most of the focus has been on the utilization of the final outputs from the electronic structure computations carried out by these database efforts, namely, the equilibrium structures, energies, band structures and other derivative material properties, for the purposes of materials screening and design. Less attention has been paid to the huge quantities of PES data, i.e., intermediate structures and their corresponding energies, forces, stresses, that have been amassed in the process of performing structural relaxations.

In this work, we develop the formalism for a graph-based deep learning IAP by combining physics-based many-body features of traditional IAPs with those of flexible graph material representations. Utilizing the largely untapped dataset of more than 187,000 energies, 16,000,000 forces and 1,600,000 stresses from structural relaxations performed by the Materials Project since its inception in 2011, we trained a universal material graph with three-body interactions neural network (M3GNet) IAP for 89 elements of the periodic table with low energy, force, and stress errors. We demonstrate the applications of M3GNet in

the calculations of phonon and elasticity, structural relaxations, etc. We further relaxed ∼30 millions of hypothetical structures for new materials discovery.

Materials Graphs with Many-Body Interactions

Mathematical graphs are a natural representation for crystals and molecules, with nodes and edges representing the atoms and the bonds between them, respectively. In traditional graph neural network (GNN) models for materials, information flows between the node, edge, and, optionally, state vector attributes via successive application of graph convolutional or update operations.[25–28] Typically, the input bond attribute is based on an interatomic pair distance measure, such as a Gaussian basis expansion. While such graph deep learning models have proven to be exceptionally effective for general materials property predictions,[25–28] they are not suitable as IAPs due to the lack of physical constraints such as continuity of energies and forces with changes with the length and number of bonds.

Here, we develop a new materials graph architecture that explicitly incorporates many-body interactions (Figure 1). The materials graph is represented as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{X}, [\boldsymbol{M}, \boldsymbol{u}])$, where $\boldsymbol{v}_i \in \mathcal{V}$ is atom information for i, $\boldsymbol{e}_{ij} \in \mathcal{E}$ is the bond information for bond connected by atom i and j, and $\boldsymbol{u}$ is the optional global state information, as temperature, pressure etc. A key difference with prior materials graph implementations is the addition of $\boldsymbol{x}_i \in \mathcal{X}$, the coordinates for atom i, and $\boldsymbol{M}$, the optional 3×3 lattice matrix in crystals, which are necessary for obtaining tensorial quantities such as forces and stresses via auto-differentiation.

The neighborhood of atom i is denoted as $\mathcal{N}_i$. Taking inspiration from traditional IAPs such as the Tersoff bond order potential,[29] we consider all other bonds emanating from atom i when calculating the bond interaction of $\boldsymbol{e}_{ij}$. To incorporate n-body interactions, each $\boldsymbol{e}_{ij}$ is updated using all distinct combinations of $n-2$ neighbors in the neighborhood of atom i

excluding atom j, i.e., $\mathcal{N}_i/j$, denoted generally as follows:

$$\tilde{\boldsymbol{e}}_{ij} = \sum_{\substack{k_1,k_2,...,k_{n-2}\in\mathcal{N}_i/j \\ k_1!=k_2!=...k_{n-2}}} \phi_n(\boldsymbol{e}_{ij}, \boldsymbol{r}_{ij}, \boldsymbol{v}_j, \boldsymbol{r}_{ik_1}, \boldsymbol{r}_{ik_2}, ..., \boldsymbol{r}_{ik_{n-2}}, \boldsymbol{v}_{k_1}, \boldsymbol{v}_{k_2}, ..., \boldsymbol{v}_{k_{n-2}}) \tag{1}$$

where ϕ_n is the update function and $\boldsymbol{r}_{ik}$ is the vector pointing from atom i to atom k. In practice, this n-body information exchange involves the calculation of distances, angles, dihedral angles, improper angles, etc., which escalates combinatorially with the order n as $(M_i - 1)!/(M_i - n + 1)!$ where M_i is the number of neighbors in $\mathcal{N}_i$. For brevity, we will denote this materials graph with n-body interactions neural network as MnGNet. In this work, we will focus on the incorporation of three-body interactions only, i.e., M3GNet.

Let θ_{jik} denote the angle between bonds $\boldsymbol{e}_{ij}$ and $\boldsymbol{e}_{ik}$. Here, we expand the three-body angular interactions using an efficient complete and orthogonal spherical Bessel function and spherical harmonics basis set, as proposed by Klicpera et al.[12] The bond update equation can then be rewritten as:

$$\tilde{\boldsymbol{e}}_{ij} = \sum_k j_l(z_{ln}\frac{r_{ik}}{r_c}) Y_l^0(\theta_{jik}) \odot \sigma(\boldsymbol{W}_v \boldsymbol{v}_k + \boldsymbol{b}_v) f_c(r_{ij}) f_c(r_{ik}) \tag{2}$$

$$\boldsymbol{e}'_{ij} = \boldsymbol{e}_{ij} + g(\tilde{\boldsymbol{W}}_2 \tilde{\boldsymbol{e}}_{ij} + \tilde{\boldsymbol{b}}_2) \odot \sigma(\tilde{\boldsymbol{W}}_1 \tilde{\boldsymbol{e}}_{ij} + \tilde{\boldsymbol{b}}_1) \tag{3}$$

where $\boldsymbol{W}$ and $\boldsymbol{b}$ are learnable weights from the network, j_l is the spherical Bessel function with the roots at z_{ln}, r_c is the cutoff radius, Y_l^0 is the spherical harmonics function with $m = 0$, $\odot$ is the element-wise product, σ is the sigmoid activation function, $f_c(r) = 1 - 6(r/r_c)^5 + 15(r/r_c)^4 - 10(r/r_c)^3$ is the cutoff function ensuring the functions vanishes smoothly at the neighbor boundary,[30] and $g(x) = x\sigma(x)$ is the nonlinear activation function.[31] $\tilde{\boldsymbol{e}}_{ij}$ is a vector of length $n_{max}l_{max}$ expanded by indices $l = 0, 1, ..., l_{max} - 1$ and $n = 0, 1, ..., n_{max} - 1$.

Following the n-body interaction update, several graph convolution steps are carried out sequentially to update the bond, atom and, optionally, state information, as follows:

$$\boldsymbol{e}'_{ij} = \boldsymbol{e}_{ij} + \phi_e(\boldsymbol{v}_i \oplus \boldsymbol{v}_j \oplus \boldsymbol{e}_{ij} \oplus \boldsymbol{u})\boldsymbol{W}_e^0 \boldsymbol{e}_{ij}^0 \quad (4)$$

$$\boldsymbol{v}'_i = \boldsymbol{v}_i + \sum_j \phi'_e(\boldsymbol{v}_i \oplus \boldsymbol{v}_j \oplus \boldsymbol{e}'_{ij} \oplus \boldsymbol{u})\boldsymbol{W}_e^{0\prime} \boldsymbol{e}_{ij}^0 \quad (5)$$

$$\boldsymbol{u}' = g(\boldsymbol{W}_2^u g(\boldsymbol{W}_1^u(\frac{1}{N_v}\sum_i^{N_v} \boldsymbol{v}_i \oplus \boldsymbol{u}) + \boldsymbol{b}_1^u) + \boldsymbol{b}_2^u) \quad (6)$$

where $\phi_e(x)$ and $\phi'_e(x)$ are gated multi-layer perceptrons as in Equation 9, $\oplus$ is the concatenation operator, N_v is the number of atoms, and $\boldsymbol{e}_{ij}^0$ are the distance-expanded basis functions with values, first and second derivatives smoothly go to zero at the cutoff boundary (see Methods). Such a design ensures that the target values and their derivatives up to second order change smoothly with changes in the number of bonds. $\boldsymbol{u}$ inputs and updates are optional to the models since not all structures or models have state attributes.

Each block of multi-step updates (n-body, bond, atom, state) can be repeated to construct models of arbitrary complexity, similar to previous materials graph network architectures.[26]

M3GNet Interatomic Potential

To develop an IAP using the M3GNet architecture, crystal structures with corresponding energies (E), forces ($\boldsymbol{f}$) and stresses ($\boldsymbol{\sigma}$) as targets were used as training data. The model generate trainable targets via auto-differentiation with $\boldsymbol{f} = -\partial E/\partial \boldsymbol{x}$ and $\boldsymbol{\sigma} = \partial E/V\partial \boldsymbol{\epsilon}$, where $\boldsymbol{x}$ are the atomic coordinates, V is the volume, and ϵ is the strain.

Benchmark on IAP datasets

As an initial benchmark, we selected a diverse DFT dataset of elemental energies and forces previously generated by Zuo et al.[10] for fcc Ni, fcc Cu, bcc Li, bcc Mo, diamond Si and diamond Ge. From Table 1, the M3GNet IAPs significantly outperform classical many-body potentials such as the embedded atom method (EAM) and modified EAM (MEAM)

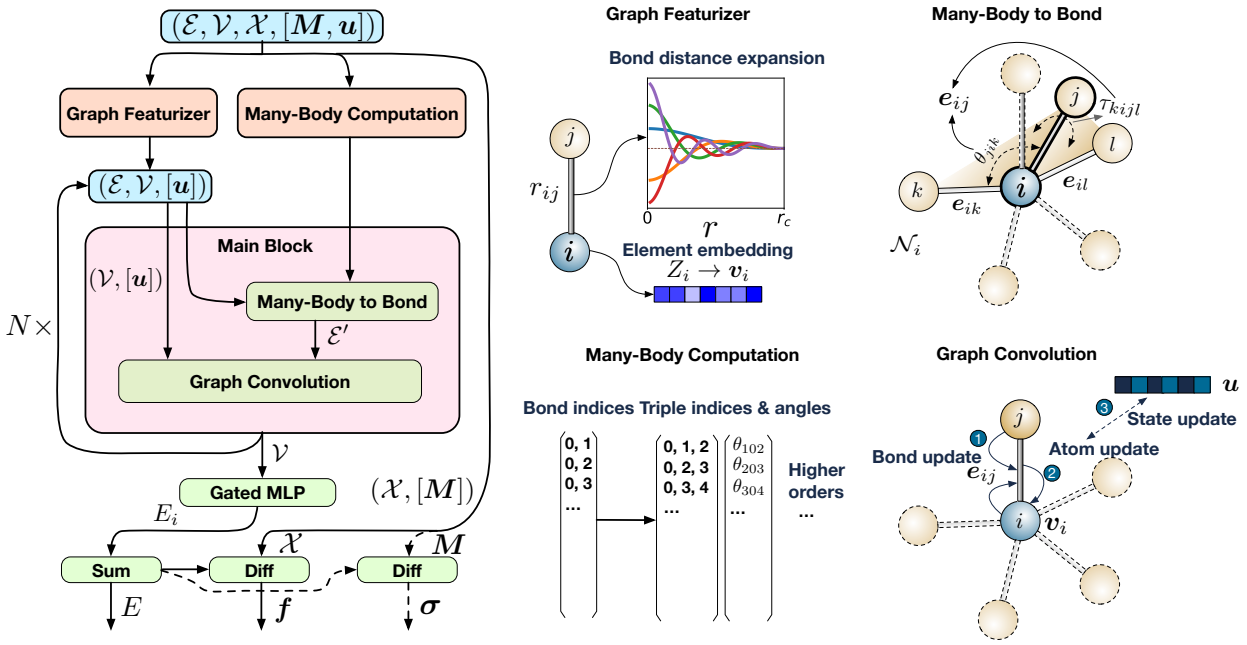

Figure 1: **Schematic of the many-body graph potential and the major computational blocks.** The model architecture starts from a position-included graph, and then goes through a featurization process, followed by main blocks, and the readout module with energy, force and stress outputs. The featurization process consists of the graph featurizer and the many-body computation module. In the graph featurizer, the atomic number of elements were embedded into a learnable continuous feature space, and the pair bond distances were expanded onto a basis set with values and derivatives up to second order going to zero at boundary. The many-body computation module calculates the three-body and many-body interaction atom indices and the associated angles. The main block consists of two main steps, namely the many-body to bond module and standard graph convolution. The many-body to bond step calculates the new bond information $\boldsymbol{e}_{ij}$ by considering the full bonding environment $\mathcal{N}_i$ of atom i via many-body angles such as θ_{jik}, τ_{kijl}, etc., and the bond length r_{ik}, r_{ij}, r_{il}, etc. The standard graph convolution updates bond, atom, and the optional state information iteratively. During the readout stage, atom information in the graph was passed to a gated MLP for obtaining atomic energy, which sums to the total energy. The derivatives of the total energy give force and stress outputs.

and performs comparably to local environment-based ML-IAPs such as the Behler-Parinello neural network potential (NNP)[5] and moment tensor potential (MTP).[8] It should be noted that while ML-IAPs can achieve slightly lower energy and force errors than M3GNet IAPs, it comes at a substantial loss in flexibility in handling multi-element chemistries. Incorporating multiple elements in ML-IAPs results in a combinatorial explosion in number of regression coefficients and corresponding data requirements. For instance, the MTP requires $\mathcal{O}(n_{ele}^2)$ regression coefficients alone to describe the element interactions, where n_{ele} is the number of elements. In contrast, the M3GNet architecture represents the elemental information for each atom (node) as a learnable embedding vector. Such a framework is readily extendable to multi-component chemistries. For instance, the M3GNet-all IAP trained on all six elements perform similarly to the M3GNet IAPs trained on individual elements. The M3GNet framework, like other GNNs, is able to capture long-range interactions without the need to increase the cutoff radius for bond construction (Figure S1). At the same time, unlike the previous GNN models, the M3GNet architecture still retains a continuous variation of energy, force and stress with changes of the number of bonds (Figure S2), a crucial requirement for IAPs.

Table 1: M3GNet models errors compared to the existing models EAM, MEAM, NNP, and MTP on the single-element dataset from Zuo et al.[10] In each cell, the errors are reported in root mean squared error (RMSE) by averaging results from three independent model training and are written in energy (10^{-3} eV atom^{-1}), force (10^{-3} eV Å^{-1}) pairs. The M3GNet-all model trains all six elements in one model.

Element	M3GNet	M3GNet-all	EAM	MEAM	NNP	MTP
Ni	0.9, 37.4	1.9, 37.0	8.5, 110	23.0, 330	2.3, 67.3	0.8, 26.9
Cu	1.8, 17.0	2.3, 16.9	7.5, 120	10.5, 240	1.7, 63.0	0.5, 13.5
Li	2.5, 22.1	4.7, 24.5	368.6, 140	-	1.0, 63.4	0.7, 13.2
Mo	6.3, 193.7	6.8, 271.4	68.0, 520	36.4, 220	5.7, 198.7	3.9, 148.1
Si	9.6, 102.8	6.8, 126.2	-	111.7, 400	9.9, 174.2	3.0, 88.1
Ge	9.4, 76.4	5.9, 78.4	-	-	11.0, 124.3	3.7, 70.3

Universal Interatomic Potential for the Periodic Table

To develop an IAP for the entire periodic table, we leveraged on one of the largest open databases of DFT crystal structure relaxations in the world - the Materials Project.[21] The Materials Project performs a sequence of two relaxation calculations[20] with the Perdew-Burke Ernzerhof (PBE)[32] generalized gradient approximation (GGA) functional or the GGA+U method[33] for every unique input crystal, typically obtained from an experimental database such as the Inorganic Crystal Structure Database (ICSD).[34] Our initial dataset comprises a sampling of the energies, forces and stresses from the first and middle ionic steps of the first relaxation and the last step of the second relaxation for calculations in the Materials Project database that contains "GGA Structure Optimization" or "GGA+U Structure Optimization" task types as of Feb 8, 2021. The snapshots that have a final energy per atom greater than 50 eV atom^{-1} or atom distance less than 0.5 Å were excluded, since those tend to be the result of errors in the initial input structure. In total, this "MPF.2021.2.8" dataset contains 187,687 ionic steps of 62,783 compounds, with 187,687 energies, 16,875,138 force components, and 1,689,183 stress components. The dataset covers an energy, force and stress range of [-28.731, 49.575] eV atom^{-1}, [-2570.567, 2552.991] eV Å^{-1} and [-5474.488, 1397.567] GPa, respectively (Figure 2a,b). While the distribution of forces is relatively symmetric, the stress data contains a slightly higher proportion of negative (compressive) stresses than positive stresses due to the well-known tendency of the PBE functional to underbind. The radial distribution function $g(r)$ (Figure 2c) shows that the dataset also spans a broad range of interatomic distances, including small distances of less than 0.6 Å that are essential for the M3GNet model to learn the repulsive forces at close distances. The dataset encompasses 89 elements of the periodic table. More information about the MPF.2021.2.8 data distribution is provided in Table S1. This dataset is then split into the training, validation and test data in the ratio of 90%, 5% and 5%, respectively. Three independent data splits were performed.

In principle, an IAP can be trained on only energies, or a combination of energies and forces. In practice, the M3GNet IAP trained only on energies (M3GNet-E) was unable to

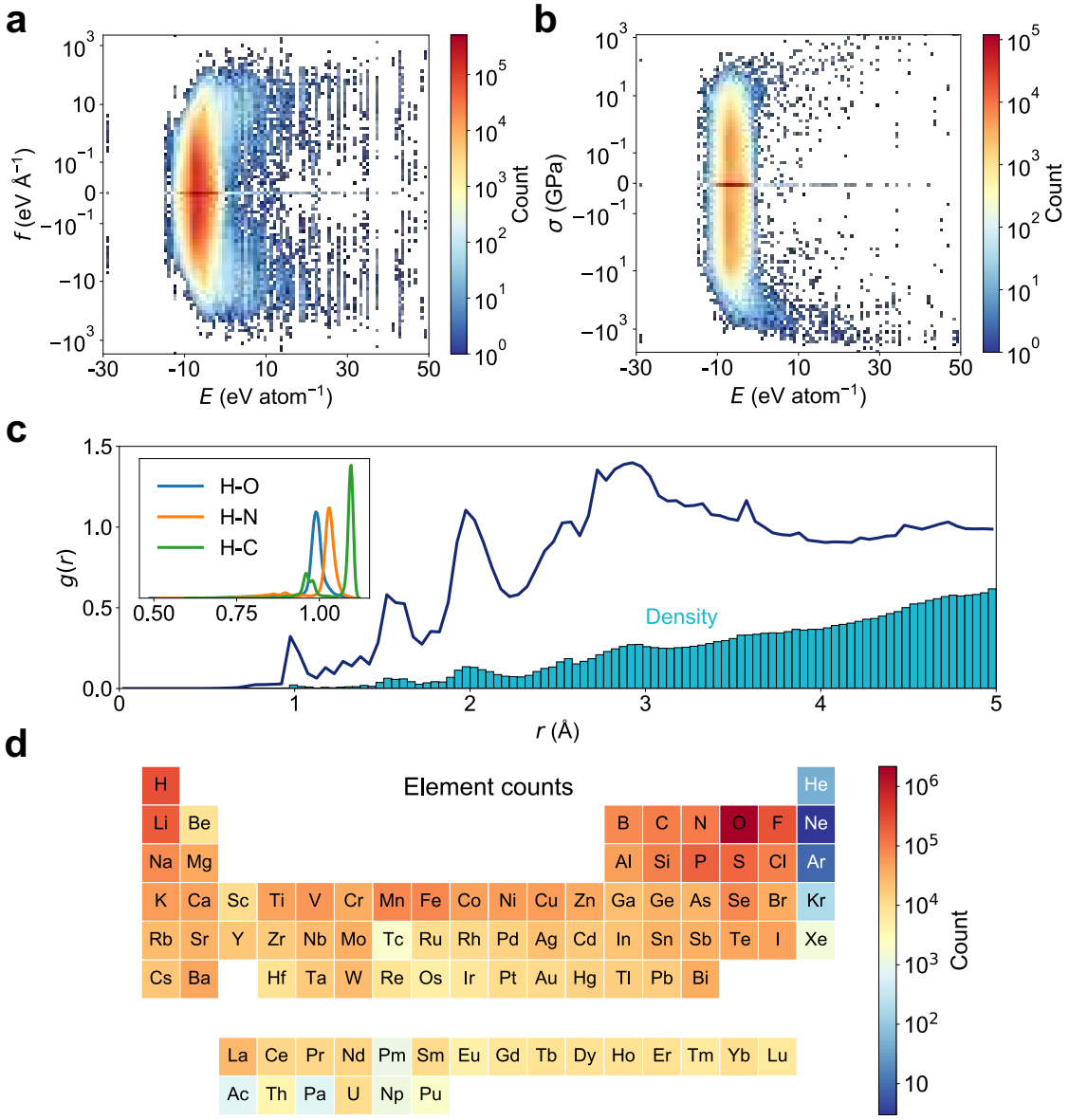

Figure 2: **The distribution of the MPF.2021.2.8 dataset.** Structural energy per atom versus force components (**a**) and stress components (**b**) distributions. **c,** The radial distribution function $g(r)$ and pair atom distance distribution density. The short distance (<1.1 Å) density is made of mostly hydrogen bonded with O, C and N, illustrated in the inset. **d,** Element counts for all atoms in the dataset, covering 89 elements across the periodic table.

achieve reasonable accuracies for predicting either forces or stresses, with mean absolute errors (MAEs) great than even the mean absolute deviation of the data (Table S2). This is the result of the amplification of errors when calculating the derivatives when only energy data is used. The M3GNet models trained with energies + forces (M3GNet-EF) and energies + forces + stresses (M3GNet-EFS) achieved relatively similar energy and force MAEs, but the MAE in stresses of the M3GNet-EFS was about half that of the M3GNet-EF model. Accurate stress predictions are necessary for applications that involve lattice changes, such as structural relaxations or NpT molecular dynamics (MD) simulations. Our results suggest that it is critical to include all three properties (energy, force, and stress) in the model training to obtain a practical IAP. The final M3GNet-EFS IAP (henceforth, referred to simple as the M3GNet model for brevity) achieved an average of 0.035 eV atom^{-1}, 0.072 eV Å^{-1}, and 0.41 GPa for energy, force, and stress test MAE, respectively.

We further investigated the test error distributions of one final M3GNet model. Generally, the model predictions and the DFT ground truth match well as revealed by the high linearity and the R^2 values for the linear fitting between DFT and model predictions (Figure 3a-c) The cumulative distribution of the model errors indicate that 50% of the data has energy, force, and stress errors lower than 0.01 eV atom^{-1}, 0.033 eV Å^{-1} and 0.042 GPa, respectively (Figure 3d-f). Even more stringent tests were carried out using phonon and elasticity calculations, which were not part of the original training data. The M3GNet model can reproduce accurate phonon dispersion curves and density of states (DOS) of β-cristobalite, stishovite, and α-quartz SiO_2 (Figure S3) to quantitative agreements with expensive DFT computations.[35] The M3GNet-predicted phonon DOS centers $\bar{\omega}$ are also in good agreement with DFT computed values with a MAE of 44.2 cm^{-1} (Figure 3g).[35] The systematic underestimation by the M3GNet model relative to DFT is likely due to the different choices of pseudopotentials; the DFT phonon calculations were performed using the PBEsol[37] functional while the M3GNet training data comprised of PBE/PBE+U calculations.[38,39] This systematic underestimation can be corrected with a constant shift of 31.6 cm^{-1} and the

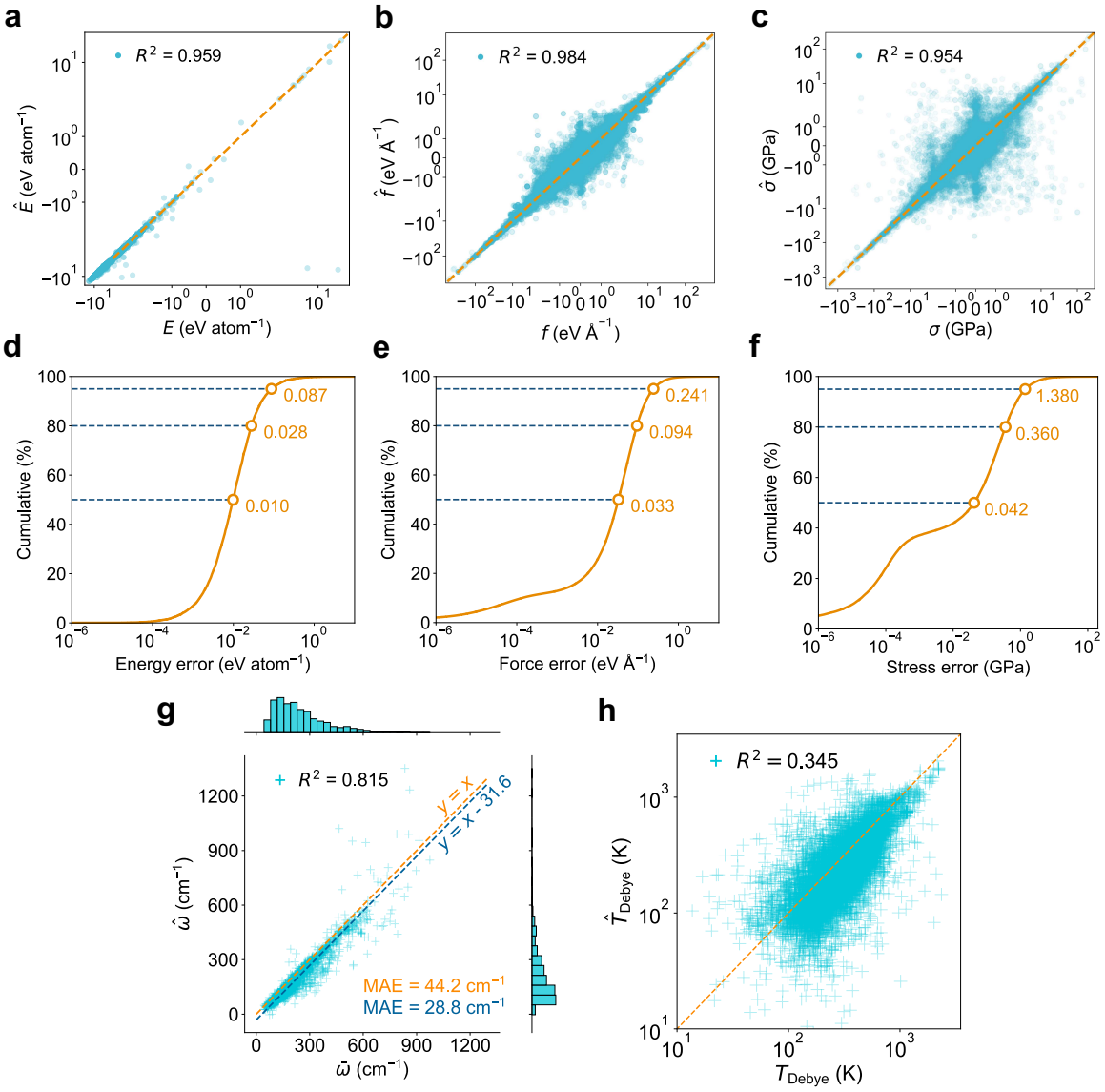

Figure 3: **The model predictions on the test dataset compared to DFT calculations. a-c**, The parity plots for energy, force and stress, respectively. The model predicted results are $\hat{E}$, $\hat{f}$, and $\hat{\sigma}$. The dashed lines are $y = x$ as guides for the eye. **d-e**, The cumulative distribution of errors for energy, force and stress, respectively. The horizontal dashed lines indicate the model errors, from bottom to top, at 50%, 80% and 95%. **g,** The comparison between model calculated 1,521 phonon density of state (DOS) center data ($\hat{\omega}$) and the PBEsol DFT calculations ($\bar{\omega}$) from Petretto et al.[35] and **h,** the 11,848 Debye temperatures (excluding negative moduli) calculated from M3GNet model ($\hat{T}_{\text{Debye}}$) and PBE DFT elastic tensors from de Jong et al.[36]

MAE reduces to 28.8 cm^{-1}. Such errors are even lower than a state-of-the-art phonon DOS peak position model which reported MAE of 36.9 cm^{-1}.[40] Similar to DFT, the relationship $\bar{\omega} \propto 1/(\overline{m})^2$, where $\overline{m}$ is the average atomic mass, is obtained (Figure S4). The M3GNet-calculated Debye temperatures are less accurate (Figure 3h), which can be attributed to relative poor M3GNet predictions of the shear moduli ($R^2 = 0.134$) (Figure S5), though the the bulk moduli predictions ($R^2 = 0.757$) are reasonable.

The M3GNet model was used to perform relaxations of the initial structures from the test dataset of 3,140 materials. M3GNet relaxation yields crystals that have volumes much closer to the DFT reference volumes (Figure 4a). While 50% and 5% of the initial input structures have volumes that differ from the final DFT relaxed crystals by more than 2.4% and 22.2%, respectively, these errors are reduced to 0.6% and 6.6% via M3GNet relaxation. Correspondingly, the errors in the predicted energies $\hat{E}$ are also much smaller (Figure 4b). Using the initial structures for direct model predictions, the energy differences distribute broadly, with considerable amount of structures having errors larger than 0.1 eV $atom^{-1}$. The overall MAE is 0.147 eV $atom^{-1}$ with $\sim$ 20% of the structures having errors greater than 0.107 eV $atom^{-1}$ (Figure 4b). These errors are far too large for reliable estimations of materials stability, given that 90% of all inorganic crystals in the ICSD has an energy above the convex hull of less than 0.067 eV $atom^{-1}$.[41] In contrast, energy calculations on the M3GNet-relaxed structures yield a MAE of 0.041 eV $atom^{-1}$ with an error distribution that is very close to energies calculated using the DFT-relaxed structures, which have MAE of 0.037 eV $atom^{-1}$ (Figure 4a). An example of M3GNet relaxation is shown in Figure S6 for $K_{57}Se_{34}$ (mp-685089), a material with one of the largest energy change during relaxation. Convergence is achieved after about 100 steps when the forces falls under 0.1 eV $Å^{-1}$. The X-ray diffraction (XRD) pattern of the M3GNet-relaxed structure also resembles the counterpart from DFT relaxation (Figure S6g). This relaxation can be performed on a laptop in about 22 seconds on a single CPU core of Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz, while the corresponding DFT relaxation took 15 hours on 32 cores in the original Materials

Project calculations.

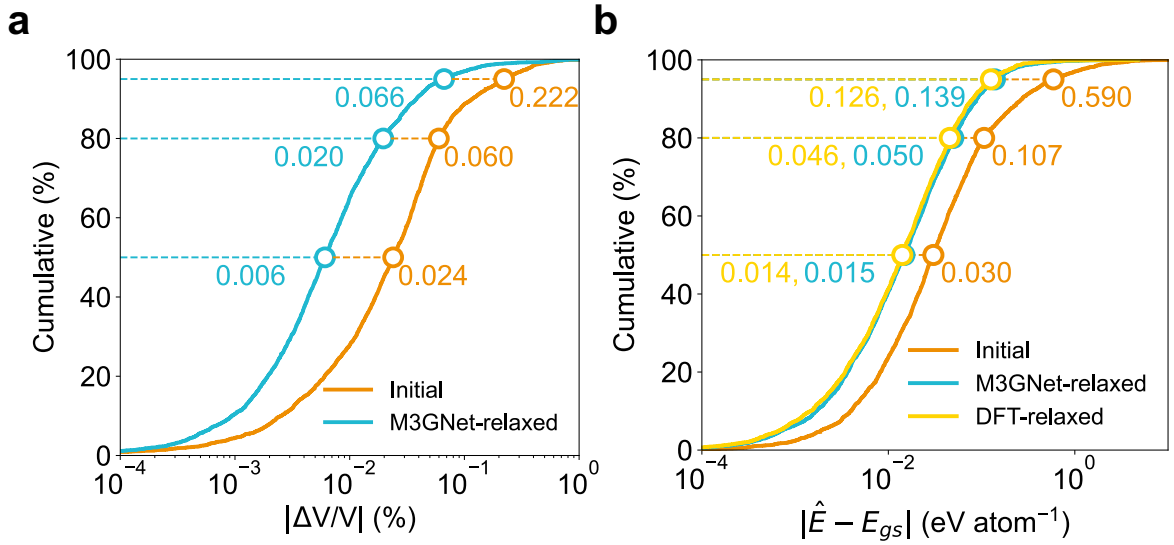

Figure 4: **Relaxation of crystal structures with M3GNet**. **a**, Distribution of the absolute percentage error in volumes of M3GNet-relaxed structures relative to DFT-relaxed structures. **b**, The differences between M3GNet-predicted energies $\hat{E}$ and ground state (gs) energies E_{gs} using the initial, M3GNet-relaxed and DFT-relaxed structures. E_{gs} is defined as the DFT energy of the DFT-relaxed crystal. The horizontal lines mark the 50th, 80th, and 95th percentiles of the distributions and the corresponding x axis values are annotated.

New Materials Discovery

The ability of M3GNet to accurately and rapidly relax arbitrary crystal structures and predict their energies makes it ideal for large-scale materials discovery. To generate hypothetical materials, combinatorial isovalent ionic substitutions based on the common oxidation states of non-noble-gas element were performed on 5,283 binary, ternary and quaternary structural prototypes in the 2019 version of the ICSD[34] database. Only prototypes with less than 51 atoms were selected for computational speed considerations. Further filtering was performed to exclude structures with non-integer or zero-charged atoms. A total of 31,664,858 hypothetical materials candidates were generated, more than 200 times the total number of unique crystals in the Materials Project today. All structures were relaxed using the M3GNet model and their signed energy distance to the Materials Project convex hull were

calculated using the M3GNet IAP-predicted energy ($E_{\mathrm{hull-m}}$). A zero or negative E_{hull} means that the material is predicted to be potentially stable compared to known materials in MP. The more negative the E_{hull}, the greater the probability that a material is likely to be stable after accounting for uncertainties in the M3GNet-predicted energies. In total, 1,849,096 materials have $E_{\mathrm{hull-m}}$ less than 0.001 eV atom^{-1}. To exclude materials with high energy uncertainties, a formation energy model based on the Matbench[40] Materials Project data was developed. Materials with a difference in the signed energy distance to the Materials Project convex hull from this model ($E_{\mathrm{hull-f}}$) and $E_{\mathrm{hull-m}}$ greater than 0.2 eV atom^{-1} were then discarded. We also excluded materials that have non-metal ions in multiple valence states, e.g., materials containing Br^+ and Br^- at the same time, etc. It is well-known that PBE overbinds single-element molecules such as O_2, S_8, Cl_2, etc. and negative anion energy corrections are applied to ionic compounds in Materials Project to offset such errors.[42] However the corrections are based mostly on composition, which may artificially over-stabilize materials with multi-valence non-metal ions.

The top-1000 lowest $E_{\mathrm{hull-m}}$ materials from any chemistry as well as the top-1000 metal oxides with elements from the first five rows (excluding Tc due to radioactivity and Rb due to high dominance) were then selected for validation via DFT relaxation and energy calculations. Only the most stable polymorphs were selected for each composition. It was found that the distribution in the DFT calculated $E_{\mathrm{hull-dft}}$ matches well with the distributions of $E_{\mathrm{hull-m}}$ (Figure 5a). For most computational materials discovery efforts, a positive threshold, typically around 0.05-0.1 eV atom^{-1}, is applied to identify synthesizable materials. This positive threshold accounts for both errors in DFT calculated energies as well as the fact that some thermodynamically meta-stable materials can be realized experimentally. Of the top-1000 materials from any chemistry, 999 were found to have a $E_{\mathrm{hull-dft}} < 0.001$ eV atom^{-1} (Figure 5b) and none of them were in the Materials Project database. For the top-1000 oxides, 579, 826, and 935 were found to be synthesizable based on $E_{\mathrm{hull-dft}}$ thresholds of 0.001, 0.05 and 0.1 eV atom^{-1}, respectively (Figure 5b). Out of the 579 DFT-stable oxides, only

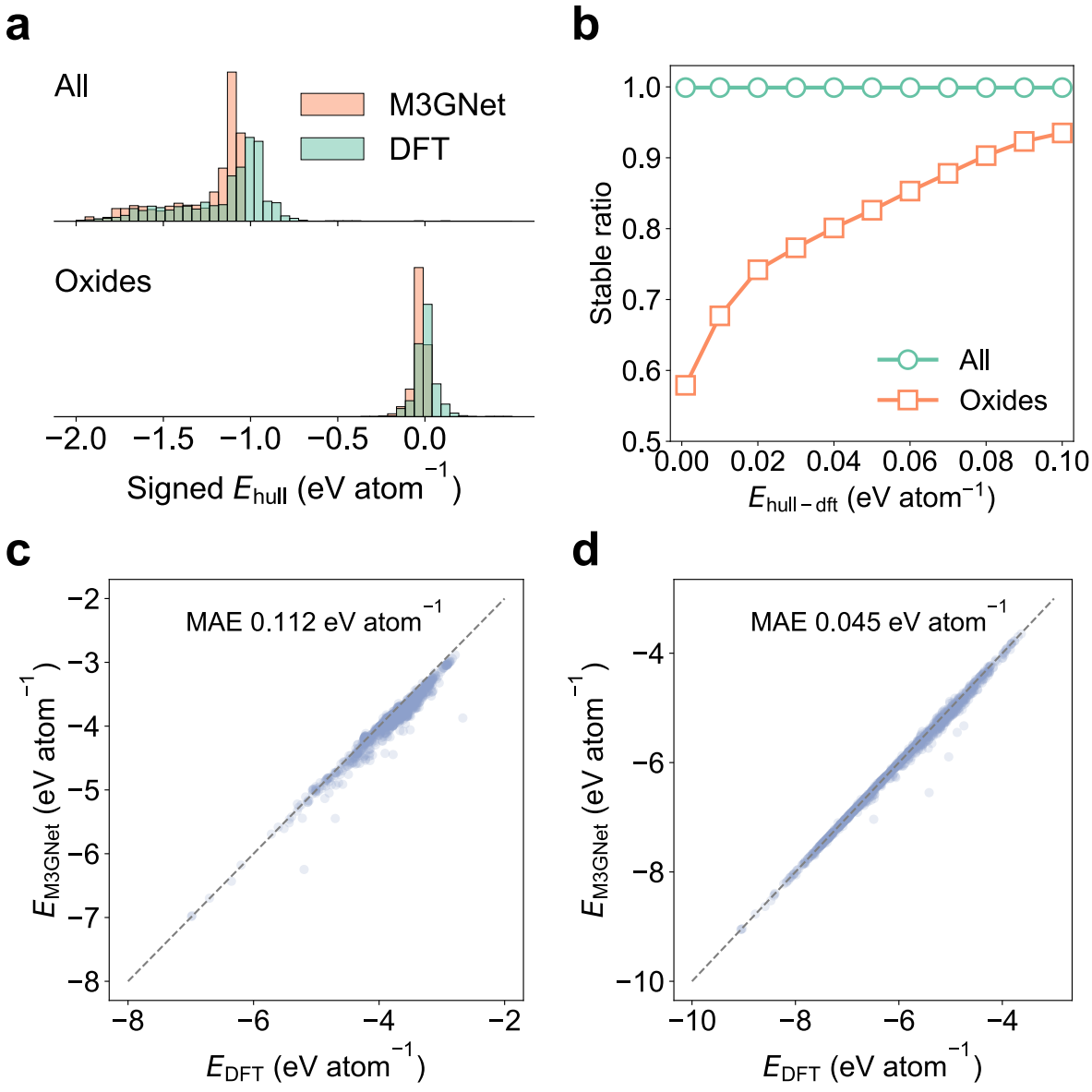

Figure 5: **Discovery of stable materials using M3GNet. a,** The signed E_{hull} distribution for the top-1000 lowest $E_{\text{hull}-\text{m}}$ materials from any chemistry (All) and oxides only (Oxides). **b,** Fraction of materials below $E_{\text{hull}-\text{dft}}$ among top-1000 materials in the All and Oxides categories. **c, d**, Plot of the final M3GNet predicted energy against final DFT energy for the (c) All and (d) Oxides categories.

five, namely $Mg_4Nb_2O_9$, $Sr_3V_2O_8$, K_2SnO_2, $Cd(RhO_2)_2$, $CoMnO_4$, were previously known and matched with the Materials Project structures. The effectiveness of the M3GNet IAP relaxations can be seen in Figures S7, which show that the energy changes during subsequent DFT relaxations (of the MEG3Net-relaxed structures) are at least one order of magnitude smaller than the energy changes during M3GNet relaxation. We also note that the final M3GNet-relaxed energies are in excellent agreement with the final DFT-relaxed energies, with MAEs of 0.112 and 0.045 eV atom^{-1} for the top 1000 materials in any chemistry and the oxide chemistry, respectively (Figures 5c-d).

Discussion

A universal IAP such as M3GNet has applications beyond crystal structure relaxation and stability predictions. For instance, a common application of IAPs is in molecular dynamics (MD) simulations to obtain transport properties such as diffusivity and ionic conductivity. The Arrhenius plot from MD simulations at multiple temperatures of the recently discovered Li superionic conductor Li_3YCl_6 is shown in Figure S8. The results agree well with the ionic conductivity and activation barriers from previous *ab initio* MD simulations.[43] Training an IAP for a complex multi-component systems such as Li_3YCl_6 is typically a highly-involved process,[44] while the M3GNet IAP can be universally applied to any material without further retraining.

It should be noted that the current M3GNet IAP reported in this work is merely the best that can be done at present with available data. Further improvements in accuracy can be achieved through several efforts. First, the training data for the M3GNet IAP comes from DFT relaxation calculations in the Materials Project, which were performed with less stringent convergence criteria such as a lower energy cutoff and sparser k-point grids. For IAP development, a best practice is to obtain accurate energies, forces and stresses via single-point, well-converged DFT calculations for training data. Building such a database is an

extensive effort that is planned for future developments in the Materials Project. Second, active learning strategies, for instance, by using the DFT relaxation data from the M3GNet-predicted stable crystals in a feedback loop, can be used to systematically improve the M3GNet IAP, especially in under-explored chemical spaces with the greatest potential for novel materials discoveries. Nevertheless, about 1.8 million of the 30 million candidates were predicted to be potentially stable or meta-stable by M3GNet against materials in the Materials Project, which already expands the potential exploration pool by an order of magnitude over the ∼140,000 crystals in the Materials Project database today.

Finally, the M3GNet framework is not limited to crystalline IAPs or even IAPs in general. The formalism of the M3GNet potential without lattice inputs and stress outputs is naturally suited for molecular force fields. When benchmarked on MD17 and MD17-CCSD(T) molecular force field data (Table S3 and S4),,[45–47] the M3GNet models were found to be more accurate than the embedded atom neural network (EANN) force field[48] and perform similarly to the state-of-the-art message-passing networks and equivariant neural network models. Moreover, by changing the readout section from summed atomic energy as in Figure 1 to intensive property readout, the M3GNet framework can be used to develop surrogate models for property prediction. We trained M3GNet models on the Matbench materials data covering nine general crystal materials properties (Table S5).[40] In all cases, the M3GNet models achieved excellent accuracies.

Acknowledgement

This work was primarily supported by the Materials Project, 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. This work used the Expanse supercomputing cluster at the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science

Foundation grant number ACI-1548562.

Author contributions

C.C. and S.P.O. conceived the idea and designed the work. C.C. implemented the models and performed the analysis. C.C. and S.P.O. wrote the manuscript and contributed to the discussion and revision.

Ethics Declaration

Competing Interests

The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to C.C. (email: chc273@eng.ucsd.edu) or S.P.O. (email: ongsp@eng.ucsd.edu).

Methods

Model construction

Neural network definition

If we denote one layer of the perceptron model as

$$\mathcal{L}_g^k : x \mapsto g(\boldsymbol{W}_k x + \boldsymbol{b}_k) \tag{7}$$

then the K-layer multi-layer perceptron (MLP) can be expressed as

$$\xi_K(x) = (\mathcal{L}_g^K \circ \mathcal{L}_g^{K-1} \circ ... \mathcal{L}_g^1)(x) \tag{8}$$

The K-layer gated MLP becomes

$$\phi_K(x) = ((\mathcal{L}_g^K \circ \mathcal{L}_g^{K-1} \circ ...\mathcal{L}_g^1)(x)) \odot ((\mathcal{L}_\sigma^K \circ \mathcal{L}_g^{K-1} \circ ...\mathcal{L}_g^1)(x)) \tag{9}$$

where $\mathcal{L}_\sigma^K(x)$ replaces the activation function $g(x)$ of $\mathcal{L}_g^K(x)$ to sigmoid function $\sigma(x)$ and $\odot$ denotes element-wise product. The gated MLP consists of the normal MLP before $\odot$ and the gate network after $\odot$.

Model architecture

Materials graphs were constructed using a radial cutoff of 5 Å. For computational efficiency considerations, the three-body interactions were limited to within a cutoff of 4 Å. The graph featurizer converts the atomic number into embeddings of dimension 64. The bond distances were expanded using the continuous and smooth basis function proposed by Kocer et al.,[49] which ensures that the first and second derivatives vanish at the cutoff radius.

$$h_m(r) = \frac{1}{\sqrt{d_m}} \left[f_m(r) + \sqrt{\frac{e_m}{d_{m-1}}} h_{m-1}(r) \right] \tag{10}$$

where

$$d_m = 1 - \frac{e_m}{d_{m-1}} \tag{11}$$

$$e_m = \frac{m^2(m+2)^2}{4(m+1)^4 + 1} \tag{12}$$

$$f_m(r) = (-1)^m \frac{\sqrt{2}\pi}{r_c^{3/2}} \frac{(m+1)(m+2)}{\sqrt{(m+1)^2 + (m+2)^2}} \left(sinc\left(r\frac{(m+1)\pi}{r_c} \right) + sinc\left(r\frac{(m+2)\pi}{r_c} \right) \right) \tag{13}$$

$$sinc(x) = \frac{\sin x}{x} \tag{14}$$

$\boldsymbol{e}_{ij}^0$ is a vector formed by m basis functions of $h(r)$.

$$\boldsymbol{e}_{ij}^0(r_{ij}) = [h_1(r_{ij}), h_2(r_{ij}), ..., h_m(r_{ij})] \tag{15}$$

In this work, we used three basis functions for the pair distance expansion.

The main blocks consist of three three-body information exchange and graph convolutions ($N = 3$ in Figure 1). By default, the $\boldsymbol{W}$'s and $\boldsymbol{b}$'s in the perceptron model gives output dimensions of 64. Each gated MLP ($\phi_e(x)$ and $\phi'_e(x)$ in Equations 4 and 5) have two layers with 64 neurons in each layer.

For the prediction of extensive properties such as total energies, a three-layer gated MLP (Equation 9) was used on the atom attributes after the graph convolution and sum the outputs as the final prediction, i.e.,

$$p_{\text{ext}} = \sum_i \phi_3(\boldsymbol{v}_i) \tag{16}$$

The gated MLP $\phi_3(x)$ has a layer neuron configuration of [64, 64, 1] and no activation in the last layer of the normal MLP part.

For the prediction of intensive properties, the readout step was performed as follows after the main blocks.

$$p_{\text{int}} = \xi_3(\sum_i w_i \xi_2(\boldsymbol{v}_i) \oplus \boldsymbol{u}) \tag{17}$$

with weights w_i summing to 1 and defined as

$$w_i = \frac{\xi'_3(\boldsymbol{v}_i)}{\sum_i \xi'_3(\boldsymbol{v}_i)} \tag{18}$$

ξ_3 and ξ'_3 have neuron configurations of [64, 64, 1] to ensure the output is scalar. For regression targets, there is no activation in the final layer of MLP, while for classification targets, the last layer activation is chosen as the sigmoid function.

In the training of MPF.2021.2.8 data, the M3GNet model contains three main blocks with 227,549 learnable weights.

Model training

The Adam optimizer[50] was used with initial learning rate of 0.001, with a cosine decay to 1% of the original value in 100 epochs. During the optimization, the validation metric values were used to monitor the model convergence, and training was stopped if the validation metric did not improve for 200 epochs. For the elemental IAP training, the loss function was the mean squared error (MSE). For other properties, the Huber loss function[51] with δ set to 0.01 was used. For the universal IAP training, the total loss function includes the loss for energy, forces, and, in inorganic compounds, also the stresses. Batch size of 32 was used in model training.

$$L = \ell(e, e_D) + w_f \ell(\boldsymbol{f}, \boldsymbol{f_D}) + w_\sigma \ell(\boldsymbol{\sigma}, \boldsymbol{\sigma_D}) \tag{19}$$

where ℓ is the Huber loss function, e is energy per atom, $\boldsymbol{f}$ is the force vector, $\boldsymbol{\sigma}$ is the stress, and w's are the scalar weights. The subscript D indicates data from DFT.

Before M3GNet IAP fitting, we fit the elemental reference energies using linear regression of the total energies. We first featurize a composition into a vector $c = [c_1, c_2, c_3, ..., c_{89}]$ where c_i is the number of atoms in the composition that has the atomic number i. The composition feature vector c is mapped to the total energy of the material E via $E = \sum_i c_i E_i$, where E_i is the reference energy for element with atomic number i that can be obtained by linear regression of the training data. Then, the elemental reference energies were subtracted from the total energies to improve M3GNet model training stability. We set $w_f = 1$ and $w_\sigma = 0.1$ during training the MPF.2021.2.8 data.

Software implementation

The M3GNet framework was implemented using the Tensorflow[52] package. All crystal and molecular structure processing were performed using the Python Materials Genomics (pymatgen)[17] package. The structural optimization was performed using the FIRE[53] algorithm

implemented in the atomic simulation environment (ASE).[54] The MD simulations were performed in the NVT ensemble using ASE.[54] Phonon calculations were performed using the Phonopy package.[55]

Data Availability

Data will be published on figshare `https://xxxx`

Code Availability

Code will be published at `https://github.com/materialsvirtuallab/m3gnet` upon paper acceptance.

References

(1) Weiner, P. K.; Kollman, P. A. AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. *Journal of Computational Chemistry* **1981**, *2*, 287–303.

(2) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *Journal of the American Chemical Society* **1984**, *106*, 765–784.

(3) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *Journal of Computational Chemistry* **2005**, *26*, 1668–1688.

(4) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *Journal of the American Chemical Society* **1992**, *114*, 10024–10035.

(5) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters* **2007**, *98*, 146401.

(6) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Physical Review Letters* **2010**, *104*, 136403.

(7) Thompson, A. P.; Swiler, L. P.; Trott, C. R.; Foiles, S. M.; Tucker, G. J. Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials. *Journal of Computational Physics* **2015**, *285*, 316–330.

(8) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Modeling & Simulation* **2016**, *14*, 1153–1173.

(9) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Physical Review Letters* **2018**, *120*, 143001.

(10) Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A. et al. Performance and Cost Assessment of Machine Learning Interatomic Potentials. *The Journal of Physical Chemistry A* **2020**, *124*, 731–745.

(11) Schütt, K. T.; Kindermans, P.-J.; Sauceda, H. E.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. SchNet: A Continuous-Filter Convolutional Neural Network for Modeling Quantum Interactions. *arXiv:1706.08566 [physics, stat]* **2017**,

(12) Klicpera, J.; Groß, J.; Günnemann, S. Directional Message Passing for Molecular Graphs. *arXiv:2003.03123 [physics, stat]* **2020**,

(13) Haghighatlari, M.; Li, J.; Guan, X.; Zhang, O.; Das, A.; Stein, C. J.; Heidar-Zadeh, F.; Liu, M.; Head-Gordon, M.; Bertels, L. et al. NewtonNet: A Newtonian Message Passing Network for Deep Learning of Interatomic Potentials and Forces. *arXiv:2108.02913 [physics]* **2021**,

(14) Park, C. W.; Kornbluth, M.; Vandermause, J.; Wolverton, C.; Kozinsky, B.; Mailoa, J. P. Accurate and Scalable Graph Neural Network Force Field and Molecular Dynamics with Direct Force Architecture. *npj Computational Materials* **2021**, *7*, 1–9.

(15) Cheon, G.; Yang, L.; McCloskey, K.; Reed, E. J.; Cubuk, E. D. Crystal Structure Search with Random Relaxations Using Graph Networks. *arXiv:2012.02920 [cond-mat, physics:physics]* **2020**,

(16) Lejaeghere, K.; Bihlmayer, G.; Björkman, T.; Blaha, P.; Blügel, S.; Blum, V.; Cal-

iste, D.; Castelli, I. E.; Clark, S. J.; Dal Corso, A. et al. Reproducibility in Density Functional Theory Calculations of Solids. *Science* **2016**, *351*, aad3000.

(17) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Computational Materials Science* **2013**, *68*, 314–319.

(18) Jain, A.; Ong, S. P.; Chen, W.; Medasani, B.; Qu, X.; Kocher, M.; Brafman, M.; Petretto, G.; Rignanese, G.-M.; Hautier, G. et al. FireWorks: A Dynamic Workflow System Designed for High-Throughput Applications. *Concurrency and Computation: Practice and Experience* **2015**, *27*, 5037–5059.

(19) Pizzi, G.; Cepellotti, A.; Sabatini, R.; Marzari, N.; Kozinsky, B. AiiDA: Automated Interactive Infrastructure and Database for Computational Science. *Computational Materials Science* **2016**, *111*, 218–230.

(20) Mathew, K.; Montoya, J. H.; Faghaninia, A.; Dwarakanath, S.; Aykol, M.; Tang, H.; Chu, I.-h.; Smidt, T.; Bocklund, B.; Horton, M. et al. Atomate: A High-Level Interface to Generate, Execute, and Analyze Computational Materials Science Workflows. *Computational Materials Science* **2017**, *139*, 140–152.

(21) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G. et al. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Materials* **2013**, *1*, 011002.

(22) Curtarolo, S.; Setyawan, W.; Wang, S.; Xue, J.; Yang, K.; Taylor, R. H.; Nelson, L. J.; Hart, G. L. W.; Sanvito, S.; Buongiorno-Nardelli, M. et al. AFLOWLIB.ORG: A Distributed Materials Properties Repository from High-Throughput Ab Initio Calculations. *Computational Materials Science* **2012**, *58*, 227–235.

(23) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies. *npj Computational Materials* **2015**, *1*, 1–15.

(24) Draxl, C.; Scheffler, M. The NOMAD Laboratory: From Data Sharing to Artificial Intelligence. *Journal of Physics: Materials* **2019**, *2*, 036001.

(25) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Physical Review Letters* **2018**, *120*, 145301.

(26) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chemistry of Materials* **2019**, *31*, 3564–3572.

(27) Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Ong, S. P. Learning Properties of Ordered and Disordered Materials from Multi-Fidelity Data. *Nature Computational Science* **2021**, *1*, 46–53.

(28) DeCost, B.; Choudhary, K. Atomistic Line Graph Neural Network for Improved Materials Property Predictions. *arXiv:2106.01829 [cond-mat]* **2021**,

(29) Tersoff, J. New Empirical Approach for the Structure and Energy of Covalent Systems. *Physical Review B* **1988**, *37*, 6991–7000.

(30) Singraber, A.; Behler, J.; Dellago, C. Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials. *Journal of Chemical Theory and Computation* **2019**, *15*, 1827–1840.

(31) Ramachandran, P.; Zoph, B.; Le, Q. V. Searching for Activation Functions. *arXiv:1710.05941 [cs]* **2017**,

(32) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865–3868.

(33) Anisimov, V. I.; Zaanen, J.; Andersen, O. K. Band Theory and Mott Insulators: Hubbard U Instead of Stoner I. *Physical Review B* **1991**, *44*, 943–954.

(34) Hellenbrandt, M. The Inorganic Crystal Structure Database (ICSD)—Present and Future. *Crystallography Reviews* **2004**, *10*, 17–22.

(35) Petretto, G.; Dwaraknath, S.; P. C. Miranda, H.; Winston, D.; Giantomassi, M.; van Setten, M. J.; Gonze, X.; Persson, K. A.; Hautier, G.; Rignanese, G.-M. High-Throughput Density-Functional Perturbation Theory Phonons for Inorganic Materials. *Scientific Data* **2018**, *5*, 180065.

(36) de Jong, M.; Chen, W.; Angsten, T.; Jain, A.; Notestine, R.; Gamst, A.; Sluiter, M.; Krishna Ande, C.; van der Zwaag, S.; Plata, J. J. et al. Charting the Complete Elastic Properties of Inorganic Crystalline Compounds. *Scientific Data* **2015**, *2*, 150009.

(37) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Physical Review Letters* **2008**, *100*, 136406.

(38) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Physical Review B* **1993**, *47*, 558–561.

(39) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Computational Materials Science* **1996**, *6*, 15–50.

(40) Dunn, A.; Wang, Q.; Ganose, A.; Dopp, D.; Jain, A. Benchmarking Materials Property Prediction Methods: The Matbench Test Set and Automatminer Reference Algorithm. *npj Computational Materials* **2020**, *6*, 1–10.

(41) Sun, W.; Dacek, S. T.; Ong, S. P.; Hautier, G.; Jain, A.; Richards, W. D.; Gamst, A. C.; Persson, K. A.; Ceder, G. The Thermodynamic Scale of Inorganic Crystalline Metastability. *Science Advances* **2016**,

(42) Wang, L.; Maxisch, T.; Ceder, G. Oxidation Energies of Transition Metal Oxides within the $\mathrm{GGA}+\mathrm{U}$ Framework. *Physical Review B* **2006**, *73*, 195107.

(43) Wang, S.; Bai, Q.; Nolan, A. M.; Liu, Y.; Gong, S.; Sun, Q.; Mo, Y. Lithium Chlorides and Bromides as Promising Solid-State Chemistries for Fast Ion Conductors with Good Electrochemical Stability. *Angewandte Chemie International Edition* **2019**, *58*, 8039–8043.

(44) Qi, J.; Banerjee, S.; Zuo, Y.; Chen, C.; Zhu, Z.; Holekevi Chandrappa, M. L.; Li, X.; Ong, S. P. Bridging the Gap between Simulated and Experimental Ionic Conductivities in Lithium Superionic Conductors. *Materials Today Physics* **2021**, *21*, 100463.

(45) Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R. Machine Learning of Accurate Energy-Conserving Molecular Force Fields. *Science Advances* **2017**, *3*, e1603015.

(46) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-Chemical Insights from Deep Tensor Neural Networks. *Nature Communications* **2017**, *8*, 13890.

(47) Chmiela, S.; Sauceda, H. E.; Müller, K.-R.; Tkatchenko, A. Towards Exact Molecular Dynamics Simulations with Machine-Learned Force Fields. *Nature Communications* **2018**, *9*, 3887.

(48) Zhang, Y.; Hu, C.; Jiang, B. Embedded Atom Neural Network Potentials: Efficient and Accurate Machine Learning with a Physically Inspired Representation. *The Journal of Physical Chemistry Letters* **2019**, *10*, 4962–4967.

(49) Kocer, E.; Mason, J. K.; Erturk, H. A Novel Approach to Describe Chemical Environments in High-Dimensional Neural Network Potentials. *The Journal of Chemical Physics* **2019**, *150*, 154102.

(50) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]* **2017**,

(51) Huber, P. J. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics* **1964**, *35*, 73–101.

(52) Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M. et al. TensorFlow: A System for Large-Scale Machine Learning. 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16). 2016; pp 265–283.

(53) Bitzek, E.; Koskinen, P.; Gähler, F.; Moseler, M.; Gumbsch, P. Structural Relaxation Made Simple. *Physical Review Letters* **2006**, *97*, 170201.

(54) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Du\lak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C. et al. The Atomic Simulation Environment—a Python Library for Working with Atoms. *Journal of Physics: Condensed Matter* **2017**, *29*, 273002.

(55) Togo, A.; Tanaka, I. First Principles Phonon Calculations in Materials Science. *Scripta Materialia* **2015**, *108*, 1–5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.pdf
- ChiChenCSflat.pdf