Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information: Exploring New Approaches towards the Formability of Mixed-Ion Perovskite by DFT and Machine Learning

Heesoo Park,[*,†] Raghvendra Mall,[‡] Fahhad H Alharbi,[†,§] Stefano Sanvito,[¶] Nouar Tabet,[†,§] Halima Bensmail,[‡] and Fedwa El-Mellouhi[*,†]

†*Qatar Environment and Energy Research Institute, Hamad Bin Khalifa University, PO BOX 34110, Doha, Qatar*

‡*Qatar Computing Research Institue, Hamad Bin Khalifa University, Doha,Qatar*

¶*School of Physics, AMBER and CRANN Institute, Trinity College, Dublin 2, Ireland*

§*College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar*

E-mail: hpark@hbku.edu.qa; felmellouhi@hbku.edu.qa

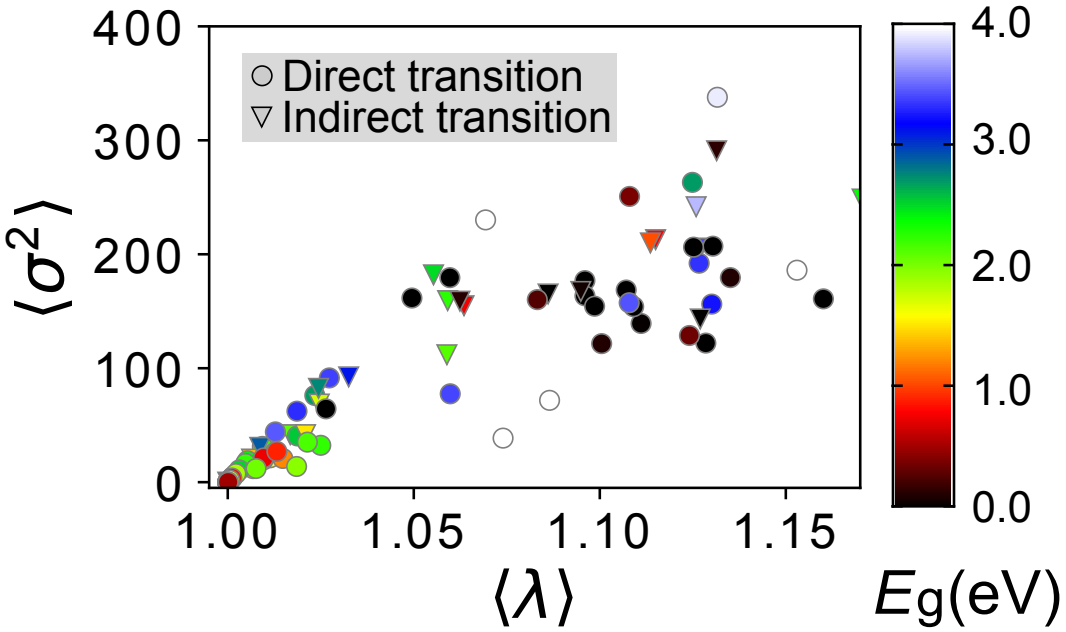

Figure S1: Octahedral distortion ($\langle\lambda\rangle$, $\langle\sigma^2\rangle$) and bandgaps (E_g) of the relaxed ABC_3 perovskites. The compounds are presented as a function of the A site and BC_3 group ions. The colour of the symbols describes the bandgaps: from 0.0=black to 4.0 eV=white. Circles (upside-down triangles) indicate that the corresponding compound has a direct (indirect) bandgap. The bandgaps are calculated at the PBE level.

Table S1: Electronic, structural and thermodynamic properties of all the compounds investigated. In the table we report: bandgap; E_g (eV), direct bandgap; E_{Dg} (eV), octahedral distortion parameters ($\langle\lambda\rangle$, $\langle\sigma^2\rangle$ and $\langle\theta^2\rangle$), lattice parameters (a, b, c, α, β and γ), and energy above the convex hull (per atom in meV).

	PBE		GLLB-sc+SOC				Lattice parameters						
	E_g	E_{Dg}	E_g	$\langle\lambda\rangle$	$\langle\sigma^2\rangle$	$\langle\theta^2\rangle$	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	ΔE_{CH}
$LiVO_3$	2.134	2.141	3.554	1.4437	436.715	14.901	10.0	9.6	8.3	85.4	88.2	86.8	-256
$NaVO_3$	2.394	2.449	4.124	1.1705	248.477	8.096	9.0	6.6	8.8	90.0	89.5	90.0	-265
KVO_3	2.652	2.652	4.460	1.1935	252.205	1.015	9.3	9.7	6.6	90.0	90.0	89.9	-238
$RbVO_3$	2.604	2.605	4.451	1.2255	261.804	0.257	6.7	10.1	9.8	90.4	90.0	90.0	-228
$CsVO_3$	1.663	1.664	3.305	NaN	NaN	NaN	9.1	9.2	12.2	84.3	104.9	98.1	-306
$LiVS_3$	0.533	0.546	0.904	1.3839	636.069	13.779	11.8	9.3	10.4	78.2	86.9	78.4	160
$NaVS_3$	0.104	0.104	0.462	1.1351	179.503	10.413	12.2	8.2	8.2	90.0	90.0	90.0	92
KVS_3	0.832	0.833	1.432	1.2184	293.506	7.057	10.2	13.0	8.1	91.4	92.1	94.1	161
$RbVS_3$	0.880	0.880	1.572	1.3095	491.377	4.930	12.1	11.7	7.8	90.0	90.0	90.1	133
$CsVS_3$	0.836	0.874	1.616	1.2421	325.126	0.253	7.9	12.3	12.2	89.8	90.0	90.0	21
$LiVSe_3$	0.000	0.000	0.203	NaN	NaN	NaN	11.4	10.3	9.6	87.9	85.4	97.8	171
$NaVSe_3$	0.000	0.000	0.000	1.1304	207.090	9.676	12.6	8.7	8.8	90.0	90.0	90.0	56
$KVSe_3$	0.000	0.000	0.006	1.1601	160.866	9.589	8.9	13.6	8.9	90.6	90.0	89.9	59

$RbVSe_3$	0.009	0.009	0.156	NaN	NaN	NaN	8.9	8.9	15.6	89.8	89.8	90.2	65
$CsVSe_3$	0.026	0.028	0.000	1.2697	432.880	2.460	8.3	12.4	13.1	98.1	90.4	90.2	46
$LiVTe_3$	0.000	0.000	0.000	1.0264	64.346	11.601	10.9	9.8	9.8	90.5	86.2	92.2	197
$NaVTe_3$	0.000	0.000	0.000	1.1252	206.237	9.432	13.6	9.4	9.5	90.0	90.2	89.5	123
$KVTe_3$	0.000	0.000	0.238	1.1071	168.611	14.432	9.6	9.7	15.3	88.3	94.2	89.4	152
$RbVTe_3$	0.000	0.000	0.000	1.0986	154.405	9.703	9.8	9.9	13.4	88.9	91.4	90.2	205
$CsVTe_3$	0.000	0.000	0.000	NaN	NaN	NaN	9.7	18.6	9.4	90.0	90.0	90.0	88
$LiNbO_3$	3.136	3.150	5.558	1.0325	92.382	17.534	8.7	10.6	8.3	88.1	75.6	95.0	136
$NaNbO_3$	2.208	2.208	4.401	1.0069	16.542	6.618	7.9	7.9	8.0	90.0	90.8	90.0	3
$KNbO_3$	1.841	1.841	3.951	1.0095	19.426	0.076	8.1	8.2	8.2	90.0	90.0	90.0	29
$RbNbO_3$	2.170	2.170	4.339	1.0214	35.150	0.043	8.5	8.1	8.6	90.0	90.0	90.0	82
$CsNbO_3$	2.317	2.317	4.495	1.3030	334.727	7.882	8.1	9.9	9.9	84.6	90.0	90.0	52
$LiNbS_3$	0.971	0.972	1.891	1.1994	307.621	17.410	13.0	9.8	10.0	102.8	71.0	86.6	144
$NaNbS_3$	0.004	0.018	0.021	1.0861	165.131	9.521	11.7	8.8	8.8	90.0	90.0	90.1	55
$KNbS_3$	0.032	0.032	0.101	1.1111	139.292	8.835	9.0	12.6	9.0	90.0	90.0	90.0	88
$RbNbS_3$	0.000	0.000	0.094	1.1285	122.055	8.523	9.1	13.1	9.1	89.7	90.0	90.2	103
$CsNbS_3$	1.557	1.557	2.815	1.2160	341.534	0.430	8.3	12.2	12.5	90.0	90.0	90.0	1
$LiNbSe_3$	0.812	0.927	1.489	1.1150	212.801	11.045	12.0	10.7	9.8	90.1	87.4	81.1	176
$NaNbSe_3$	0.000	0.000	0.000	1.0960	176.878	9.832	9.1	12.5	9.3	89.9	90.0	90.0	62
$KNbSe_3$	0.003	0.003	0.000	1.1091	153.995	9.326	13.1	9.4	9.4	90.0	89.9	90.0	88
$RbNbSe_3$	0.003	0.004	0.000	1.1270	143.223	9.092	9.4	13.7	9.5	89.9	90.0	89.9	27
$CsNbSe_3$	1.198	1.198	2.119	1.2964	525.683	4.947	13.1	12.7	8.6	90.0	90.0	90.1	0
$LiNbTe_3$	0.000	0.000	0.000	1.0013	4.497	13.459	10.2	10.2	10.2	89.9	90.0	90.1	162
$NaNbTe_3$	0.000	0.000	0.038	1.0597	179.494	16.684	11.8	11.8	10.0	90.5	106.6	87.5	209
$KNbTe_3$	0.129	0.188	0.319	1.0624	158.843	17.444	11.3	12.4	12.0	101.1	103.1	77.7	136
$RbNbTe_3$	0.151	0.170	0.458	1.1314	290.795	14.538	12.1	12.9	12.4	69.8	111.4	97.8	142
$CsNbTe_3$	0.251	0.251	0.472	NaN	NaN	NaN	11.9	9.7	14.9	98.7	89.1	86.6	119
$LiTaO_3$	3.409	3.440	5.729	1.1270	205.430	7.394	8.5	7.8	9.5	72.9	96.7	75.3	130
$NaTaO_3$	2.553	2.554	4.747	1.0000	0.034	5.791	7.9	7.9	7.9	90.0	90.0	90.0	0
$KTaO_3$	2.106	2.106	4.145	1.0001	0.324	0.006	8.1	8.1	8.1	90.0	90.0	90.0	31
$RbTaO_3$	2.053	2.053	4.164	1.0076	11.970	0.010	8.1	8.1	8.4	90.0	90.0	90.0	84
$CsTaO_3$	3.093	3.095	5.438	1.2095	216.186	8.489	10.1	10.4	7.6	90.0	90.0	98.8	100
$LiTaS_3$	1.726	1.745	2.875	1.0246	68.555	13.507	10.3	11.1	10.4	92.4	87.8	100.1	118
$NaTaS_3$	0.278	0.278	0.551	1.0832	160.196	8.775	8.9	8.9	11.6	90.0	90.0	90.0	13
$KTaS_3$	0.116	0.116	0.513	1.1005	121.585	8.231	9.0	9.0	12.4	90.1	90.0	90.0	87
$RbTaS_3$	0.347	0.347	0.780	1.1241	128.715	7.535	9.2	12.9	9.1	90.0	90.0	90.0	97
$CsTaS_3$	1.159	1.164	2.141	1.2425	359.515	9.160	15.2	11.1	9.5	76.5	66.6	70.0	0
$LiTaSe_3$	1.056	1.081	1.774	1.1136	209.946	21.097	10.5	12.1	10.3	82.8	86.9	77.7	143
$NaTaSe_3$	0.775	0.791	1.323	1.0635	154.812	15.327	10.7	11.1	9.4	93.9	88.6	94.8	113
$KTaSe_3$	1.063	1.093	1.792	1.2814	509.647	15.318	10.3	12.5	11.5	78.3	102.4	97.3	49
$RbTaSe_3$	0.061	0.062	0.287	1.0949	167.374	13.730	15.8	9.5	9.3	91.1	106.4	89.3	2

$CsTaSe_3$	0.998	1.001	-	1.2783	366.410	13.321	11.7	12.6	14.2	72.6	84.2	106.8	99
$LiTaTe_3$	0.000	0.000	0.000	1.0014	4.753	13.144	10.3	10.3	10.1	90.1	90.0	90.0	169
$NaTaTe_3$	0.000	0.000	0.038	1.0495	161.592	18.437	11.2	9.3	12.0	86.5	98.4	94.1	227
$KTaTe_3$	0.000	0.000	0.012	1.0960	163.541	8.802	10.2	10.1	13.9	91.0	88.0	90.8	209
$RbTaTe_3$	0.396	0.396	0.680	1.1080	250.752	16.743	9.8	12.7	13.5	84.5	94.1	82.6	167
$CsTaTe_3$	0.427	0.427	0.729	1.4733	530.823	17.410	14.7	9.6	12.2	91.6	103.0	95.6	155
$LiGeF_3$	4.582	4.665	7.417	NaN	NaN	NaN	9.1	8.0	10.0	88.2	109.4	104.4	77
$NaGeF_3$	4.203	4.203	7.191	1.2142	282.365	15.629	8.9	8.9	9.0	90.8	93.2	88.1	111
$KGeF_3$	4.375	4.375	7.588	1.1530	186.018	10.993	9.8	9.4	9.4	93.4	93.1	90.5	106
$RbGeF_3$	4.185	4.185	7.223	1.0865	71.819	9.037	9.5	9.5	9.7	92.7	92.6	90.9	108
$CsGeF_3$	4.227	4.227	7.119	1.0740	38.664	7.277	9.7	9.8	9.8	92.0	89.4	90.7	95
$LiGeCl_3$	3.839	3.865	6.028	1.1825	318.496	16.490	11.1	10.9	10.5	87.7	91.4	90.1	28
$NaGeCl_3$	3.765	3.767	5.823	1.1259	241.698	13.935	10.2	10.4	10.4	99.0	90.1	89.9	-1
$KGeCl_3$	0.848	0.848	1.455	1.0000	0.046	0.087	10.5	10.5	10.5	90.0	90.0	90.0	72
$RbGeCl_3$	0.927	0.927	1.583	1.0000	0.020	0.025	10.6	10.6	10.6	90.0	90.0	90.0	66
$CsGeCl_3$	1.963	1.963	3.124	1.0185	13.763	0.101	11.0	11.0	11.0	90.9	89.1	90.9	22
$LiGeBr_3$	2.136	2.149	3.501	1.0589	111.808	11.153	11.1	11.0	11.2	87.6	88.5	92.7	71
$NaGeBr_3$	2.252	2.252	3.624	1.0250	32.344	10.714	11.0	11.1	10.9	93.8	85.8	94.2	21
$KGeBr_3$	0.565	0.565	1.027	1.0000	0.035	0.066	11.1	11.1	11.1	90.0	90.0	90.0	65
$RbGeBr_3$	0.631	0.631	1.131	1.0000	0.018	0.023	11.1	11.1	11.1	90.0	90.0	90.0	51
$CsGeBr_3$	1.233	1.233	2.101	1.0148	20.817	0.156	11.4	11.4	11.4	89.0	89.0	88.9	18
$LiGeI_3$	1.616	1.616	2.425	1.0114	21.411	11.381	11.3	11.3	11.2	96.3	83.6	96.3	68
$NaGeI_3$	1.382	1.382	2.031	1.0068	11.908	11.699	11.5	11.3	11.3	93.8	90.0	89.9	67
$KGeI_3$	1.525	1.541	2.213	1.0209	41.703	5.789	12.1	12.1	12.1	93.3	93.2	86.7	62
$RbGeI_3$	0.756	0.756	1.087	1.0096	21.191	0.132	12.0	12.0	12.0	90.9	89.1	90.9	52
$CsGeI_3$	0.900	0.900	1.322	1.0132	26.972	0.102	12.1	12.1	12.1	91.0	88.9	91.0	24
$LiSnF_3$	3.471	3.548	5.846	NaN	NaN	NaN	9.4	8.8	8.7	101.7	73.2	105.5	108
$NaSnF_3$	3.349	3.349	5.834	1.1267	192.171	17.163	8.5	9.8	8.6	90.2	90.6	90.5	98
$KSnF_3$	3.256	3.256	5.634	1.1301	156.259	11.218	10.8	8.9	8.6	89.9	90.1	90.3	113
$RbSnF_3$	3.440	3.440	5.998	1.1079	157.440	8.993	9.0	9.6	10.6	91.0	89.4	91.0	90
$CsSnF_3$	3.413	3.413	5.867	1.0598	77.550	5.563	9.6	10.5	9.5	89.1	90.3	92.4	70
$LiSnCl_3$	3.376	3.376	5.320	1.1882	312.271	12.556	11.1	11.0	11.0	82.0	95.5	83.1	59
$NaSnCl_3$	2.572	2.572	3.981	1.0187	40.586	11.135	10.8	10.7	10.7	84.8	94.7	94.5	38
$KSnCl_3$	2.095	2.095	3.201	1.0127	26.038	9.685	11.0	11.0	11.0	89.9	89.8	90.0	41
$RbSnCl_3$	1.008	1.008	1.552	1.0002	0.051	3.585	11.1	11.1	11.2	90.1	90.2	90.0	55
$CsSnCl_3$	0.962	0.962	1.444	1.0000	0.071	0.015	11.2	11.2	11.2	90.0	90.0	90.0	32
$LiSnBr_3$	2.700	2.700	4.285	1.1249	263.104	15.848	11.0	10.9	11.1	89.2	84.0	92.0	59
$NaSnBr_3$	2.104	2.126	3.247	1.0168	41.685	12.201	11.4	11.4	11.0	90.0	90.0	88.7	48
$KSnBr_3$	1.298	1.298	2.033	1.0003	0.963	9.724	11.4	11.4	11.4	90.0	93.5	90.0	27
$RbSnBr_3$	0.641	0.641	1.006	1.0002	0.062	3.498	11.6	11.6	11.7	90.0	89.9	90.0	49
$CsSnBr_3$	0.603	0.603	0.687	1.0000	0.030	0.030	11.7	11.7	11.7	90.0	90.0	90.0	27

$LiSnI_3$	2.273	2.312	3.323	1.0591	159.299	15.894	11.3	11.4	11.4	84.8	90.0	89.9	56
$NaSnI_3$	1.380	1.385	1.898	1.0064	20.285	11.615	12.0	12.0	11.8	90.1	90.4	89.0	59
$KSnI_3$	1.056	1.056	1.455	1.0001	0.329	9.913	12.1	12.1	12.1	90.0	86.6	90.0	33
$RbSnI_3$	0.821	0.821	1.039	1.0011	3.633	7.312	12.5	12.3	12.3	90.0	90.0	90.0	52
$CsSnI_3$	0.508	0.508	0.457	1.0001	0.042	2.612	12.6	12.5	12.5	90.0	90.3	90.0	37
$LiPbF_3$	3.927	3.927	6.362	1.1316	337.781	12.902	8.7	8.8	8.8	87.2	86.2	75.0	113
$NaPbF_3$	4.293	4.293	6.695	1.0693	230.153	14.966	8.7	8.7	8.6	90.1	90.1	86.6	64
$KPbF_3$	3.460	3.460	5.493	1.0128	44.102	9.211	9.3	9.4	9.4	90.1	89.9	89.9	71
$RbPbF_3$	3.443	3.443	5.023	1.0010	3.445	8.001	9.5	9.5	9.5	90.0	90.0	88.8	59
$CsPbF_3$	3.021	3.021	4.520	1.0002	0.695	4.505	9.7	9.7	9.7	90.1	90.3	90.1	45
$LiPbCl_3$	3.388	3.388	4.587	1.0273	91.476	11.848	10.4	10.5	10.5	97.7	98.5	82.1	36
$NaPbCl_3$	3.329	3.329	4.259	1.0186	62.300	11.866	10.8	10.8	10.7	89.6	86.1	89.9	47
$KPbCl_3$	2.917	2.926	3.546	1.0086	30.553	10.058	11.0	11.0	11.1	90.0	90.0	90.1	35
$RbPbCl_3$	2.523	2.523	3.072	1.0031	10.773	7.315	11.1	11.4	11.1	90.0	91.0	90.0	38
$CsPbCl_3$	2.243	2.243	2.589	1.0001	0.024	3.232	11.3	11.5	11.3	89.8	90.0	90.0	18
$LiPbBr_3$	2.712	2.712	3.592	1.0234	76.094	12.083	11.1	11.4	11.2	87.7	99.3	90.1	52
$NaPbBr_3$	2.770	2.788	3.527	1.0244	82.330	12.086	11.1	11.2	11.3	88.5	90.6	91.3	49
$KPbBr_3$	2.429	2.429	2.876	1.0053	18.789	9.901	11.7	11.5	11.6	90.4	90.0	90.0	33
$RbPbBr_3$	2.255	2.255	2.603	1.0048	16.424	8.324	11.6	11.8	11.7	90.0	90.0	89.9	39
$CsPbBr_3$	2.062	2.062	2.360	1.0002	0.680	6.589	11.9	11.9	11.8	91.6	90.0	90.0	17
$LiPbI_3$	2.503	2.619	3.286	1.0553	181.928	15.832	11.5	11.2	11.5	90.1	94.1	89.9	77
$NaPbI_3$	2.234	2.234	2.519	1.0093	31.375	11.887	12.0	12.1	12.2	89.8	90.6	89.2	80
$KPbI_3$	1.966	1.966	2.170	1.0005	1.769	10.769	12.2	12.3	12.3	86.2	90.1	90.0	53
$RbPbI_3$	1.872	1.872	1.855	1.0022	7.499	8.486	12.6	12.5	12.4	90.0	90.0	90.0	55
$CsPbI_3$	1.793	1.793	1.753	1.0005	1.860	7.260	12.7	12.6	12.6	87.2	90.0	90.0	36

Table S2: Calculated bandgaps and octahedral distortion parameters for the hybrid perovskites investigated. For each composition three different inequivalent molecule orientations have been studied.

Compound	E_g	$\langle\lambda\rangle$	$\langle\sigma^2\rangle$	$\langle\theta^2\rangle$	Volume (Å^3)
MA-PbI_3 (1)	1.59	1.003	8.779	4.340	2100
MA-PbI_3 (2)	1.67	1.003	11.503	3.960	2102
MA-PbI_3 (3)	1.74	1.006	19.203	4.372	2087
FA-PbI_3 (1)	1.62	1.003	11.207	2.077	2166
FA-PbI_3 (2)	1.66	1.003	10.889	1.937	2167
FA-PbI_3 (3)	1.67	1.005	16.772	1.846	2165
Gau-PbI_3 (1)	1.80	1.007	14.651	0.590	2305
Gua-PbI_3 (2)	1.84	1.010	18.966	1.134	2331
Gua-PbI_3 (3)	1.87	1.009	15.006	0.885	2350

Table S3: Five runs of 5-fold cross-validation of random forests. Mean absolute error (MAE), square root of mean squared error (RMSE), and R^2.

Response variable	Models	MAE	RMSE	R^2
$\langle\lambda\rangle$	l-1	0.0478	0.0706	0.47
	l-2	0.0504	0.0771	0.40
	l-3	0.0498	0.0767	0.43
	l-4	0.0455	0.0681	0.53
	l-5	0.0516	0.0831	0.39
$\langle\sigma^2\rangle$	s-1	65.6	96.6	0.52
	s-2	69.1	102.7	0.42
	s-3	66.3	96.1	0.56
	s-4	62.8	92.9	0.53
	s-5	64.0	95.3	0.54
$\langle\theta^2\rangle$	t-1	2.94	3.54	0.52
	t-2	2.98	3.55	0.53
	t-3	2.88	3.48	0.49
	t-4	2.82	3.29	0.54
	t-5	2.85	3.44	0.56
E_{g}	e-1	0.317	0.397	0.89
	e-2	0.317	0.398	0.90
	e-3	0.323	0.417	0.90
	e-4	0.335	0.410	0.90
	e-5	0.306	0.391	0.91

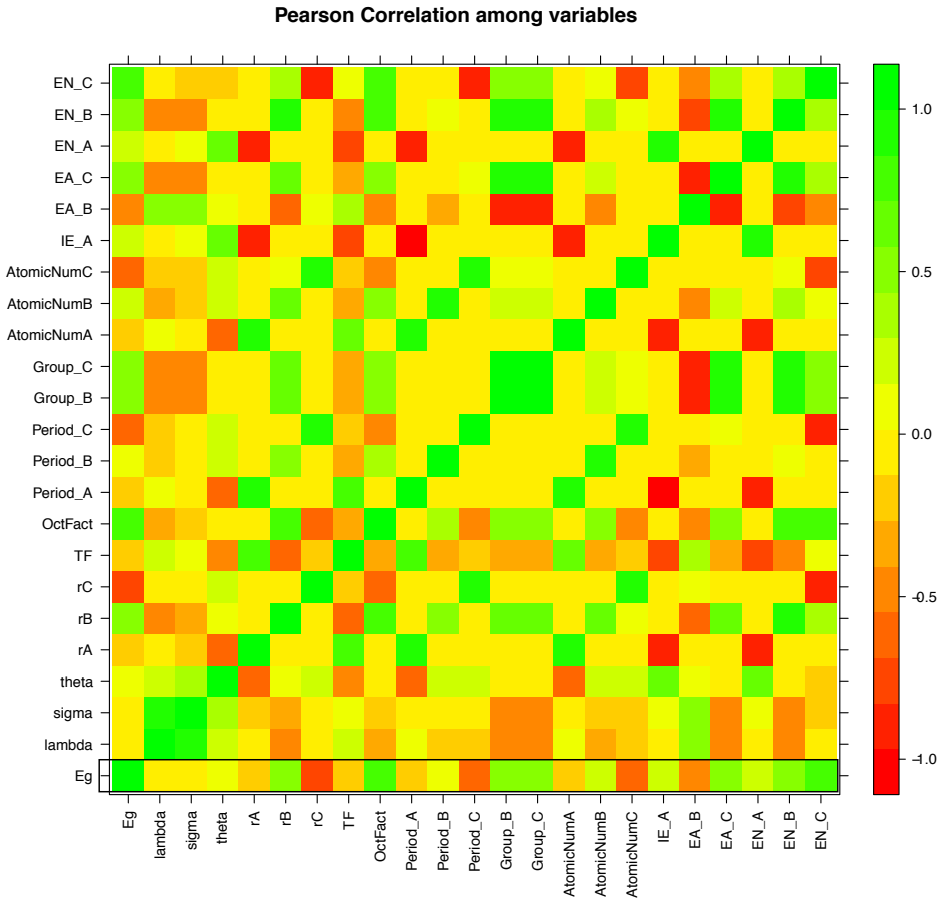

Figure S2: Pearson correlation coefficients between input features.

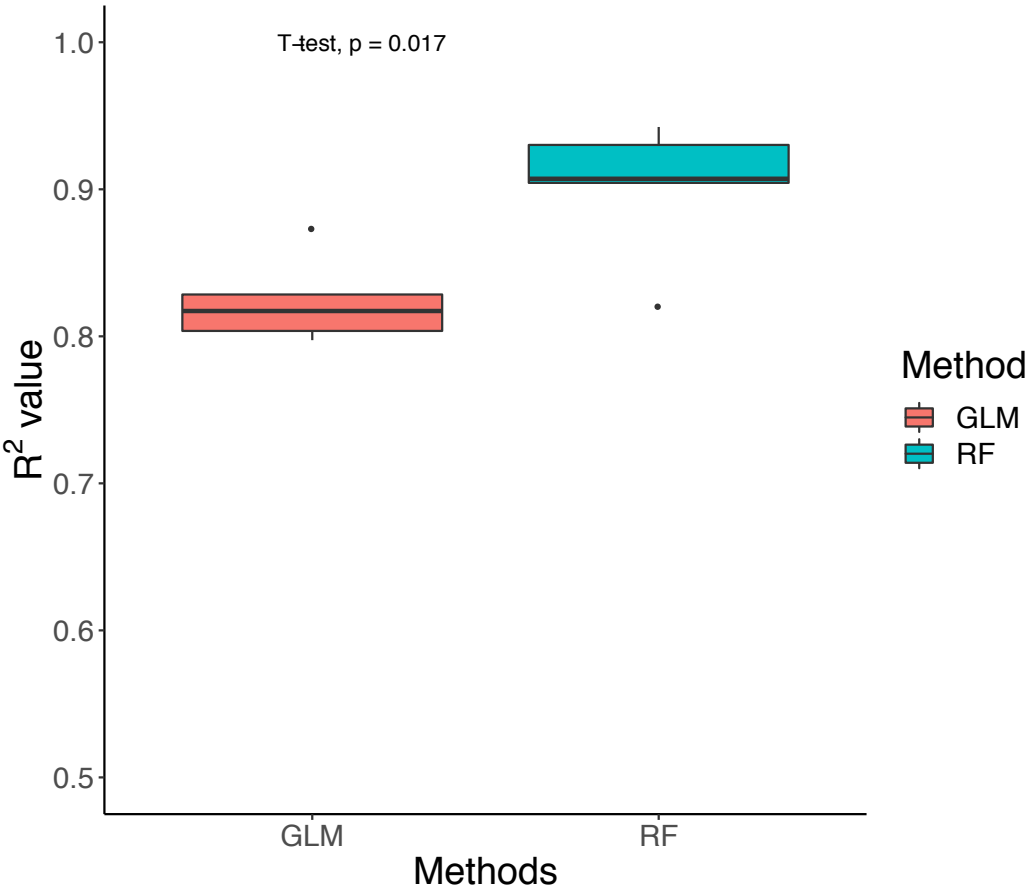

Figure S3: Comparison of R^2 values in the bandgap regression between general linear method (GLM) and random forest (RF).

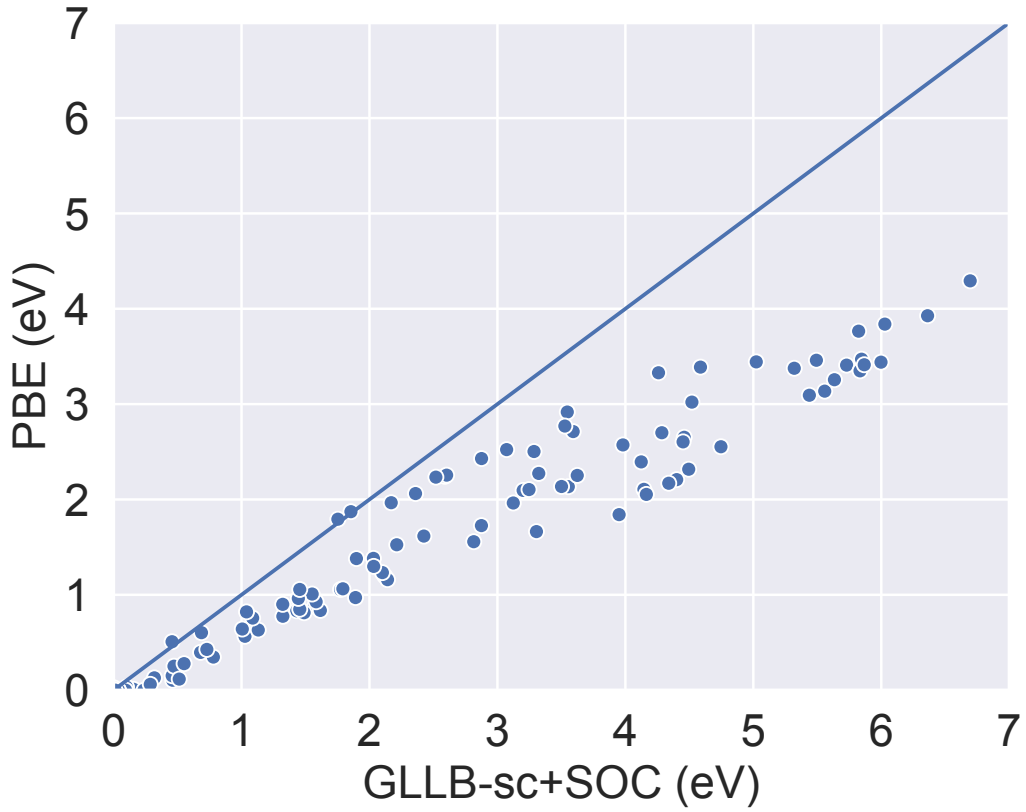

Figure S4: Comparison of the calculated bandgap at the PBE and GLLB-sc+SOC. The GLLB-sc bandgap was corrected by spin-orbit coupling.

Table S4: Coordinates in the t-SNE clustering plot of Figure 6 in the main manuscript.

Compound	x	y	Deformation	Compound	x	y	Deformation
$LiVO_3$	8.4	-16.1	Non-perovskite	$LiGeF_3$	6.3	-16.7	Non-perovskite
$NaVO_3$	12.0	-17.6	Non-perovskite	$NaGeF_3$	10.2	-15.6	Non-perovskite
KVO_3	11.4	-17.0	Non-perovskite	$KGeF_3$	11.0	-11.1	Non-perovskite
$RbVO_3$	10.9	-16.2	Non-perovskite	$RbGeF_3$	8.7	-10.3	Distorted
$CsVO_3$	-0.8	-6.5	Non-perovskite	$CsGeF_3$	7.9	-9.0	Distorted
$LiVS_3$	-0.8	-5.6	Non-perovskite	$LiGeCl_3$	-0.2	0.4	Non-perovskite
$NaVS_3$	10.9	-12.3	Distorted	$NaGeCl_3$	3.5	-3.7	Distorted
KVS_3	2.5	-3.3	Non-perovskite	$KGeCl_3$	8.6	-0.9	Close-to-ideal-cubic
$RbVS_3$	-0.4	-4.6	Non-perovskite	$RbGeCl_3$	8.5	-0.3	Close-to-ideal-cubic
$CsVS_3$	0.5	-1.2	Non-perovskite	$CsGeCl_3$	-3.1	12.2	Distorted
$LiVSe_3$	-1.3	-5.9	Non-perovskite	$LiGeBr_3$	-5.0	8.5	Distorted
$NaVSe_3$	10.3	-8.1	Distorted	$NaGeBr_3$	-2.3	11.8	Distorted
$KVSe_3$	5.6	-4.3	Non-perovskite	$KGeBr_3$	-3.6	13.1	Close-to-ideal-cubic
$RbVSe_3$	-1.7	-5.3	Non-perovskite	$RbGeBr_3$	-4.5	13.0	Close-to-ideal-cubic
$CsVSe_3$	1.1	2.6	Non-perovskite	$CsGeBr_3$	-8.7	12.4	Distorted
$LiVTe_3$	8.5	-4.7	Distorted	$LiGeI_3$	-5.3	11.7	Distorted

$NaVTe_3$	-2.2	-0.4	Distorted	$NaGeI_3$	-7.4	12.7	Close-to-ideal-cubic
$KVTe_3$	-6.4	6.5	Non-perovskite	$KGeI_3$	-14.8	12.4	Distorted
$RbVTe_3$	-3.9	2.0	Non-perovskite	$RbGeI_3$	-14.3	11.7	Close-to-ideal-cubic
$CsVTe_3$	1.4	6.2	Non-perovskite	$CsGeI_3$	-15.5	11.4	Distorted
$LiNbO_3$	9.6	-13.2	Non-perovskite	$LiSnF_3$	6.3	-16.7	Non-perovskite
$NaNbO_3$	15.2	-18.3	Close-to-ideal-cubic	$NaSnF_3$	11.0	-14.5	Non-perovskite
$KNbO_3$	14.4	-17.6	Close-to-ideal-cubic	$KSnF_3$	10.2	-11.6	Distorted
$RbNbO_3$	14.5	-16.9	Distorted	$RbSnF_3$	10.2	-9.4	Distorted
$CsNbO_3$	9.1	-15.7	Non-perovskite	$CsSnF_3$	8.3	-8.1	Distorted
$LiNbS_3$	1.0	-1.9	Non-perovskite	$LiSnCl_3$	-0.3	1.5	Non-perovskite
$NaNbS_3$	10.9	-9.9	Distorted	$NaSnCl_3$	7.7	0.6	Distorted
$KNbS_3$	7.3	-5.9	Distorted	$KSnCl_3$	-3.1	11.4	Distorted
$RbNbS_3$	6.9	-3.9	Distorted	$RbSnCl_3$	-4.2	12.2	Close-to-ideal-cubic
$CsNbS_3$	0.5	0.7	Non-perovskite	$CsSnCl_3$	-5.8	13.0	Close-to-ideal-cubic
$LiNbSe_3$	-2.2	0.1	Non-perovskite	$LiSnBr_3$	-1.3	1.6	Distorted
$NaNbSe_3$	5.9	-5.3	Distorted	$NaSnBr_3$	-6.3	11.2	Distorted
$KNbSe_3$	5.3	-2.7	Distorted	$KSnBr_3$	-7.7	12.0	Close-to-ideal-cubic
$RbNbSe_3$	-3.3	-0.6	Distorted	$RbSnBr_3$	-11.1	12.8	Close-to-ideal-cubic
$CsNbSe_3$	1.2	3.9	Non-perovskite	$CsSnBr_3$	-12.3	12.8	Close-to-ideal-cubic
$LiNbTe_3$	9.1	-3.0	Close-to-ideal-cubic	$LiSnI_3$	-7.1	7.3	Distorted
$NaNbTe_3$	-3.5	2.9	Non-perovskite	$NaSnI_3$	-14.1	12.9	Close-to-ideal-cubic
$KNbTe_3$	-9.3	7.6	Non-perovskite	$KSnI_3$	-15.5	13.1	Close-to-ideal-cubic
$RbNbTe_3$	-11.1	7.4	Non-perovskite	$RbSnI_3$	-16.9	11.8	Close-to-ideal-cubic
$CsNbTe_3$	1.3	6.3	Non-perovskite	$CsSnI_3$	-17.8	11.7	Close-to-ideal-cubic
$LiTaO_3$	12.4	-16.8	Non-perovskite	$LiPbF_3$	10.1	-16.8	Non-perovskite
$NaTaO_3$	14.5	-18.4	Close-to-ideal-cubic	$NaPbF_3$	11.6	-15.8	Distorted
$KTaO_3$	15.5	-17.7	Close-to-ideal-cubic	$KPbF_3$	8.4	-11.3	Distorted
$RbTaO_3$	15.2	-17.0	Close-to-ideal-cubic	$RbPbF_3$	7.6	-10.5	Close-to-ideal-cubic
$CsTaO_3$	10.9	-13.5	Non-perovskite	$CsPbF_3$	7.4	-9.7	Close-to-ideal-cubic
$LiTaS_3$	7.3	-0.8	Non-perovskite	$LiPbCl_3$	7.1	-2.8	Distorted
$NaTaS_3$	10.2	-10.1	Distorted	$NaPbCl_3$	7.1	0.8	Distorted
$KTaS_3$	8.0	-6.2	Distorted	$KPbCl_3$	-2.7	12.7	Close-to-ideal-cubic
$RbTaS_3$	6.7	-4.5	Distorted	$RbPbCl_3$	-5.8	12.3	Close-to-ideal-cubic
$CsTaS_3$	0.3	2.4	Non-perovskite	$CsPbCl_3$	-8.0	13.1	Close-to-ideal-cubic
$LiTaSe_3$	-2.7	1.2	Non-perovskite	$LiPbBr_3$	-5.2	9.9	Distorted
$NaTaSe_3$	5.8	-3.6	Distorted	$NaPbBr_3$	-5.9	9.8	Distorted
$KTaSe_3$	1.3	3.5	Non-perovskite	$KPbBr_3$	-10.3	12.4	Close-to-ideal-cubic
$RbTaSe_3$	-4.0	2.7	Non-perovskite	$RbPbBr_3$	-11.7	12.1	Close-to-ideal-cubic
$CsTaSe_3$	-12.2	7.4	Non-perovskite	$CsPbBr_3$	-13.2	12.3	Close-to-ideal-cubic
$LiTaTe_3$	9.0	-2.5	Close-to-ideal-cubic	$LiPbI_3$	-7.5	6.8	Distorted
$NaTaTe_3$	-3.3	0.1	Non-perovskite	$NaPbI_3$	-15.5	12.1	Close-to-ideal-cubic

$KTaTe_3$	-6.4	7.1	Non-perovskite	$KPbI_3$	-16.4	12.7	Close-to-ideal-cubic
$RbTaTe_3$	-10.7	7.4	Non-perovskite	$RbPbI_3$	-17.7	12.8	Close-to-ideal-cubic
$CsTaTe_3$	0.6	5.9	Non-perovskite	$CsPbI_3$	-18.3	12.5	Close-to-ideal-cubic
				$MAPbI_3$	-19.0	11.8	Close-to-ideal-cubic
				$FAPbI_3$	-19.4	12.6	Close-to-ideal-cubic
				$GauPbI_3$	-20.1	12.1	Close-to-ideal-cubic

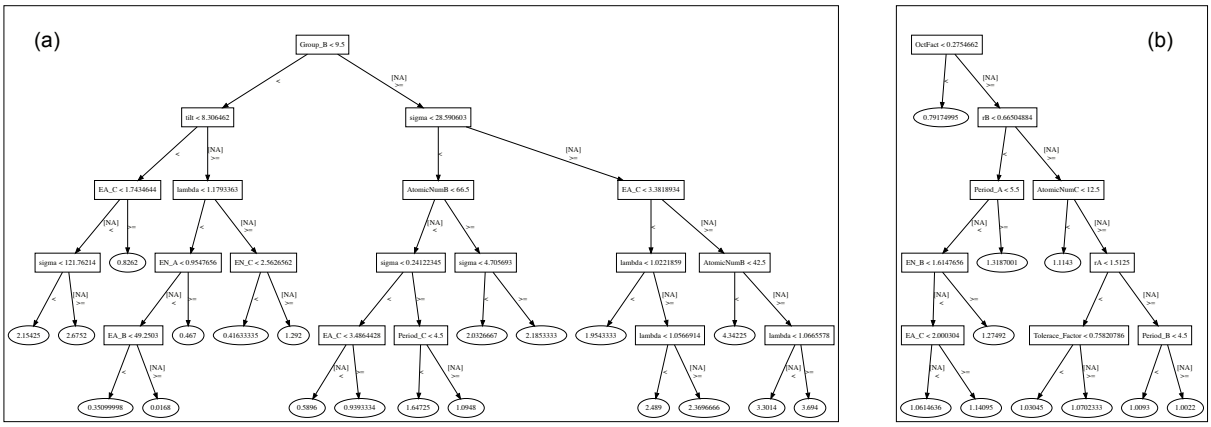

Figure S5: Visualization of the decision trees. These trees are the part of the RF regression on (a) E_{g} and (b) λ.