Supplementary Materials

Accelerated development of hard high entropy alloys by data-driven high-throughput experiments

Yi Liu[*1,2], Jiong Wang[1], Bin Xiao[1], and Jintao Shu[1]
[1]Materials Genome Institute, Shanghai University, Shanghai 200444, China
[2]Zhejiang Laboratory, Hangzhou 311100, China
[*]Email：YiLiu@shu.edu.cn

Text S1 Linear regression models of hardness

We first fit the H48 dataset in the exp-1.1 using Mo and W compositions as variables, and obtained a linear equation eq. (1), correlating with the experimental data with R^2= 0.675. The comparison between linear regression prediction and experiment results is shown in Figure S1 of Supplementary Materials (SM).

$$\mathrm{HV} = 67.37 * \mathrm{Mo} - 121.61 * \mathrm{W} + 782.5 \quad \text{eq. (S1)}$$

We then included all the 111 composition-hardness data (H111) in the exp-1.1 and exp-1.2 into the linear regression fitting using all five components as variables. The fitted linear equation is shown as eq. (2) where R^2= 0.363. The linear regression prediction and experiment results are shown in Figure S2 of SM.

$$\mathrm{HV} = -137.33 * \mathrm{Co} + 414.89 * \mathrm{Cr} - 220.38 * \mathrm{Ti} - 52.81 * \mathrm{Mo} - 590.36 * \mathrm{W} + 793.46 \quad \text{eq. (S2)}$$

It is found that the conventional linear regression models may work reasonably well (R^2= 0.675) for a few variables [e.g., W and Mo in eq. (1)], however, the prediction accuracy dropped quickly (R^2= 0.362) as the number of variables increases [e.g., five components in eq. (2)]. Also, the negative coefficients of the components may not have physical meanings since many of them are considered to have positive contributions to the hardness. Even worse, the coefficients of Mo terms had opposite signs in eq. (1) and eq. (2), respectively, showing contradictory influence and little generality.

Table S1 Definition of 29 descriptors used in this work: AD (1-29): all descriptors; EMF(1-5): element mole fraction; VDSHOT(6-11): six phase descriptors; EP(12-29): elementary properties of elements.

No.	Symbol	Definition	No.	Symbol	Definition
1	Co	Cobalt mole fraction	16	DC	Distance, core electron (Schubert) Molar
2	Cr	Chromium mole fraction	17	DV	Distance, valence electron (Schubert) Moment
3	Ti	Titanium mole fraction	18	EN	Electronegativity (Martynov and Batsanov) First oxidation
4	Mo	Molybdenum mole fraction	19	EC	Energy, cohesive Brewer
5	W	Tungsten mole fraction	20	EI	Energy of ionization, first
6	VEC	Valence electron concentration	21	EM	Enthalpy of melting
7	ΔH_{mix}	Enthalpy of mixing	22	MP	Mendeleev Pettifor
8	ΔS_{mix}	Entropy of mixing	23	NC	Nuclear charge, effective
9	T_m	Average melting point	24	QN	Quantum number
10	Ω	omega = T_m*ΔSmix/ |Δ Hmix|	25	RC	Radius, covalent
11	δ	Atomic size difference	26	RM	Radius, metal (Waber)
12	AN	Atomic number	27	TM	Average melting point
13	AW	Atomic weight	28	TN	Thermal neutron capture cross section
14	CH	Charge, nuclear effective (Clementi)	29	BM	Bulk modulus
15	DS	Density			

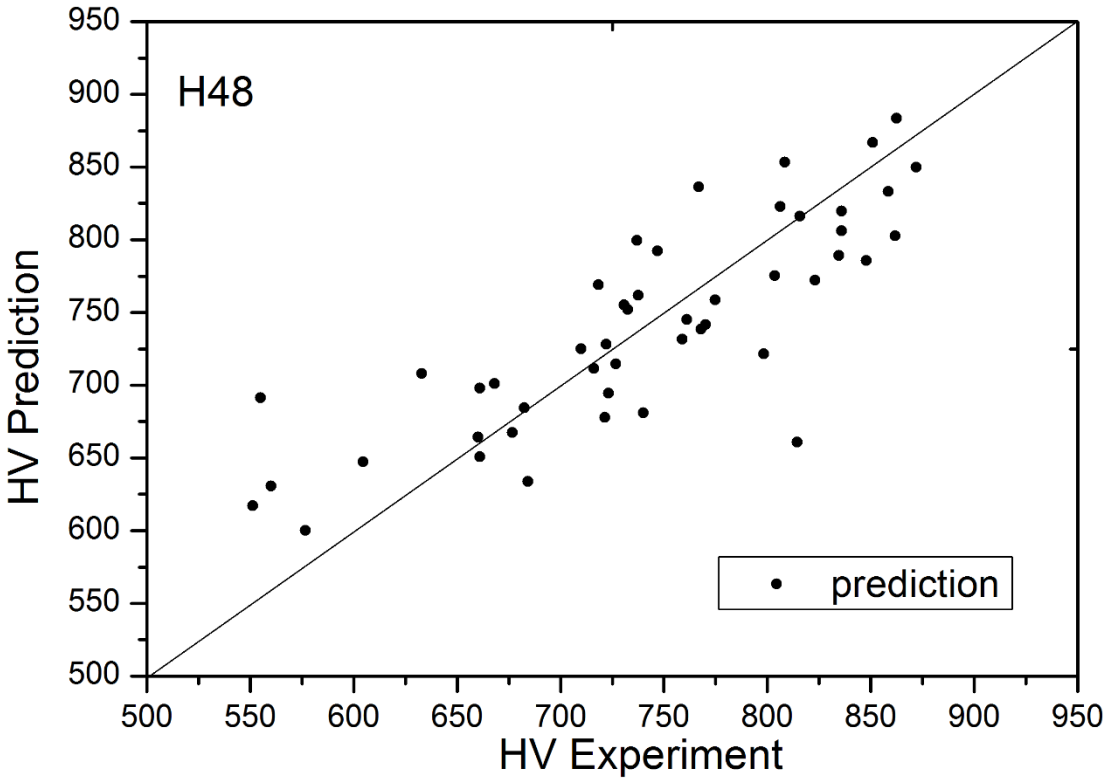


Figure S1 Linear regression fitting of H48 dataset (only W and Mo vary).

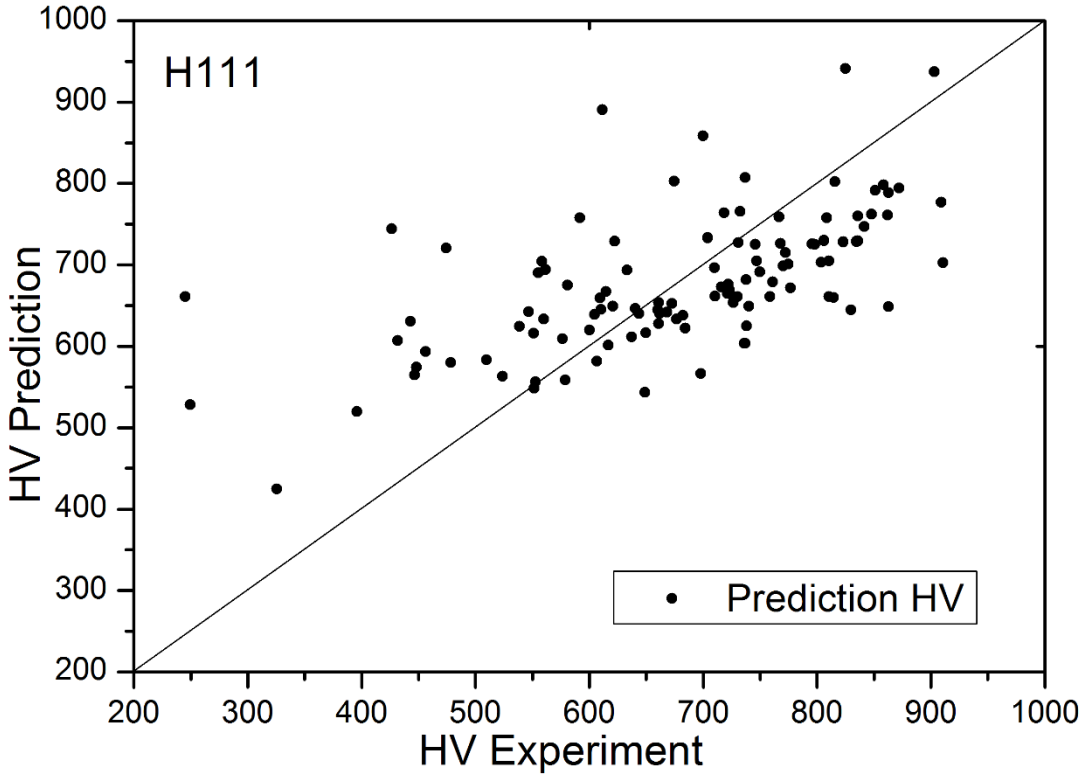
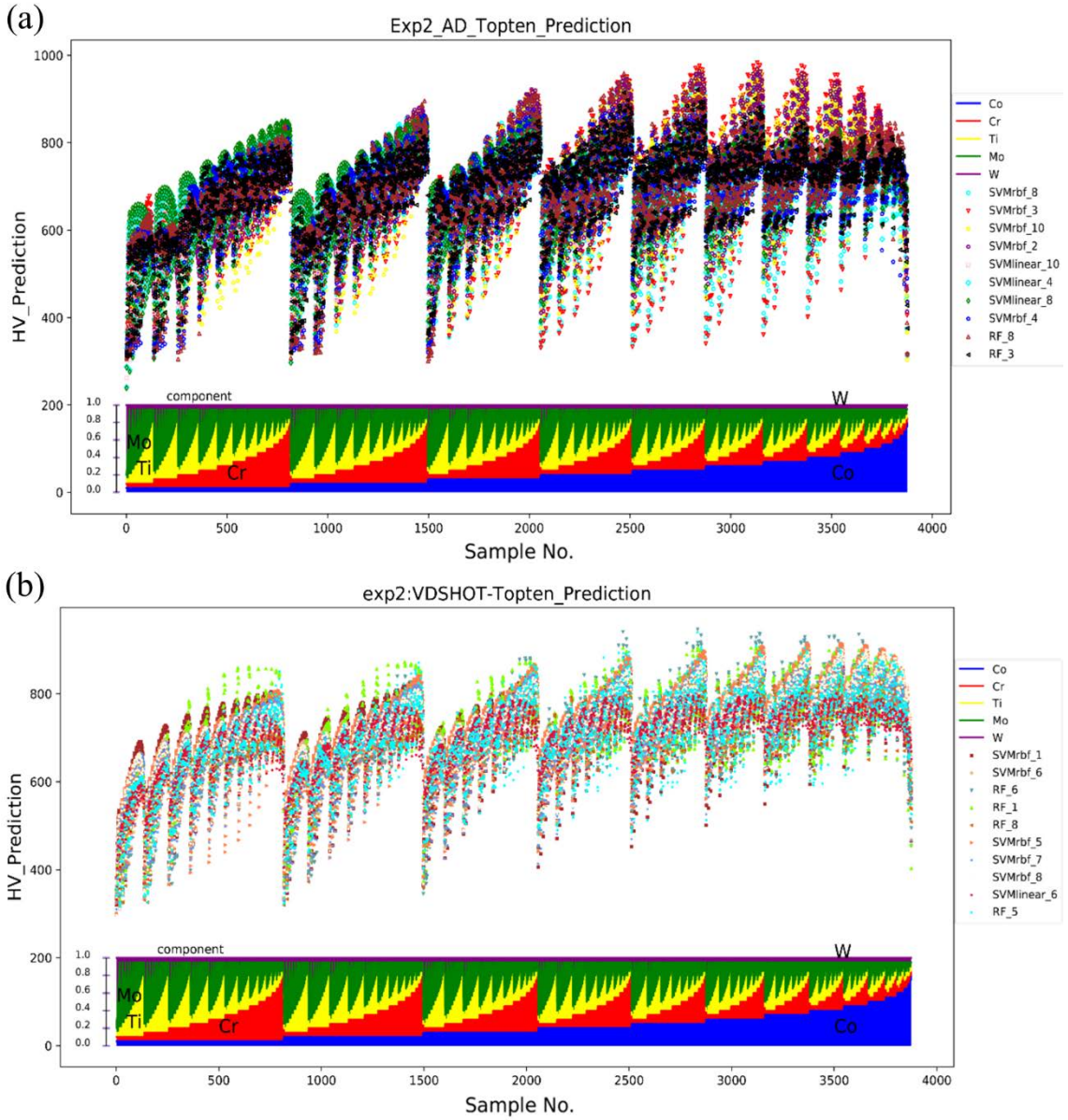


Figure S2 Linear regression fitting of H111 dataset at the experiment stage I.


(a)
Exp2_AD_Topten_Prediction
HV_Prediction
Sample No.
component
W
Co
Cr
Mo
Ti
Co
Cr
Ti
Mo
W
SVMrbf_8
SVMrbf_3
SVMrbf_10
SVMrbf_2
SVMlinear_10
SVMlinear_4
SVMlinear_8
SVMrbf_4
RF_8
RF_3
(b)
exp2:VDSHOT-Topten_Prediction
HV_Prediction
Sample No.
component
W
Co
Cr
Mo
Ti
Co
Cr
Ti
Mo
W
SVMrbf_1
SVMrbf_6
RF_6
RF_1
RF_8
SVMrbf_5
SVMrbf_7
SVMrbf_8
SVMlinear_6
RF_5

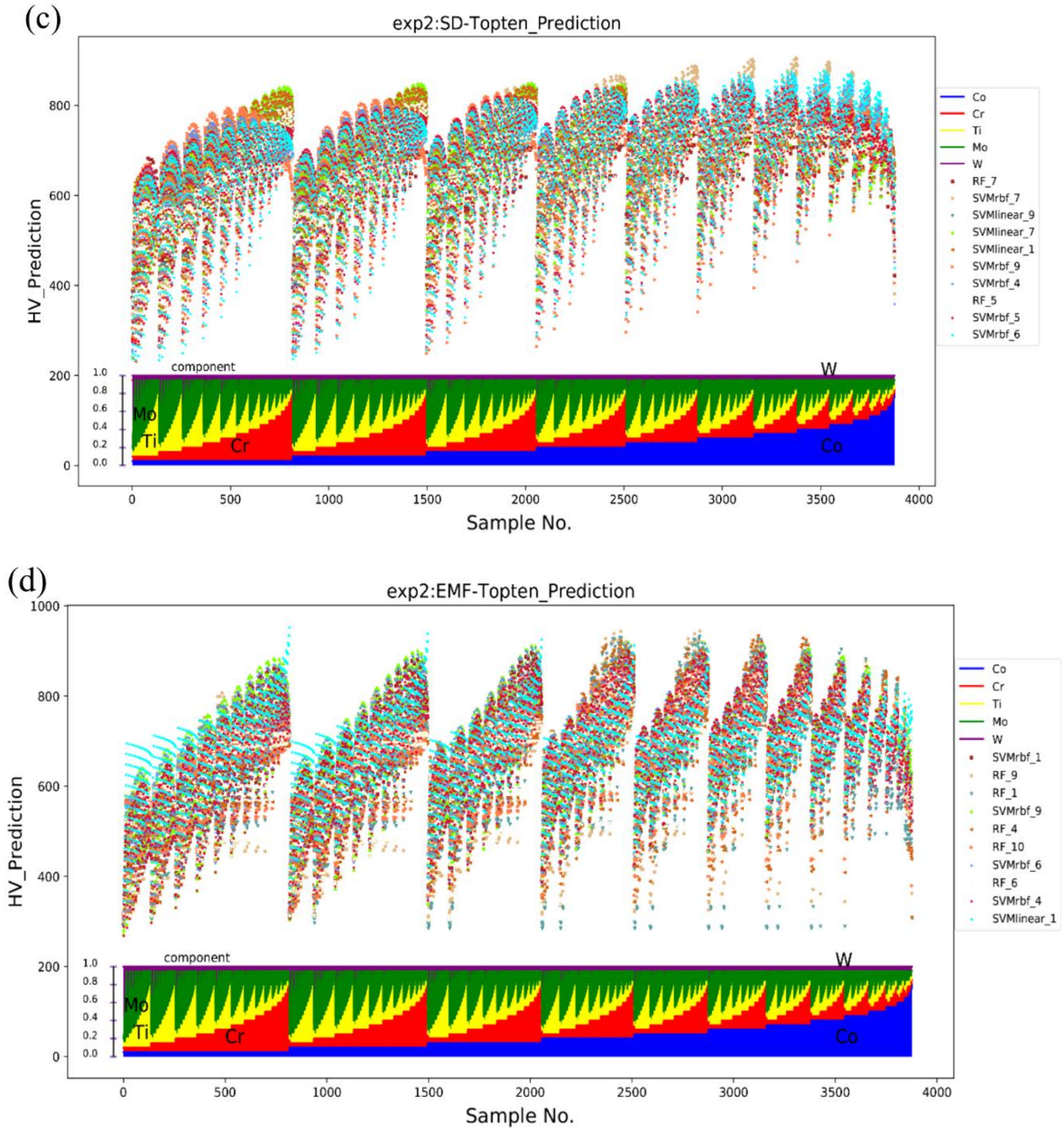


Figure S3 Hardness of the 3876 hypothetical HEAs predicted by the top 10 best ML-2 models (H138) from the four descriptor groups, respectively: (a) all descriptors (AD); (b) VDSHOT; (c) selected descriptors (SD); (d) element mole fraction (EMF). The lower insets show the compositions of the HEAs.

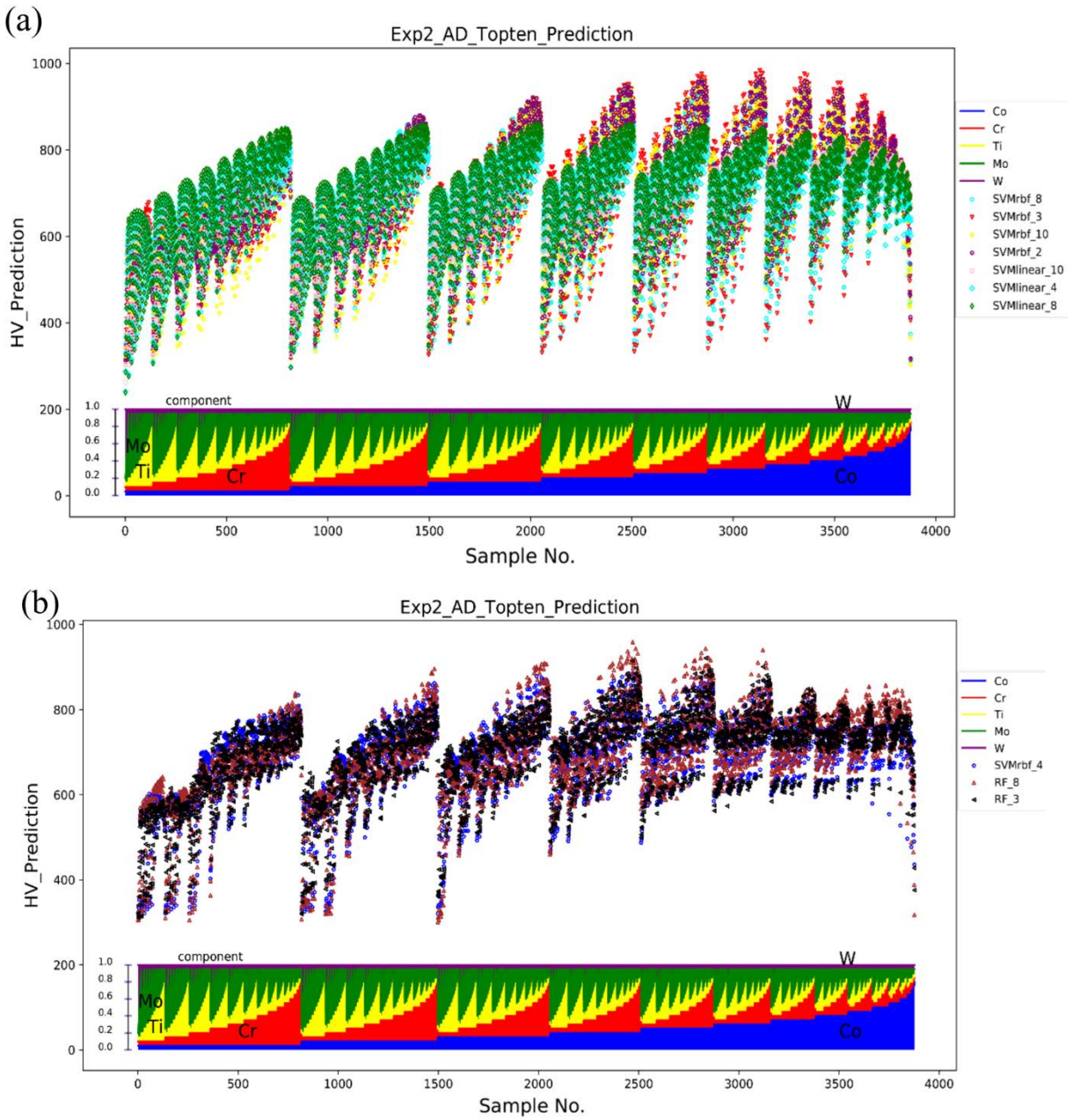
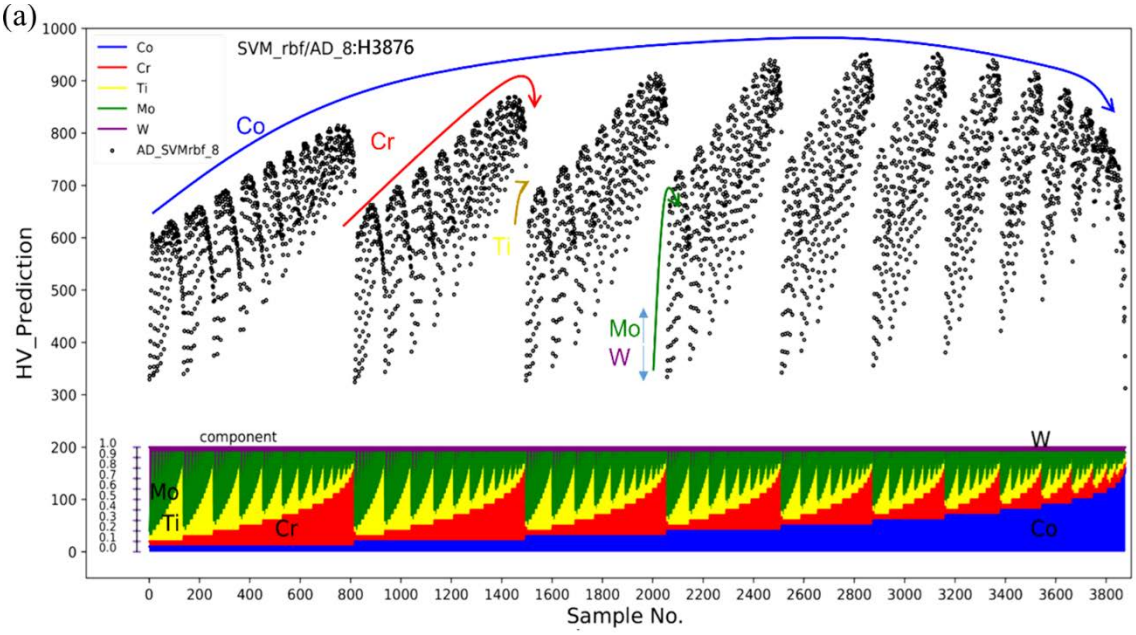
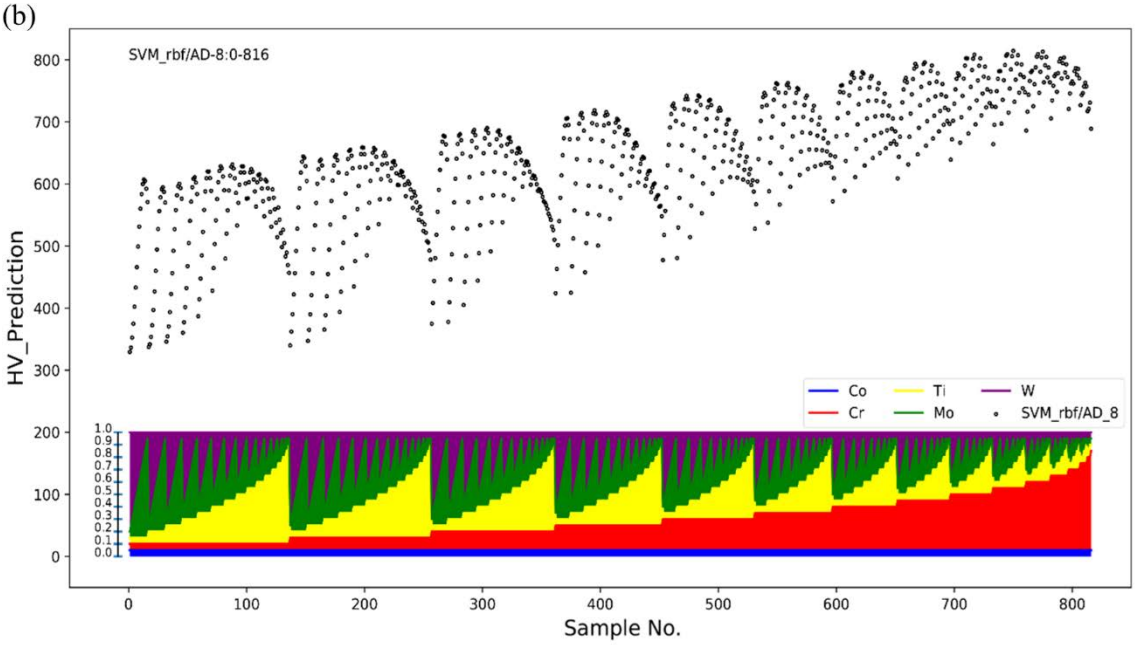


Figure S4 Hardness of the 3876 hypothetical HEAs predicted by the top 10 best ML-2 models (H138) using all descriptors (AD): (a) the 7 ML models with the qualitatively similar trends; (b) the 3 ML models with the qualitatively different trends. The lower insets show the compositions of the HEAs.

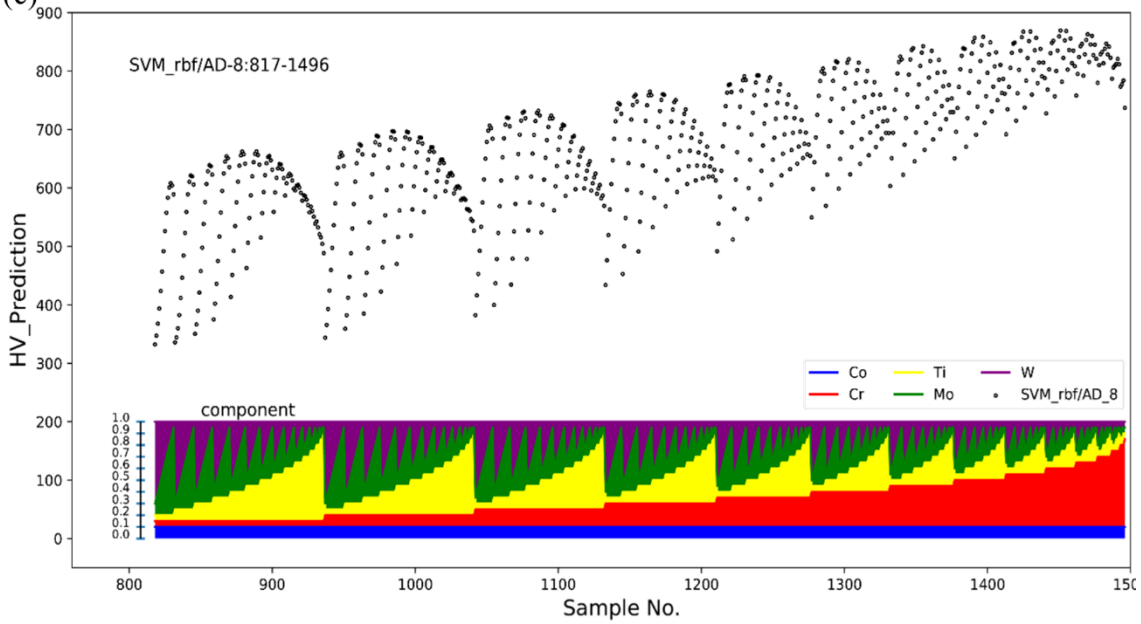
(a)


SVM_rbf/AD_8:H3876
Co
Cr
Ti
Mo
W
AD_SVMrbf_8
HV_Prediction
component
Sample No.

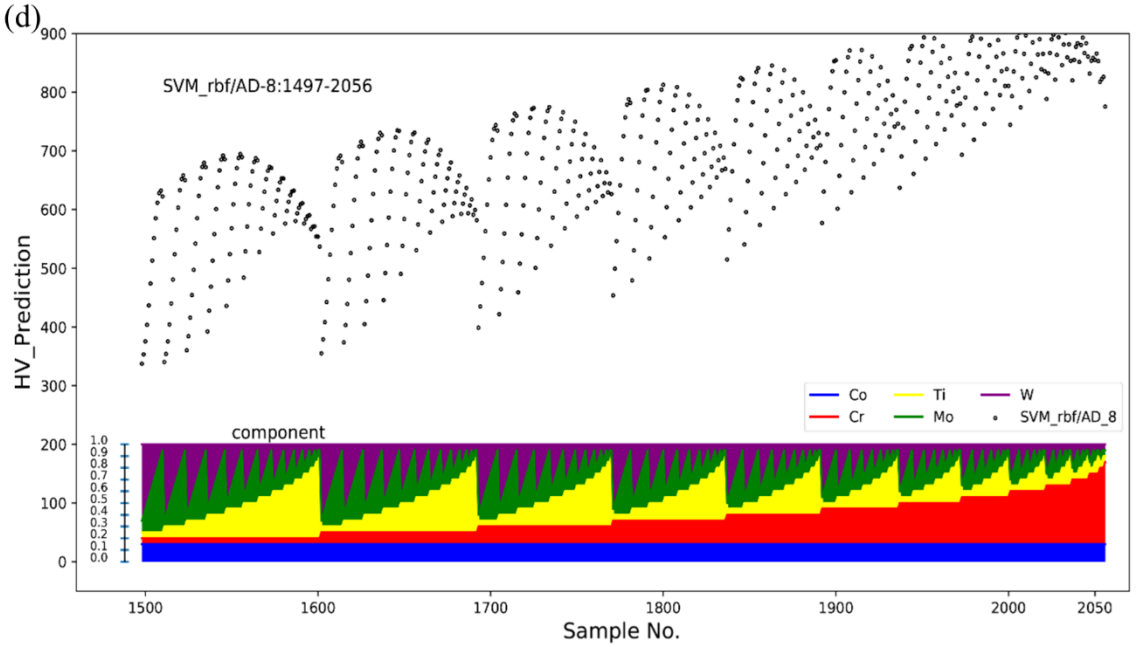
(b)


SVM_rbf/AD-8:0-816
HV_Prediction
Co
Cr
Ti
Mo
W
SVM_rbf/AD_8
Sample No.

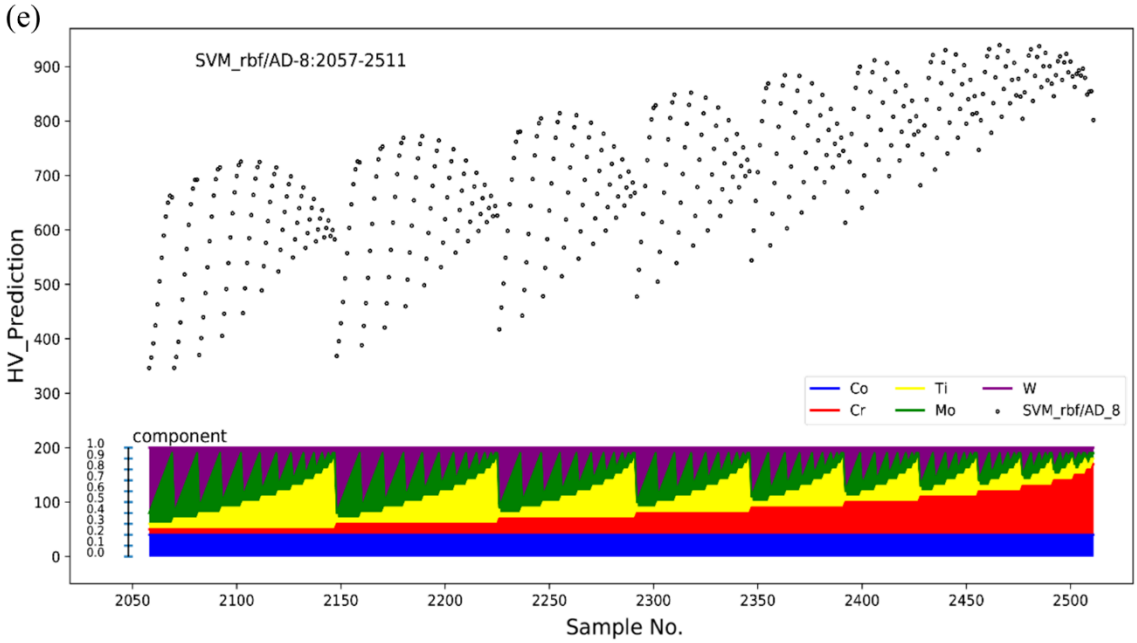
(c)


SVM_rbf/AD-8:817-1496
HV_Prediction
Sample No.
component
Co
Cr
Ti
Mo
W
SVM_rbf/AD_8

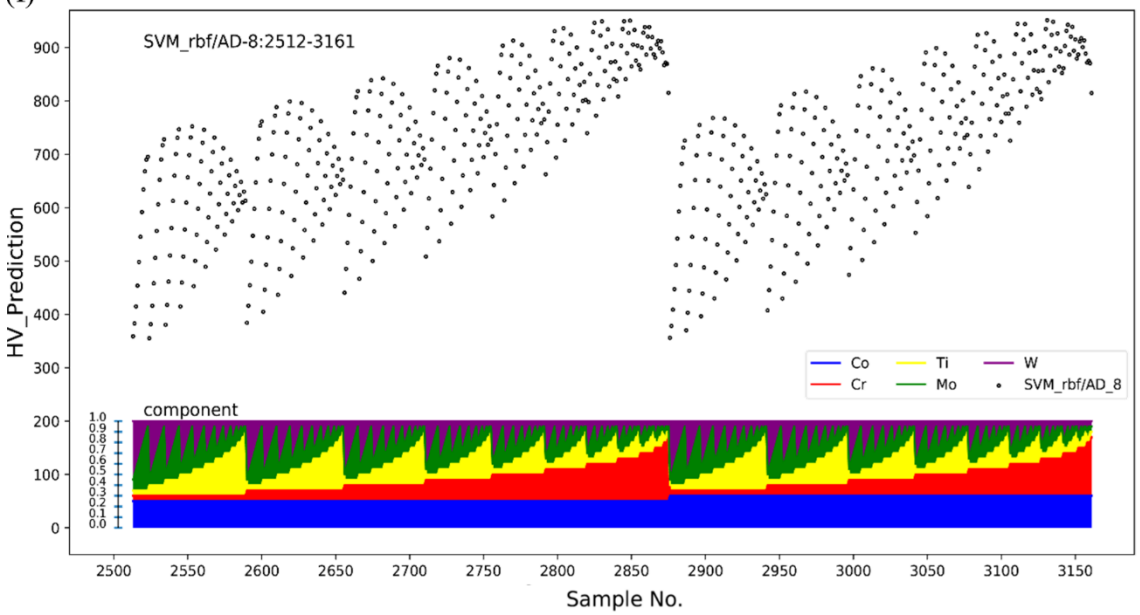
(d)


SVM_rbf/AD-8:1497-2056
HV_Prediction
Sample No.
component
Co
Cr
Ti
Mo
W
SVM_rbf/AD_8

(e)


SVM_rbf/AD-8:2057-2511
HV_Prediction
Sample No.
component
Co
Cr
Ti
Mo
W
SVM_rbf/AD_8

(f)


SVM_rbf/AD-8:2512-3161
HV_Prediction
Sample No.
component
Co
Cr
Ti
Mo
W
SVM_rbf/AD_8

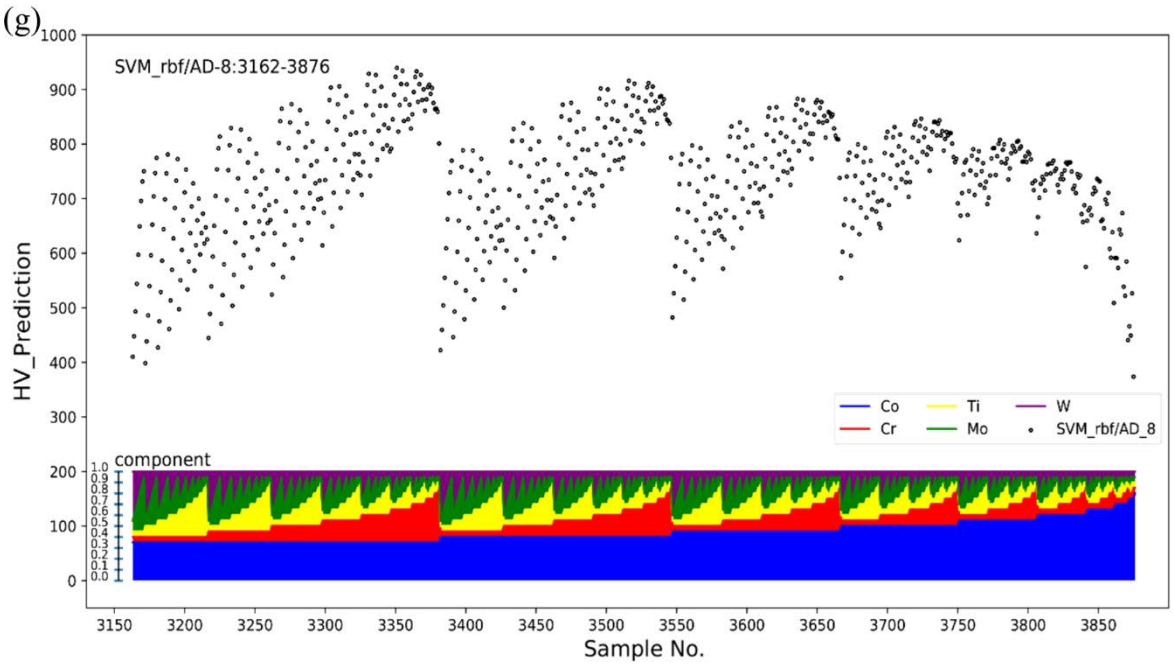


Figure S5 Hardness of the 3876 hypothetical HEAs predicted by the best ML-2 model (H138) SVM_rbf/AD_8: (a) the whole H3876 dataset; (b)-(g) the H3876 dataset at the separated composition ranges.

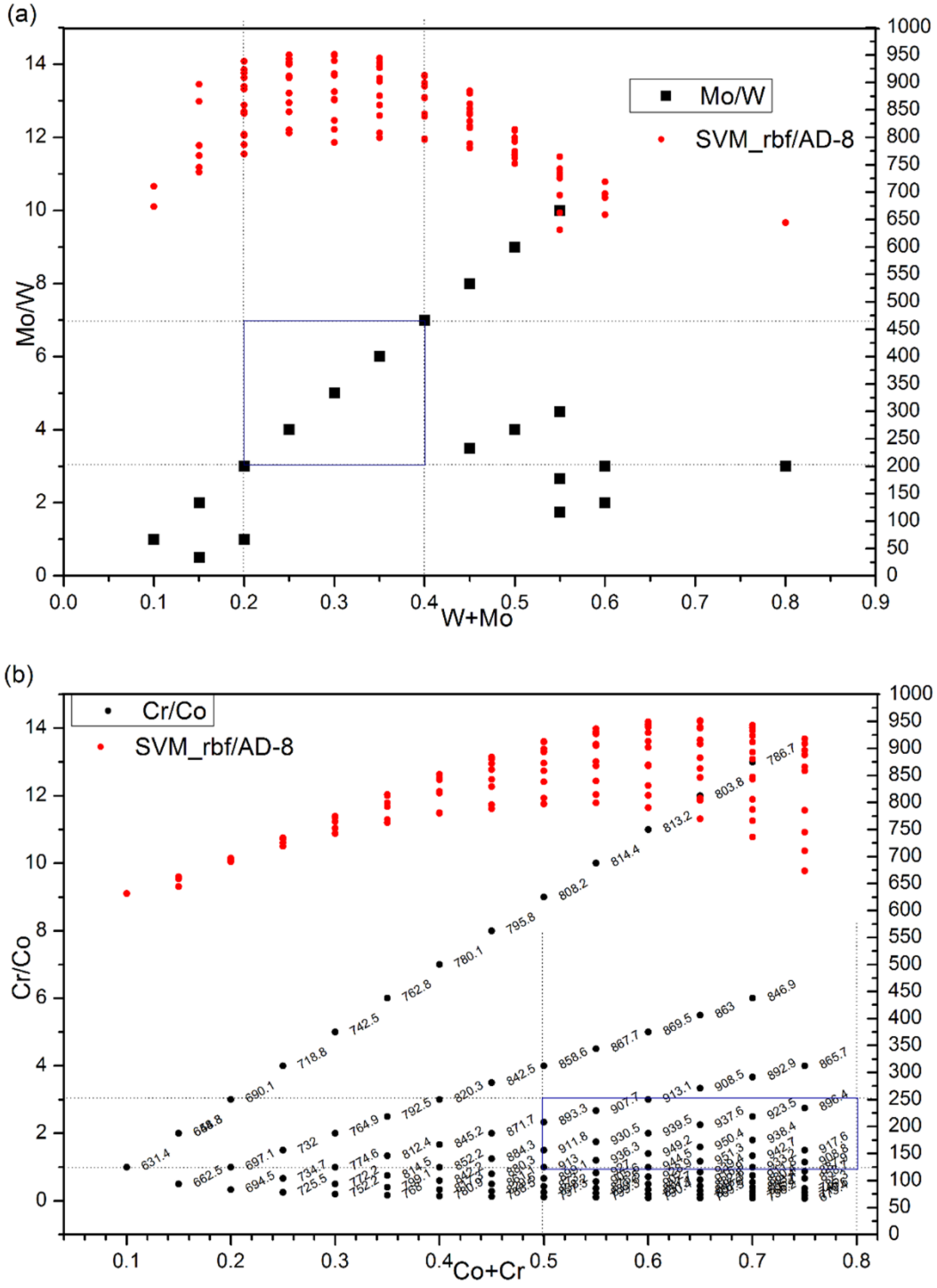
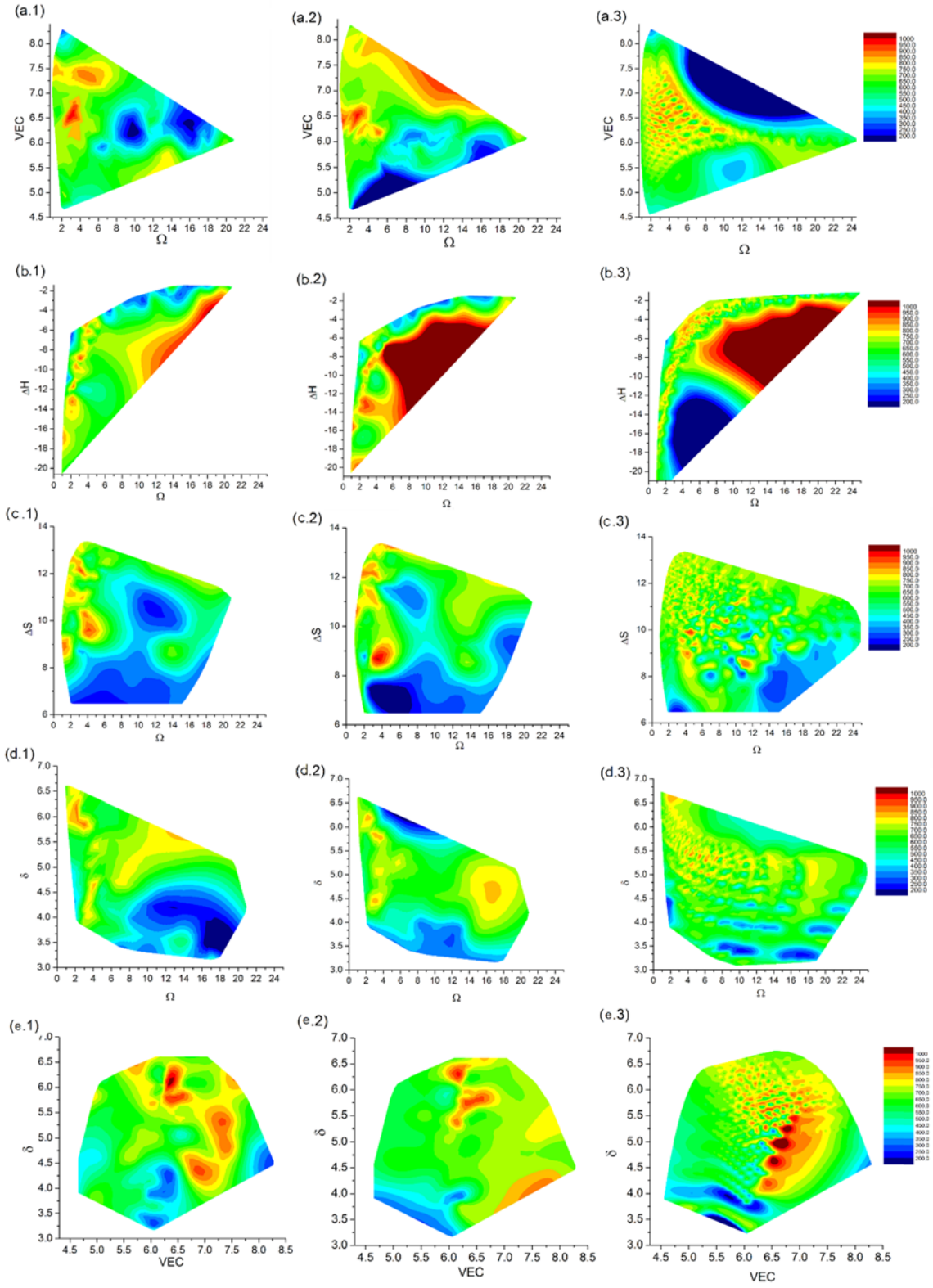


Figure S6 Selected maximum hardness as functions of (a) Mo/W= 3-7 and W+Mo= 0.2-0.4; (b) Cr/Co= 1-3 and Co+Cr= 0.5-0.8 of the hypothetical HEAs.


(a.1)
(a.2)
(a.3)
VEC
Ω
(b.1)
(b.2)
(b.3)
ΔH
Ω
(c.1)
(c.2)
(c.3)
ΔS
Ω
(d.1)
(d.2)
(d.3)
δ
Ω
(e.1)
(e.2)
(e.3)
δ
VEC

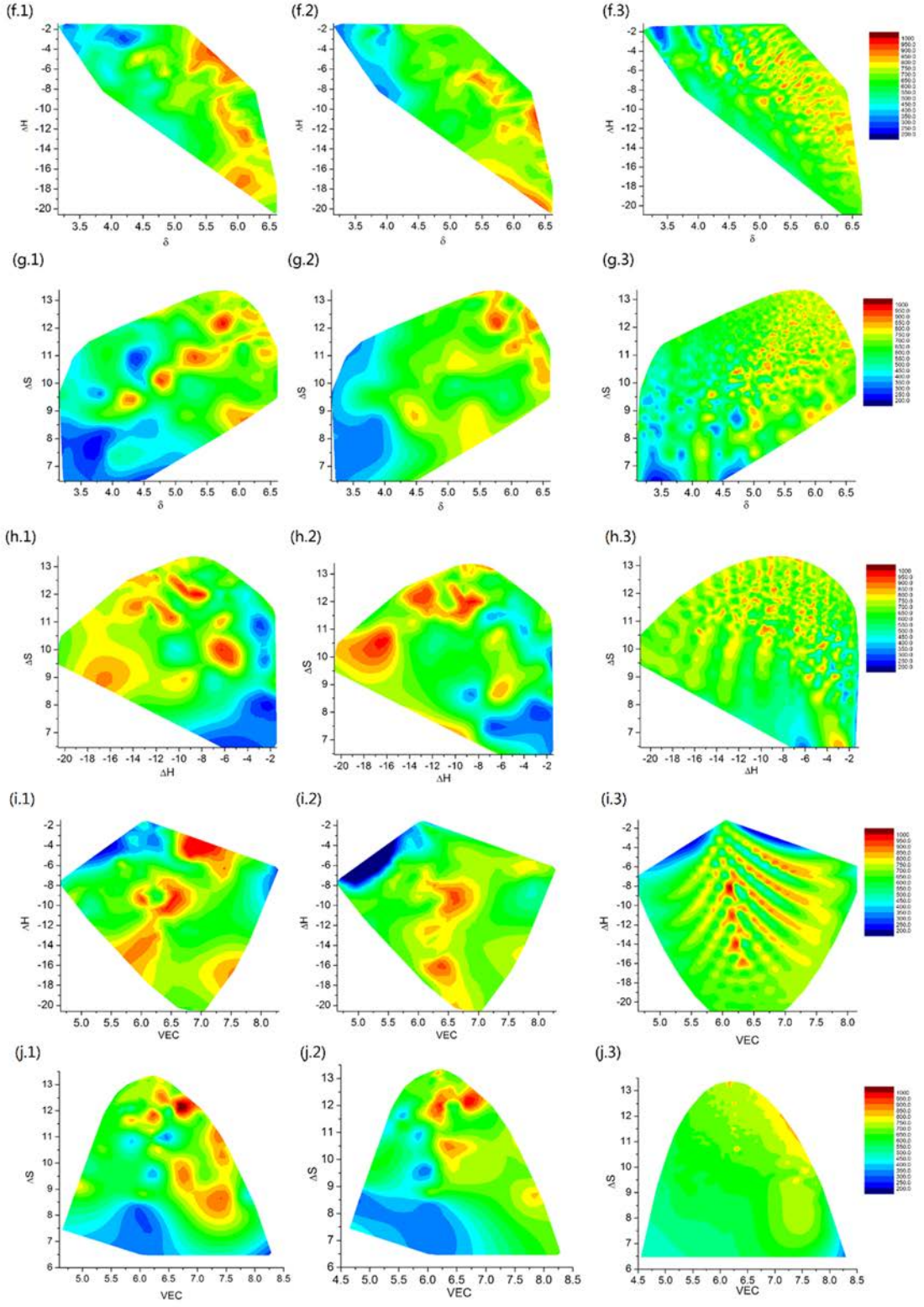


Figure S7 Descriptor-hardness contour maps for (a.1)-(i.1) the experimental hardness (H138); (a.2)-(i.2) the H138 dataset and (a.3)-(i.3) the H3876+H48 dataset predicted by the best ML-2 model (H138), SVM_rbf/AD_8. The 15 pairs of descriptors are chosen from VEC, ΔH, ΔS, δ, and Ω.