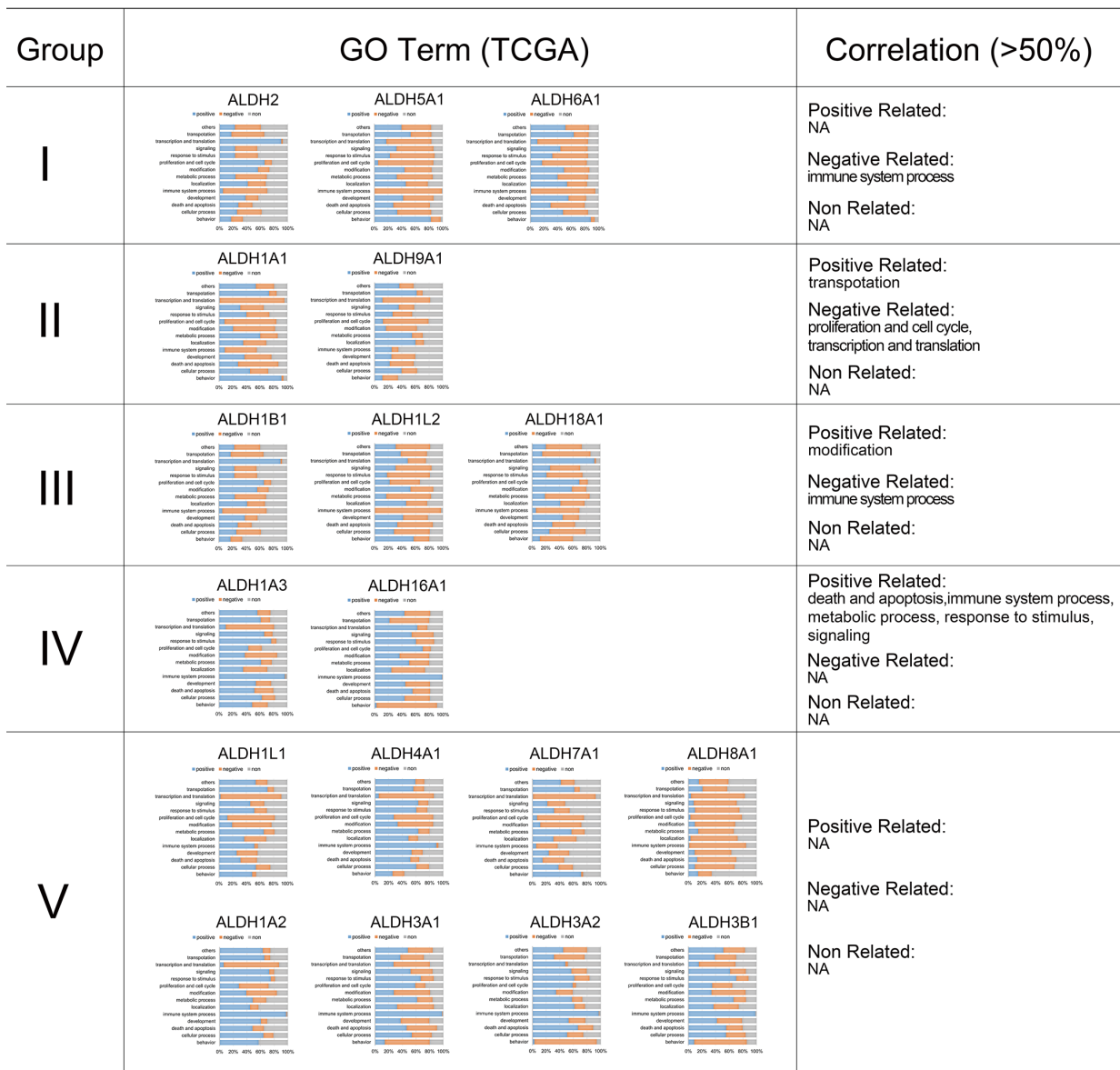
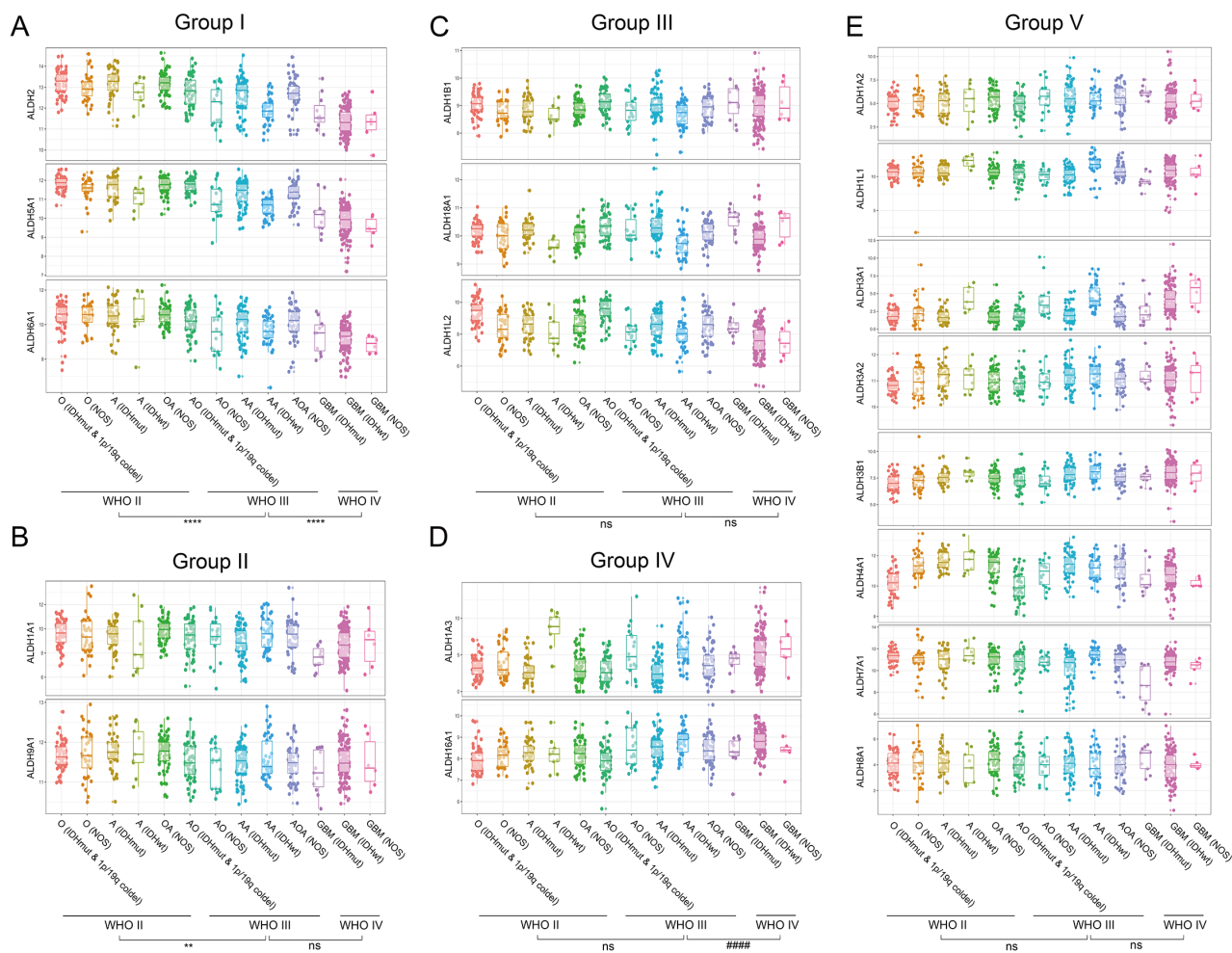


**Figure S1** Multi-platform cluster analysis of ALDH in TCGA database. (A,B) Unsupervised clustering of all ALDH isoforms by gene mRNA expression and gene methylation level in TCGA database. (C,D) Unsupervised biological clustering of all ALDH isoforms in TCGA database. ALDH isoforms can be classified into 4 groups according to their biological functions and 4 groups according to their signal pathway activations. (E) ALDH isoforms can be classified into 5 groups by integrated Cluster-Of-Cluster assignments analysis according to their biological functions and signal pathway activations.



**Figure S2** Landscape of biological functions of all ALDH isoforms in 5 groups in CGGA database. Most related biological functions of ALDH isoforms in each group. Only all ALDH isoforms in a group are significantly (not) related to more than 50% biological functions of one classification is defined as most (non) related.



**Figure S3** The clinical features of all ALDH isoforms in 5 groups. (A) In TCGA database, the expression of all ALDH isoforms in group I decreases as the WHO grade increases. \*\*\*\*:  $P < 0.0001$ , \*\*\*:  $P < 0.001$ . (B) In TCGA database, the expression of all ALDH isoforms in group II decreases as the WHO grade increases. \*\*:  $P < 0.01$ , ns:  $P > 0.05$ . (C) In TCGA database, the expression of ALDH isoforms in group III does not share the same changing trend with the increase of WHO grade. ns:  $p > 0.05$ . (D) In TCGA database, the expression of all ALDH isoforms in group IV increases as the WHO grade increases. ####:  $P < 0.0001$ , ns:  $P > 0.05$ . (E) In TCGA database, the expression of ALDH isoforms in group V does not share the same changing trend with the increase of WHO grade. ns:  $P > 0.05$ .

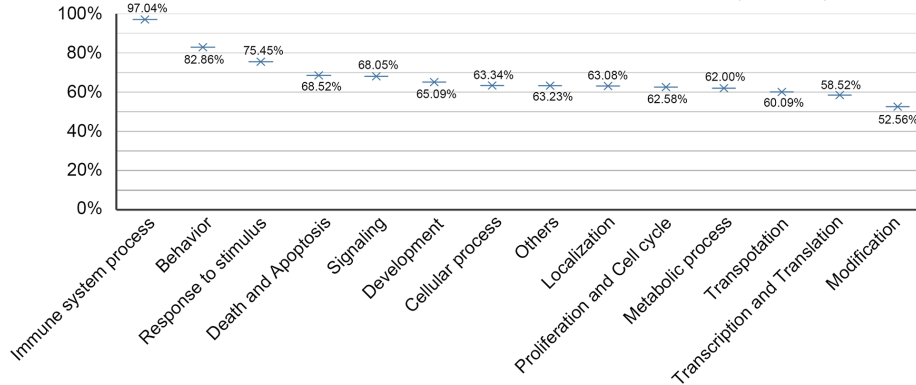
A

### DNA repair functions (TCGA)

Group	I			II		III			IV		V							
	ALDH2	ALDH5A1	ALDH6A1	ALDH1A1	ALDH8A1	ALDH1B1	ALDH1L2	ALDH18A1	ALDH1A3	ALDH16A1	ALDH1A2	ALDH1L1	ALDH3A1	ALDH3A2	ALDH3B1	ALDH4A1	ALDH7A1	ALDH8A1
Poly(ADP-ribose) polymerase (PARP) enzymes that bind to DNA	0.822	0.664	0.815	0.301	0.281	0.063	0.340	0.013	-0.292	-0.386	-0.090	0.885	-0.334	-0.140	-0.358	0.030	0.180	0.073
NER-related	0.822	0.709	0.873	0.111	0.203	0.100	0.453	0.224	-0.395	-0.445	-0.021	-0.063	-0.575	-0.071	-0.293	0.021	0.006	0.006
TFIIH/Catalyzes unwinding in preinitiation complex	0.278	0.258	0.286	0.094	0.189	-0.114	0.003	-0.161	-0.137	-0.141	0.000	0.071	-0.281	-0.037	-0.113	0.141	0.001	0.050
Nucleotide excision repair (NER) (XP = xeroderma pigmentosum)	0.317	0.086	0.114	0.301	0.481	-0.170	-0.316	-0.281	0.103	-0.060	0.077	0.498	0.290	0.367	0.271	0.419	0.207	-0.170
Ubiquitination and modification	0.214	0.278	0.405	0.161	0.197	-0.032	0.135	-0.048	0.099	-0.576	0.154	0.068	-0.116	-0.048	-0.223	0.051	0.060	-0.209
Base excision repair (BER)/DNA glycosylases: major altered base released	0.196	-0.070	-0.237	0.163	0.159	-0.119	-0.310	-0.302	-0.001	0.249	-0.089	0.289	0.174	-0.010	0.166	0.163	0.189	0.075
Other identified genes with known or suspected DNA repair function	0.193	0.096	0.089	0.059	0.152	-0.022	-0.042	0.035	-0.001	-0.231	0.050	0.046	0.030	0.185	0.092	0.172	-0.047	-0.227
Other BER and strand break joining factors/Other BER and strand break joining factors	-0.244	-0.307	-0.273	-0.195	-0.085	0.040	-0.280	0.084	0.021	0.512	0.072	-0.096	0.127	0.128	0.291	0.160	-0.111	-0.095
Fanconi anemia/Tolerance and repair of DNA crosslinks and other adducts in DNA	-0.590	-0.345	-0.374	-0.323	-0.413	0.273	0.098	0.283	0.098	0.148	-0.137	-0.320	0.116	-0.146	-0.172	-0.393	-0.186	-0.064
Homologous recombination	-0.492	-0.490	-0.259	-0.142	-0.186	-0.161	-0.244	-0.156	0.199	0.495	0.008	0.009	0.364	0.096	0.219	0.036	0.028	0.126
Repair of DNA-topoisomerase crosslinks	-0.215	-0.258	-0.111	0.012	-0.044	-0.078	-0.314	-0.160	0.347	-0.060	0.188	0.055	0.296	-0.006	0.116	-0.039	-0.019	-0.062
Chromatin Structure and Modification	-0.220	-0.372	-0.427	-0.063	-0.178	-0.088	-0.379	-0.113	0.074	0.500	-0.073	0.025	0.262	0.014	0.132	0.104	-0.067	0.128
Direct reversal of damage	0.235	-0.062	-0.229	0.168	0.138	-0.092	-0.166	-0.225	0.042	0.146	0.025	0.142	0.109	0.013	0.203	0.142	0.106	0.066
Mismatch excision repair (MMR)	-0.098	0.281	0.098	-0.008	-0.183	-0.121	0.131	-0.009	-0.070	-0.115	-0.104	-0.058	-0.053	-0.153	-0.246	-0.220	0.037	0.095
Non-homologous end-joining	0.017	0.106	0.146	-0.100	0.099	0.285	0.281	0.373	-0.058	-0.298	0.060	-0.231	-0.230	0.030	0.004	-0.073	-0.170	-0.257
Modulation of nucleotide pools	0.120	-0.010	-0.194	0.172	0.236	-0.007	-0.064	-0.220	0.081	0.050	-0.118	0.228	0.192	0.033	0.163	-0.154	0.230	-0.096
DNA polymerases (catalytic subunits)	-0.046	0.012	-0.008	-0.211	-0.278	-0.031	0.128	0.148	-0.199	0.223	0.007	-0.157	-0.265	-0.041	0.077	-0.040	-0.135	0.362
Genes defective in diseases associated with sensitivity to DNA damaging agents	0.052	0.256	0.235	-0.038	-0.066	0.069	0.449	0.183	-0.056	-0.153	-0.042	-0.123	-0.296	-0.226	-0.213	-0.276	0.048	0.052
Esting and processing nucleases	0.023	-0.195	-0.302	0.196	-0.053	0.111	-0.169	-0.115	0.158	0.126	-0.049	0.952	0.249	-0.293	0.100	-0.007	0.021	0.520
Other conserved DNA damage response genes	-0.206	0.050	0.129	-0.165	-0.188	0.056	0.299	0.241	-0.110	-0.113	0.019	-0.172	-0.179	-0.022	-0.146	-0.196	-0.118	-0.014

B

### DNA Repair-related GO classifications (TCGA)

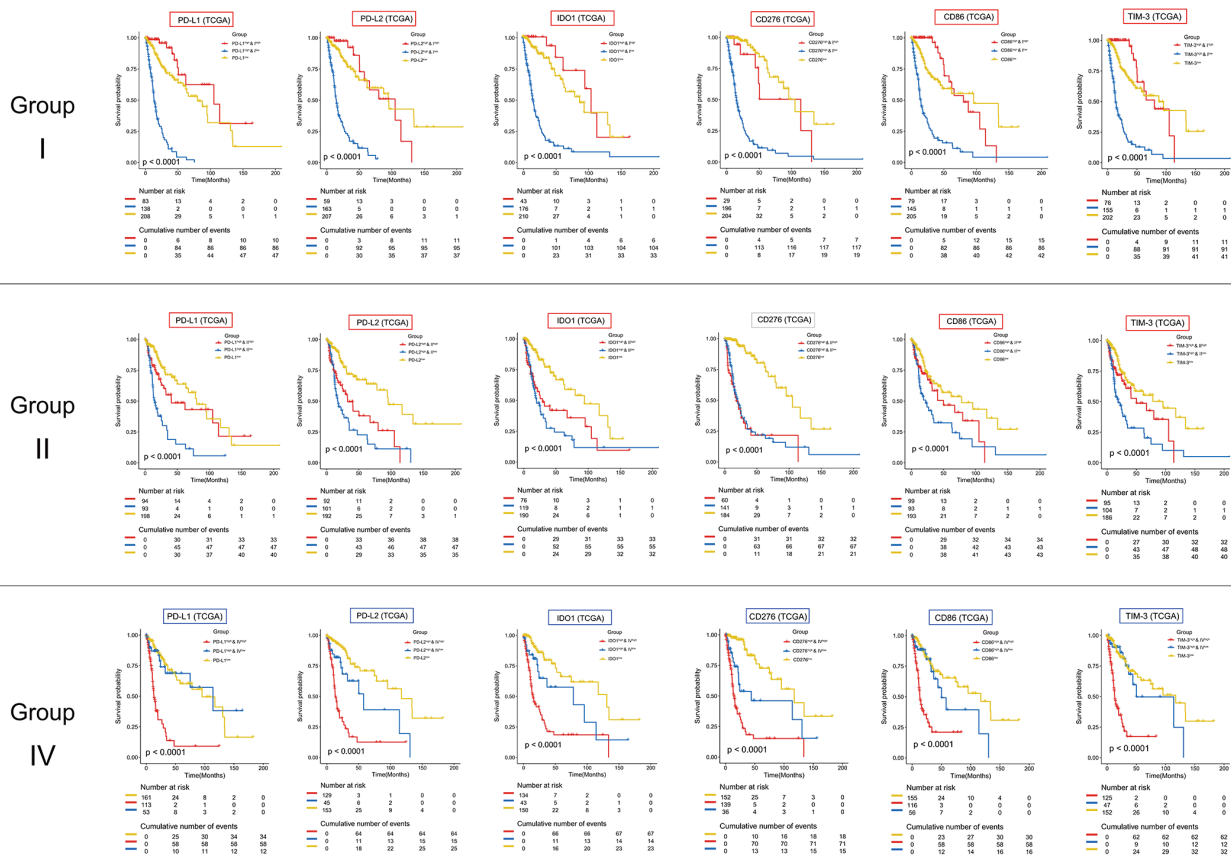


C

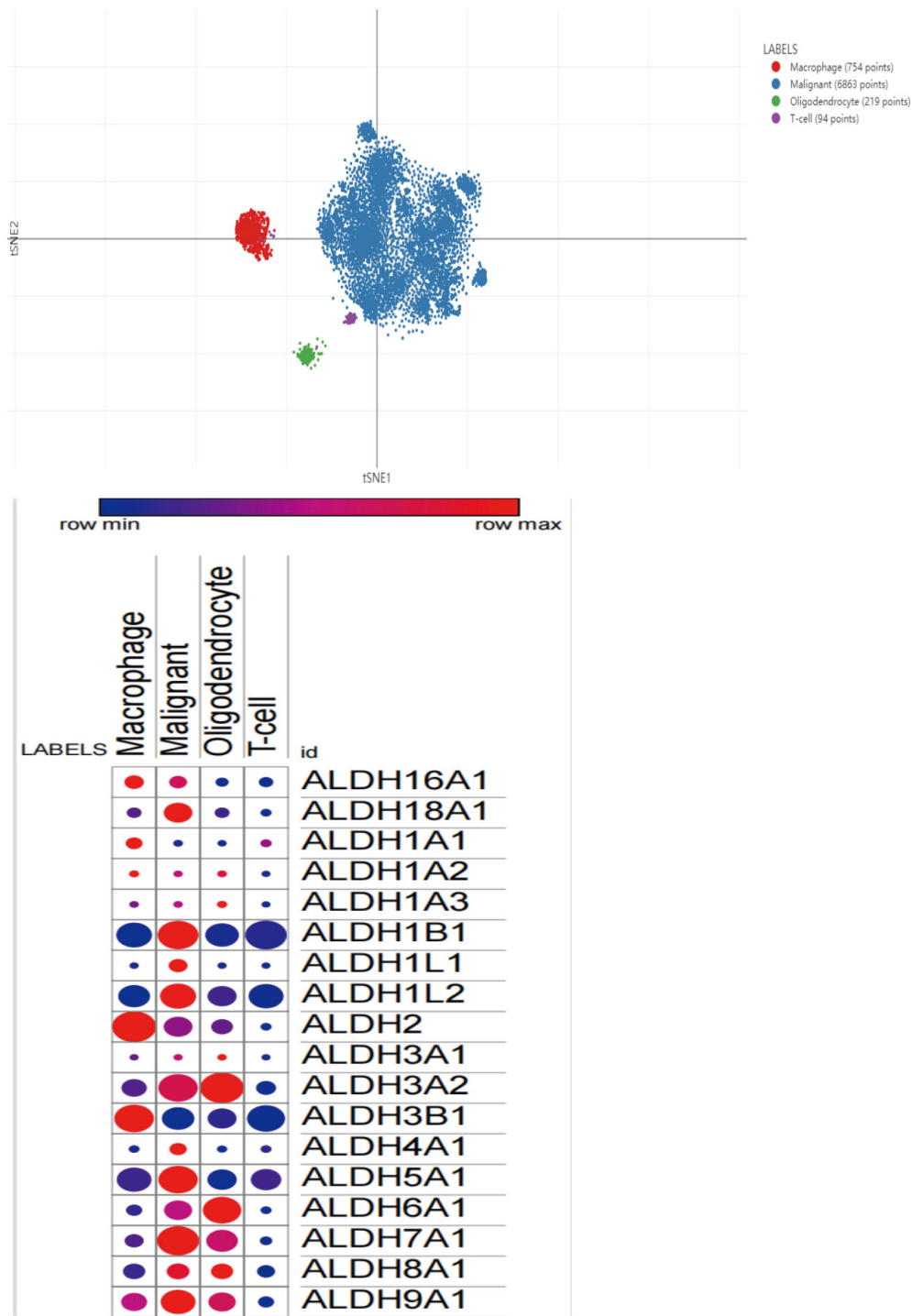
### Immune functions (TCGA)

Group	I			II		III			IV		V							
	ALDH2	ALDH5A1	ALDH6A1	ALDH1A1	ALDH8A1	ALDH1B1	ALDH1L2	ALDH18A1	ALDH1A3	ALDH16A1	ALDH1A2	ALDH1L1	ALDH3A1	ALDH3A2	ALDH3B1	ALDH4A1	ALDH7A1	ALDH8A1
immune response	-0.417	-0.591	-0.383	-0.078	0.029	-0.095	-0.421	-0.137	0.383	0.473	0.268	0.116	0.487	0.354	0.684	0.217	-0.029	-0.205
B cell activation involved in immune response	-0.842	-0.678	-0.434	-0.228	-0.105	-0.008	-0.341	0.019	0.312	0.491	0.175	-0.042	0.469	0.317	0.524	0.043	-0.116	-0.227
T cell activation involved in immune response	-0.329	-0.519	-0.302	-0.022	0.027	-0.134	-0.353	-0.120	0.300	0.469	0.327	0.073	0.411	0.328	0.690	0.267	-0.065	-0.090
cytokine production involved in immune response	-0.404	-0.583	-0.320	-0.028	0.061	-0.169	-0.436	-0.198	0.371	0.487	0.317	0.112	0.902	0.346	0.652	0.272	0.001	-0.122
cytokine secretion involved in immune response	-0.383	-0.524	-0.284	-0.043	0.082	-0.128	-0.387	-0.150	0.309	0.439	0.286	0.084	0.433	0.325	0.588	0.288	-0.039	-0.107
immune response to tumor cell	-0.385	-0.548	-0.425	-0.045	0.039	-0.033	-0.427	-0.109	0.285	0.535	0.222	0.033	0.417	0.289	0.508	0.195	-0.034	-0.054
leukocyte activation involved in immune response	-0.386	-0.584	-0.355	-0.047	0.055	-0.096	-0.444	-0.159	0.359	0.439	0.274	0.142	0.497	0.344	0.697	0.246	-0.006	-0.204
natural killer cell mediated immune response to tumor cell	-0.324	-0.506	-0.349	0.020	-0.010	-0.112	-0.443	-0.191	0.276	0.560	0.248	0.083	0.367	0.234	0.600	0.289	-0.035	0.024
T cell mediated immune response to tumor cell	0.173	0.223	-0.006	-0.133	0.188	0.223	0.229	0.321	-0.417	-0.081	-0.207	-0.200	-0.284	0.171	-0.281	-0.113	-0.116	0.004
natural killer cell activation involved in immune response	-0.104	-0.312	-0.078	0.179	0.074	-0.259	-0.311	-0.276	0.340	0.266	0.284	0.192	0.330	0.193	0.593	0.317	0.084	0.045
myeloid cell activation involved in immune response	-0.357	-0.567	-0.346	-0.034	0.074	-0.101	-0.457	-0.178	0.353	0.422	0.259	0.166	0.494	0.339	0.683	0.258	0.011	-0.206

**Figure S4** The ALDH groups are closely related to the state of genomic mutation. (A) The correlation coefficient between all ALDH isoforms and DNA repair function scores reveals that PARP enzymes that bind to DNA, NER-related and H/Catalyzes unwinding in precision complex are most positive related to the expression of ALDH isoforms in group I and group II and negative related to the expression of ALDH isoforms in group IV. In addition, Fanconi anemia/Tolerance and repair of DNA crosslinks and other adducts in DNA and Homologous recombination are most negative related to the expression of ALDH isoforms in group I and group II and positive related to the expression of ALDH isoforms in group IV. Red background represents a positive correlation. Blue background represents a negative correlation. Grey background represents no significant correlation. (B) Biological functions, most related DNA repair functions, is immune system process classification. The scale values in the graph represent the proportions of significant correlation biological functions in each biological function classifications. (C) The correlation coefficient matrix of ALDH isoforms and immune function scores. Red background represents a positive correlation. Blue background represents a negative correlation. Grey background represents no significant correlation.



**Figure S5** Prediction of therapeutic effect of immunological and enzymatic combined therapy. Kaplan-Meier survival curves reveal the therapeutic effect of immunological and enzymatic combined therapy. Red frame means elevating the expression of ALDH isoforms in the corresponding group can improve the prognosis of patients. Blue frame means knocking down the expression of ALDH isoforms in the corresponding group can improve the prognosis of patients. Grey frame means changes in ALDH isoforms expression in the corresponding group do not affect patient prognosis.



**Figure S6** Analysis on different sub-classify ALDH enzyme in distinct cell types based on scRNA-seq dataset from 28 patients with GBM. Cell subtypes clustering via t SNE method and distinct isoforms of ALDHs expression in different cell clusters based on single cell mRNA sequencing data. The color shows mRNA expression level compared to the average level in all cells. The size of plots represent the ratio of specific subtype of ALDH positive cells in distinct cell cluster.