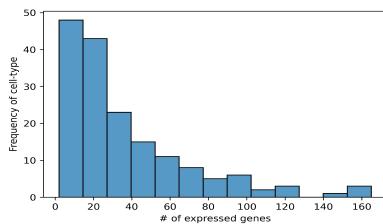
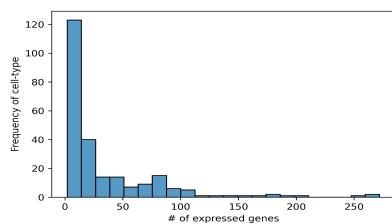


## Supplementary Figure S1

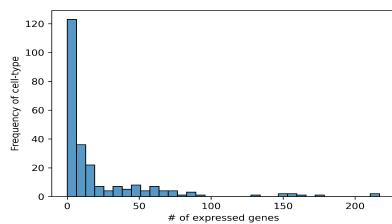
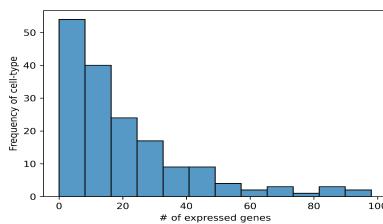
PanglaoDB



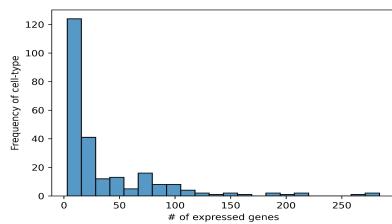
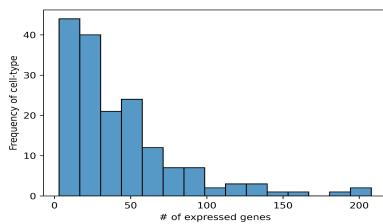
CellMarker



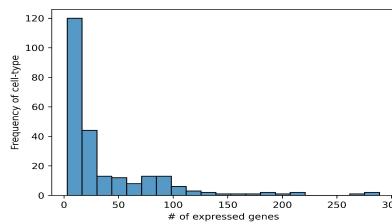
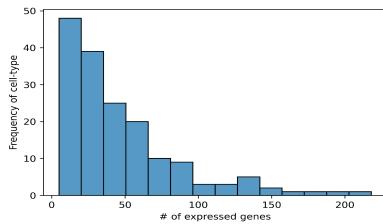
PBMC  
(Zheng et al., 2017)



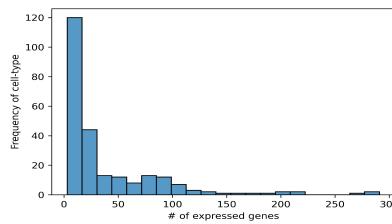
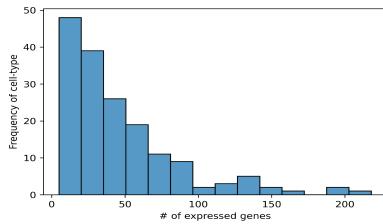
CellBench  
(Tian et al., 2019)



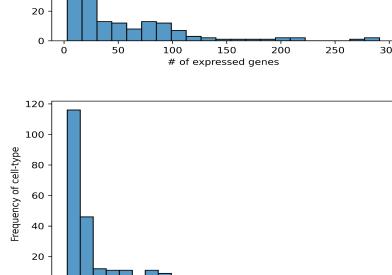
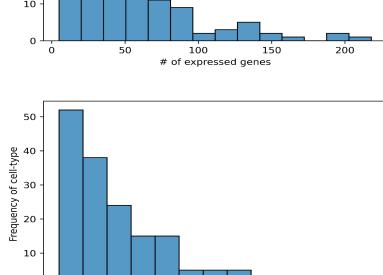
Pancreas  
(Baron et al., 2016)



Heart  
(Wang et al., 2020)



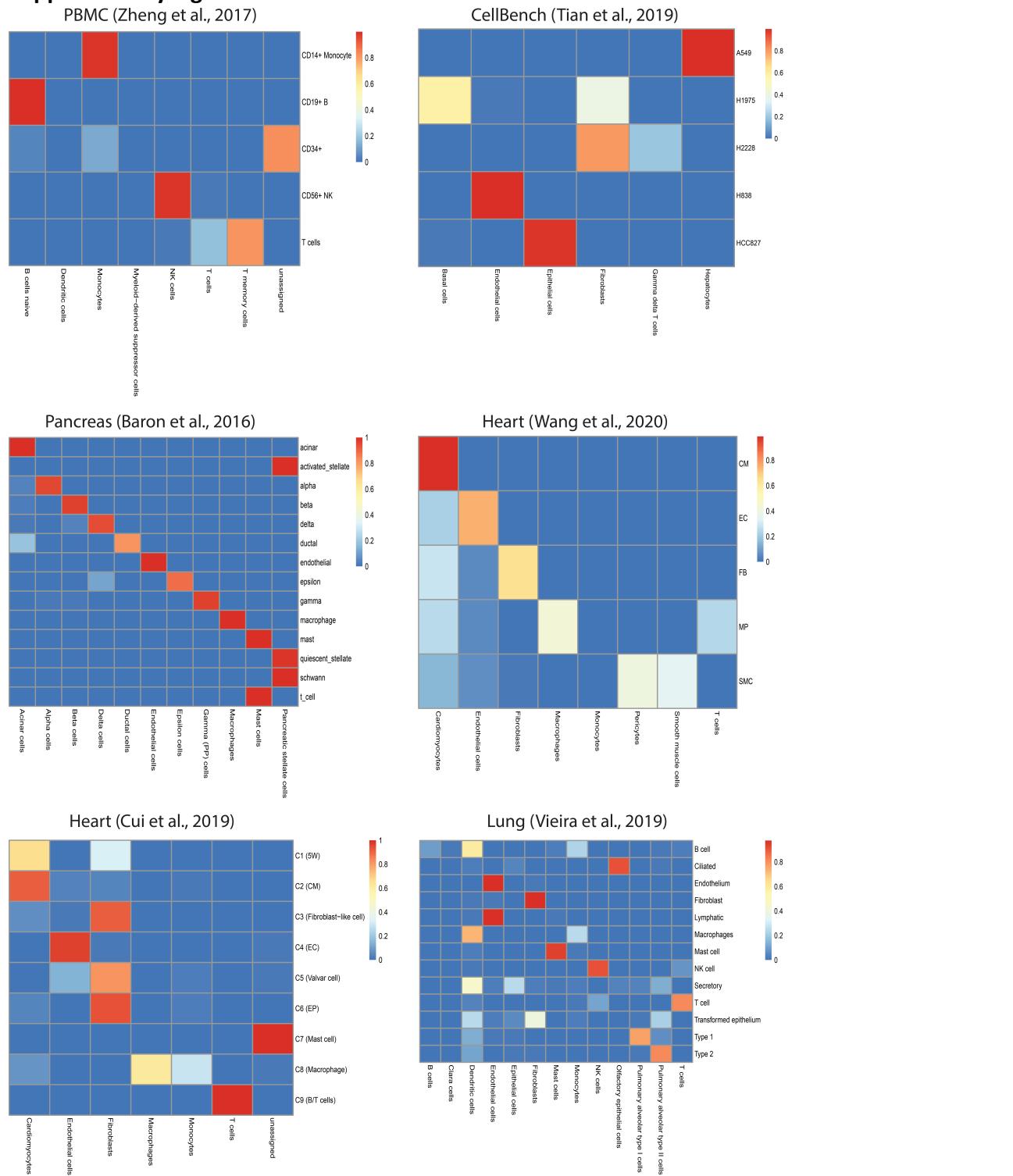
Heart  
(Cui et al., 2019)



Lung  
(Vieira et al., 2019)

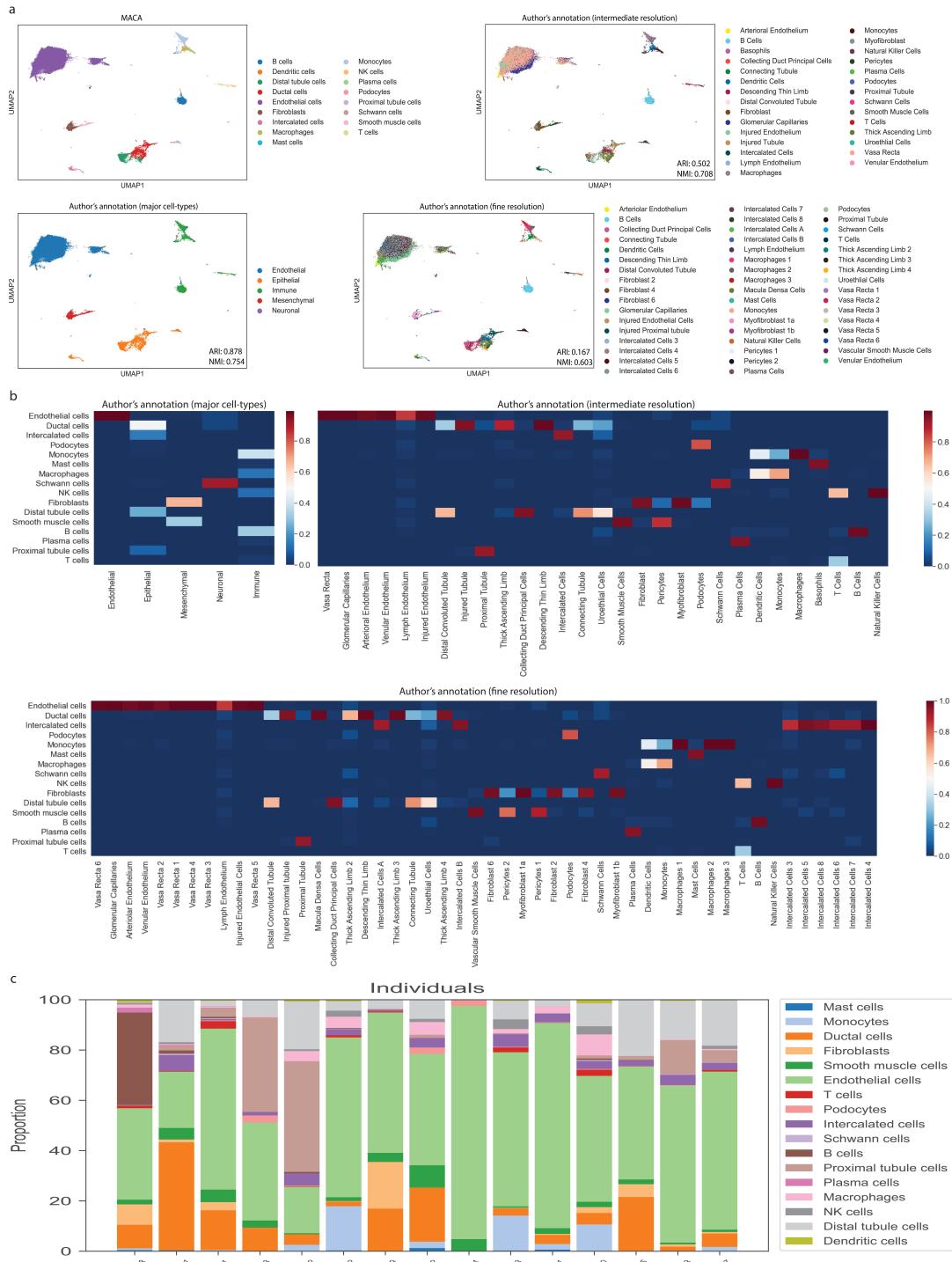
**Supplementary Figure S1.** Expressed marker genes of cell-types in PanglaoDB and CellMarker across 6 single cell datasets. Any marker genes in either PanglaoDB or CellMarker would be regarded as expressed genes if they have non-zero values in at least one cell. Histograms show the number of expressed marker genes against the number of cell-types.

## Supplementary Figure S2



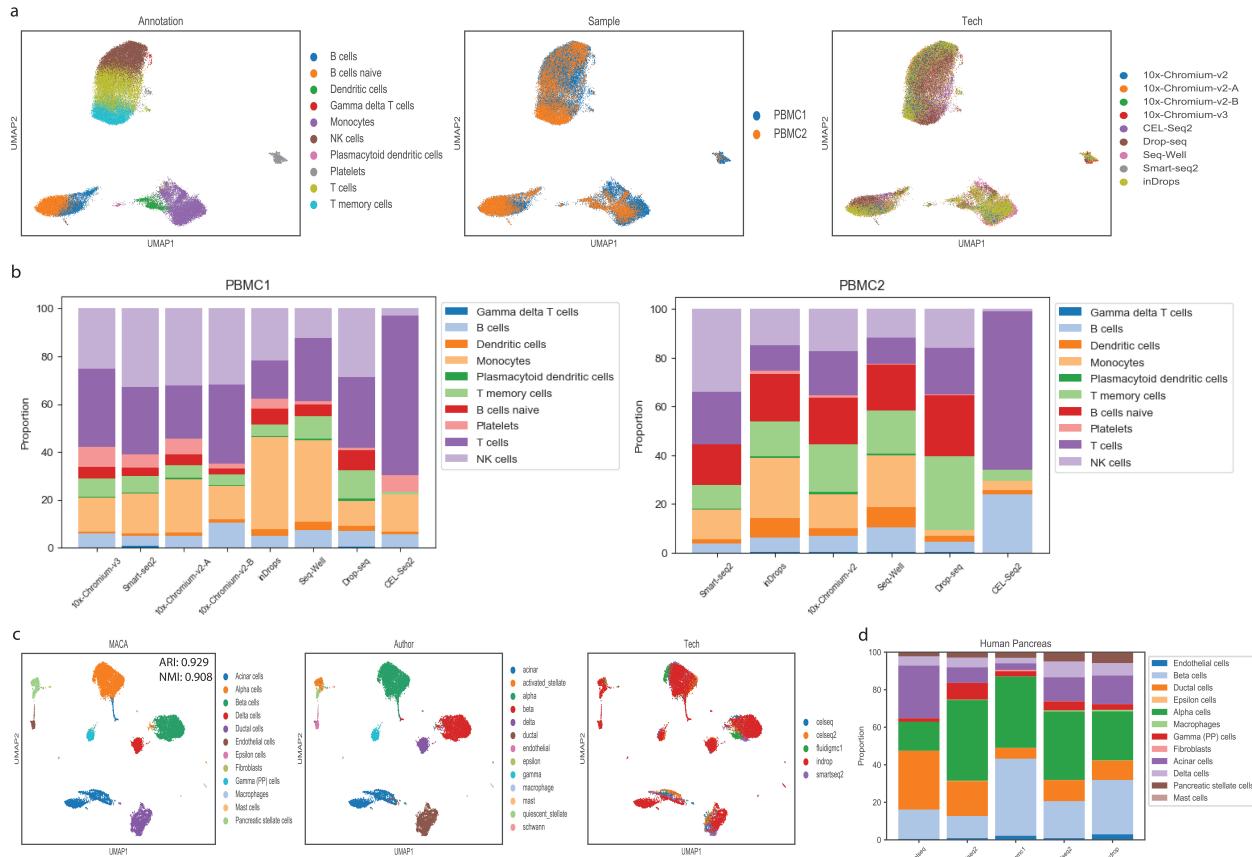
**Supplementary Figure S2.** Normalized confusion matrices between authors' annotation on y-axis and annotation by MACA on x-axis, for 6 single cell data used in this study. Each row stands for cell-type label in authors' annotation, and each column is cell-type label in MACA's annotation. Confusion matrix was normalized by dividing the sum of each row. A high value means that MACA annotation agrees well with authors' annotation.

## Supplementary Figure S3



**Supplementary Figure S3.** Integrated annotation of human kidney (CD10). a, MACA's annotation is shown in the top. Author reported 3 levels of annotations, with 5 major cell-types, 29 intermediate cell-types, and 50 fine cell-types. MACA's annotation is compared to all 3 resolutions by ARI and NMI. b, Confusion matrices between MACA's annotation and 3 author's annotations. Each column stands for cell-type label in authors' annotation, and each row is cell-type label in MACA's annotation. Confusion matrix was normalized by dividing the sum of each column. c, Cellular component analysis across different individuals.

## Supplementary Figure S4



**Supplementary Figure S4.** Integrated annotation of human PBMC and pancreas data across different single cell platforms. a, UMAP visualization of human PBMC data. Cells are colored according to annotation by MACA (left), source of sample (middle), and platform (right). UMAP dimension reduction is based on PlinerScore with PanglaoDB as marker database. b, cellular component analysis of human PBMC data. Proportion of each cell-type identified by MACA is calculated for each platform for two PBMC samples, separately. c, UMAP visualization of human pancreas data. Cells are colored according to annotation by MACA (left), author-reported cell-types (middle), and platform (right). UMAP dimension reduction is based on PlinerScore with PanglaoDB as marker database. d, cellular component analysis of human pancreas data. Proportion of each cell-type identified by MACA is calculated for each platform.

**Table S1**

Index	Species	Tissue	# of cells	# of major cell types by author	Technology	Reference	GEO	Source of processed data
1	Human	PBMC	20000	5	10X chromium	Zheng et al., 2017	NA	<a href="https://doi.org/10.5281/zenodo.3357167">https://doi.org/10.5281/zenodo.3357167</a>
2	Human	Cell lines	3803	5	10X chromium	Tian et al., 2019	GSE118767	<a href="https://doi.org/10.5281/zenodo.3357167">https://doi.org/10.5281/zenodo.3357167</a>
3	Human	Pancreas	8569	14	inDrop	Baron et al., 2016	GSE84133	<a href="https://doi.org/10.5281/zenodo.3357167">https://doi.org/10.5281/zenodo.3357167</a>
4	Human	Heart	6731	5	SMART-Seq2	Wang et al., 2020	GSE109816	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE109816">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE109816</a>
5	Human	Heart	3842	9	STRT-seq	Cui et al., 2019	GSE106118	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106118">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106118</a>
6	Human	Lung	10360	13	Drop-seq	Vieira et al., 2019	GSE130148	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE130148">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE130148</a>
7	Human	Left Atrium	75743	9	single nuclei RNA-seq	Tucker et al., 2020	NA	<a href="https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart">https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart</a>
8	Human	Right Atrium	82045	9	single nuclei RNA-seq	Tucker et al., 2020	NA	<a href="https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart">https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart</a>
9	Human	Left Ventricle	78264	9	single nuclei RNA-seq	Tucker et al., 2020	NA	<a href="https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart">https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart</a>
10	Human	Right Ventricle	51217	9	single nuclei RNA-seq	Tucker et al., 2020	NA	<a href="https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart">https://singlecell.broadinstitute.org/single_cell/study/SC498/transcriptional-and-cellular-diversity-of-the-human-heart</a>
11	Human	PBMC	3362	9	10X chromium v2	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
12	Human	PBMC	5172	9	10X chromium v2A	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
13	Human	PBMC	2657	9	10X chromium v2B	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
14	Human	PBMC	4038	8	10X chromium v3	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
15	Human	PBMC	5646	7	CEL-seq2	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
16	Human	PBMC	11095	9	Drop-seq	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
17	Human	PBMC	6038	7	Seq-Well	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
18	Human	PBMC	584	7	SMART-seq2	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
19	Human	PBMC	10710	9	inDrop	Ding et al., 2020	GSE132044	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132044</a>
20	Human	Pancreas	1004	13	CEL-seq	Grün et al., 2016	GSE81076	<a href="https://github.com/satijalab/seurat-data">https://github.com/satijalab/seurat-data</a>
21	Human	Pancreas	2285	14	CEL-seq2	Muraro et al., 2016	GSE85241	
22	Human	Pancreas	638	13	Fluidigm C1	Lawlor et al., 2017	GSE86469	
23	Human	Pancreas	2394	13	SMART-seq2	Segerstolpe et al., 2016	E-MTAB-5061	
24	Human	Kidney (CD10-)	53672	5	Smart-Seq2 and 10X Genomics 3	Kuppe et al., 2020	NA	<a href="https://doi.org/10.5281/zenodo.4059315">https://doi.org/10.5281/zenodo.4059315</a>

**Supplementary Table S1.** Summary of datasets used in this study. MACA was developed and its parameters for choosing the optimal scoring and underlying marker database were optimized on 6 datasets (PBMC, CellBench, Pancreas, Heart and Lung). Platform validation was conducted on human PBMC data across 9 single cell RNA-seq platforms and human pancreas data across 5 platforms. Furthermore, MACA's annotation resolution is compared with 3 different annotation resolutions by author in human kidney (CD10-) data. Fully optimized MACA was then tried on the human 4 chamber heart dataset obtained from Tucker et al., 2020.

**Table S2**

ARI	PBMC (Zheng et al., 2017)	CellBench (Tian et al., 2019)	Pancreas (Baron et al., 2016)	Heart (Wang et al., 2020)	Heart (Cui et al., 2019)	Lung (Vieira et al., 2019)
0.2	0.95	0.92	0.90	0.71	0.58	0.44
0.3	0.95	0.92	0.90	0.71	0.58	0.44
0.4	0.95	0.92	0.90	0.71	0.60	0.44
0.5	0.95	0.92	0.90	0.71	0.61	0.45
0.6	0.95	0.92	0.90	0.71	0.58	0.44
0.7	0.95	0.92	0.92	0.71	0.58	0.48
0.8	0.95	0.93	0.91	0.67	0.53	0.44
0.9	0.95	0.41	0.89	0.67	0.55	0.26
NMI	PBMC (Zheng et al., 2017)	CellBench (Tian et al., 2019)	Pancreas (Baron et al., 2016)	Heart (Wang et al., 2020)	Heart (Cui et al., 2019)	Lung (Vieira et al., 2019)
0.2	0.89	0.92	0.88	0.59	0.63	0.60
0.3	0.89	0.92	0.88	0.59	0.63	0.60
0.4	0.89	0.92	0.88	0.59	0.63	0.60
0.5	0.89	0.92	0.88	0.59	0.62	0.59
0.6	0.90	0.92	0.88	0.59	0.60	0.60
0.7	0.90	0.92	0.90	0.59	0.60	0.59
0.8	0.90	0.94	0.89	0.57	0.57	0.58
0.9	0.89	0.67	0.87	0.57	0.58	0.53

**Supplementary Table S2.** Evaluation of threshold of consensus between Label 1 and Label 2. The first column is the threshold used. Clusters in Label 2 will be assigned a cell-type label in Label 1, only when the proportion of cells in that cluster sharing the same cell-type label is larger than the threshold.

**Table S3**

	PBMC (Zheng et al., 2017)	CellBench (Tian et al., 2019)	Pancreas (Baron et al., 2016)	Heart (Wang et al., 2020)	Heart (Cui et al., 2019)	Lung (Vieira et al., 2019)
Runtime of Our method (in M)	1.5	1	1	1	1	1
Runtime of SCINA (in M)	18	5	16	8.5	8	19
Runtime of CellAssign (in M)	NA	30	140	120	180	160
Runtime of Cell-ID (in M)	15	2	5	5	2	10
Runtime of scCATCH (in M)	7.5	NA	1.5	1	1	6.5
	Left Atrium (Tucker et al., 2020)	Right Atrium (Tucker et al., 2020)	Left Ventricle (Tucker et al., 2020)	Right Ventricle (Tucker et al., 2020)		
Runtime of Our method (in M)	3.5	5.5	4	2.5		
Runtime of SCINA (in M)	NA	NA	NA	NA		
Runtime of CellAssign (in M)	NA	NA	NA	NA		
Runtime of Cell-ID (in M)	NA	NA	NA	NA		
Runtime of scCATCH (in M)	NA	NA	NA	NA		

**Supplementary Table S3.** Runtime of 5 annotation tools across 10 datasets. MACA was much shorter than other methods tested in this benchmark, running a workstation with 16-core CPU and 64GB memory.

**Table S4**

	PBMC (Zheng et al., 2017)	CellBench (Tian et al., 2019)		Pancreas (Baron et al., 2016)		Heart (Wang et al., 2020)		Heart (Cui et al., 2019)		Lung (Vieira et al., 2019)	
PanglaoDB+PlinerScore	<b>0.93</b>		0.99		<b>0.97</b>		<b>0.90</b>		<b>0.96</b>		<b>0.85</b>
PanglaoDB+AUCell	0.67		NA		0.89		0.85		0.89		0.57
PanglaoDB+CIM	0.77		<b>1.00</b>		<b>0.97</b>		0.83		0.91		0.73
PanglaoDB+DingScore	0.91		0.98		0.89		<b>0.92</b>		0.93		0.83
CellMarker+PlinerScore	<b>0.96</b>		0.98		0.93		<b>0.95</b>		0.90		0.83
CellMarker+AUCell	0.48		0.96		0.60		0.74		0.65		0.44
CellMarker+CIM	0.69		<b>1.00</b>		0.92		0.78		0.84		0.55
CellMarker+DingScore	0.76		0.99		0.90		0.62		<b>0.97</b>		NA
SCINA	0.80		0.74		0.91		0.43		0.90		0.78
CellAssign	NA		0.74		0.94		0.41		0.77		0.77
Cell-ID	0.82		0.36		0.80		0.53		0.78		0.70
scCATCH	0.91		0.92		0.92		0.56		<b>0.94</b>		<b>0.89</b>
Authors' annotation	<b>0.97</b>		<b>1.00</b>		<b>0.98</b>		0.89		0.93		<b>0.90</b>

**Supplementary Table S4.** Mean accuracy of 5-fold SVM classifier. SVM classifiers were trained with labels from original reports, or generated by MACA, CellAssign, SCINA, Cell-ID, and scCATCH. Each data was split into 5 folds. Classifiers were trained on 4 folds, and they used the rest 1-fold to report accuracy. Results here came from the mean accuracy of 5-fold training. The highest accuracies obtained for that dataset is shown in bold. For most datasets, PanglaoDB+PlinerScore and authors' annotations achieve the highest accuracy in SVM classification.

**Table S5**

PBMC (Zheng et al., 2017)			CellBench (Tian et al., 2019)			Pancreas (Baron et al., 2016)			Heart (Wang et al., 2020)			Heart (Cui et al., 2020)			
# of cells	ARI	NMI	Threshold	# of cells	ARI	NMI	Threshold	# of cells	ARI	NMI	Threshold	# of cells	ARI	NMI	Threshold
20000	0.94	0.87	1/9	3803	0.92	0.92	1/9	8569	0.9	0.88	1/9	6731	0.71	0.59	1/9
20000	0.94	0.87	2/9	3803	0.92	0.92	2/9	8569	0.9	0.88	2/9	6731	0.71	0.59	2/9
20000	0.94	0.87	3/9	3803	0.92	0.92	3/9	8569	0.9	0.88	3/9	6730	0.71	0.59	3/9
19983	0.94	0.87	4/9	3792	0.92	0.92	4/9	8567	0.9	0.88	4/9	6681	0.72	0.6	4/9
19830	0.95	0.88	5/9	3615	0.93	0.94	5/9	8552	0.9	0.88	5/9	6580	0.74	0.62	5/9
17228	0.94	0.88	6/9	2634	0.99	0.98	6/9	8535	0.9	0.88	6/9	6397	0.75	0.63	6/9
14299	0.95	0.9	7/9	2623	1	0.99	7/9	8359	0.91	0.9	7/9	6047	0.78	0.67	7/9
11434	0.95	0.9	8/9	2623	1	0.99	8/9	8183	0.91	0.9	8/9	5717	0.84	0.72	8/9
8469	0.94	0.91	9/9	2616	1	0.99	9/9	8008	0.93	0.92	9/9	5205	0.88	0.77	9/9
Lung (Vieira et al., 2019)			LA (Tucker et al., 2020)			RA (Tucker et al., 2020)			LV (Tucker et al., 2020)			RV (Tucker et al., 2020)			
# of cells	ARI	NMI	Threshold	# of cells	ARI	NMI	Threshold	# of cells	ARI	NMI	Threshold	# of cells	ARI	NMI	Threshold
10360	0.45	0.57	1/9	75743	0.69	0.61	1/9	82045	0.69	0.59	1/9	78264	0.85	0.71	1/9
10360	0.45	0.57	2/9	75743	0.69	0.61	2/9	82045	0.69	0.59	2/9	78264	0.85	0.71	2/9
10356	0.45	0.57	3/9	75738	0.69	0.61	3/9	81913	0.69	0.59	3/9	78260	0.85	0.71	3/9
10245	0.45	0.58	4/9	75451	0.69	0.62	4/9	81219	0.7	0.6	4/9	77985	0.85	0.71	4/9
9891	0.46	0.59	5/9	74386	0.7	0.63	5/9	79642	0.71	0.62	5/9	76786	0.86	0.73	5/9
9063	0.48	0.6	6/9	71512	0.72	0.65	6/9	77043	0.74	0.64	6/9	74658	0.88	0.76	6/9
8032	0.5	0.63	7/9	68907	0.74	0.68	7/9	73846	0.76	0.67	7/9	71417	0.9	0.8	7/9
7412	0.54	0.65	8/9	66298	0.77	0.7	8/9	70652	0.79	0.7	8/9	69049	0.92	0.83	8/9
6443	0.56	0.67	9/9	62134	0.79	0.72	9/9	65205	0.82	0.73	9/9	64215	0.94	0.87	9/9

**Supplementary Table S5.** ARI/NMI against the number of cells to be annotated when different thresholds are used. In this study, we performed 9 different settings of clustering to get 9 different cell-type annotations. MACA adopts voting for ensemble annotation. Cells are finally annotated as the cell-type that gets the greatest number of votes. Increasing threshold of voting (the number of votes) to remove cells with less consistent votes inevitably reduces the number of cells to be included in the final annotation. However, comparing annotations of rest of cells by MACA shows increasing agreement with authors' annotations (increases of ARI and NMI).

**Table S6**

ARI	LV	RV	LA	RA
MACA	0.84	0.81	0.69	0.69
NMI	LV	RV	LA	RA
MACA	0.70	0.63	0.61	0.58
SVM_Accuracy	LV	RV	LA	RA
MACA	0.85	0.85	0.85	0.82
Authors' annotation	0.90	0.88	0.86	0.83

**Supplementary Table S6.** Performance of MACA in 4 human-chamber single-nuclei RNA-seq datasets, measured by ARI, NMI, and Accuracy of SVM classifier. RA: right atrium; LA: left atrium; RV: right ventricle; LV: left ventricle. The final MACA setting is using PanglaoDB as marker reference and PlinerScore as cell-type scoring method. The performance in human 4 chamber data (Tucker et al., 2020) is quantified by ARI and NMI against author's annotations (top 4 rows). Both MACA's and author's annotation were used to train SVM classifiers. Datasets were split into 5 folds. SVM classifiers were trained on 4-fold data and tested on the rest 1-fold. Means of accuracies were reported to show how reasonable MACA's and authors' annotations are (bottom 2 rows).

**Table S7**

(rows are author's annotation, and columns are MACA's annotation)										
RA	Fibroblasts	Cardiomyocytes	Endothelial cells	Pericytes	Macrophages	Smooth muscle cells	unassigned	T memory cells	Adipocytes	
Fibroblast	25946	2040	464	187	109	75	138	14	41	
Cardiomyocyte	1571	30642	234	65	94	103	55	9	20	
Macrophage	1256	268	89	2	2764	4	316	19	2	
Endothelium	1457	788	6764	77	21	25	40	4	2	
Pericyte	556	118	167	1833	6	238	92	5	3	
Vascular Smooth Muscle	168	104	36	194	12	1096	23	1	2	
Neuronal	229	174	50	6	14	10	130	2	6	
Adipocyte	208	27	12	4	4	2	5	0	442	
Lymphocyte	38	22	27	1	27	4	22	220	0	
LA	Cardiomyocytes	Adipocytes	Smooth muscle cells	Fibroblasts	Mesangial cells	Endothelial cells	Macrophages	Pericytes	T memory cells	unassigned
Cardiomyocyte	23484	64	53	1751	9	205	145	63	25	126
Adipocyte	228	2810	1	415	2	7	15	4	0	12
Vascular Smooth Muscle	252	1	2128	340	416	84	17	183	2	37
Macrophage	499	5	7	1103	3	102	3530	18	20	84
Fibroblast	1711	67	38	25039	23	284	119	135	12	177
Endothelium	258	0	6	184	1	3470	39	27	3	88
Pericyte	429	7	147	480	12	334	35	3424	9	266
Neuronal	118	10	5	118	1	80	4	10	0	7
Lymphocyte	29	0	0	31	0	12	69	0	154	21
RV	Cardiomyocytes	Adipocytes	Fibroblasts	Endothelial cells	Macrophages	Pericytes	Smooth muscle cells	unassigned	T memory cells	
Cardiomyocyte	25291	26	301	132	54	51	25	149	0	
Adipocyte	220	2724	365	41	60	11	2	282	0	
Endothelium	363	18	1260	3745	39	34	2	158	0	
Fibroblast	151	54	8668	174	42	69	2	700	0	
Macrophage	177	16	866	169	1506	6	0	155	1	
Pericyte	108	5	249	316	9	1566	46	354	0	
Neuronal	14	11	53	11	2	16	1	7	1	
Vascular Smooth Muscle	13	1	27	13	2	60	119	23	0	
Lymphocyte	8	0	8	18	12	1	0	13	21	
LV	Cardiomyocytes	Adipocytes	Macrophages	Fibroblasts	Smooth muscle cells	Endothelial cells	Pericytes	T memory cells	unassigned	Myofibroblasts
Cardiomyocyte	36855	1	97	397	29	296	49	10	239	2
Adipocyte	64	450	7	73	1	5	3	0	152	0
Macrophage	279	2	3099	669	0	248	17	30	137	0
Fibroblast	254	4	77	14514	4	567	192	13	1115	1
Vascular Smooth Muscle	14	0	3	11	278	20	48	1	7	4
Pericyte	233	1	24	446	166	335	5900	9	511	28
Endothelium	403	0	95	354	1	7871	76	12	238	0
Neuronal	41	0	41	164	1	50	34	14	133	0
Lymphocyte	22	0	76	20	0	12	3	573	39	0

**Supplementary Table S7a.** Confusion matrix between author-reported label and MACA's annotation in 4 human-chamber sing-nuclei RNA-seq datasets. RA: right atrium; LA: left atrium; RV: right ventricle; LV: left ventricle. Each row represents the number of cells annotated by author, and each column is MACA's annotation.

(rows are author's annotation, and columns are MACA's annotation)										
RA	Fibroblasts	Cardiomyocytes	Endothelial cells	Pericytes	Macrophages	Smooth muscle cells	unassigned	T memory cells	Adipocytes	
Fibroblast	0.8943	0.0703	0.0160	0.0064	0.0058	0.0026	0.0048	0.0005	0.0014	
Cardiomyocyte	0.0479	0.9344	0.0071	0.0020	0.0029	0.0031	0.0017	0.0003	0.0006	
Macrophage	0.2661	0.0568	0.0189	0.0004	0.5856	0.0008	0.0669	0.0040	0.0004	
Endothelium	0.1587	0.0859	0.7370	0.0084	0.0023	0.0027	0.0044	0.0004	0.0002	
Vascular Smooth Muscle	0.1027	0.0636	0.0220	0.1186	0.0073	0.6699	0.0141	0.0006	0.0012	
Neuronal	0.3688	0.2802	0.0805	0.0097	0.0225	0.0161	0.2093	0.0032	0.0097	
Adipocyte	0.2955	0.0384	0.0170	0.0057	0.0057	0.0028	0.0071	0.0000	0.6278	
Lymphocyte	0.1053	0.0609	0.0748	0.0028	0.0748	0.0111	0.0609	0.6094	0.0000	
LA	Cardiomyocytes	Adipocytes	Smooth muscle cells	Fibroblasts	Mesangial cells	Endothelial cells	Macrophages	Pericytes	T memory cells	unassigned
Cardiomyocyte	0.9058	0.0025	0.0020	0.0675	0.0003	0.0079	0.0056	0.0024	0.0010	0.0049
Adipocyte	0.0653	0.8042	0.0003	0.1188	0.0006	0.0020	0.0043	0.0011	0.0000	0.0034
Vascular Smooth Muscle	0.0728	0.0003	0.6150	0.0983	0.1202	0.0243	0.0049	0.0529	0.0006	0.0107
Macrophage	0.0929	0.0009	0.0013	0.2054	0.0006	0.0190	0.6572	0.0034	0.0037	0.0156
Fibroblast	0.0620	0.0024	0.0014	0.0070	0.0008	0.0103	0.0092	0.0049	0.0004	0.0064
Endothelium	0.0633	0.0000	0.0015	0.0451	0.0002	0.8513	0.0096	0.0066	0.0007	0.0216
Pericyte	0.0834	0.0014	0.0286	0.0933	0.0023	0.0649	0.6658	0.0017	0.0517	
Neuronal	0.3343	0.0283	0.0142	0.3348	0.0028	0.2266	0.0113	0.0283	0.0000	0.0198
Lymphocyte	0.0918	0.0000	0.0000	0.0981	0.0000	0.0380	0.2184	0.0000	0.4873	0.0665
RV	Cardiomyocytes	Adipocytes	Fibroblasts	Endothelial cells	Macrophages	Pericytes	Smooth muscle cells	unassigned	T memory cells	
Cardiomyocyte	0.9716	0.0010	0.0116	0.0051	0.0021	0.0020	0.0010	0.0057	0.0000	
Adipocyte	0.0594	0.7352	0.0985	0.0111	0.0162	0.0030	0.0005	0.0761	0.0000	
Endothelium	0.0646	0.0032	0.2242	0.6665	0.0069	0.0061	0.0004	0.0281	0.0000	
Fibroblast	0.0153	0.0055	0.8791	0.0176	0.0043	0.0070	0.0002	0.0710	0.0000	
Macrophage	0.0611	0.0055	0.2990	0.0584	0.5200	0.0021	0.0000	0.0535	0.0003	
Pericyte	0.0407	0.0019	0.0939	0.1191	0.0034	0.5903	0.0173	0.1334	0.0000	
Neuronal	0.1207	0.0948	0.4569	0.9498	0.0172	0.1379	0.0086	0.0603	0.0086	
Vascular Smooth Muscle	0.0504	0.0039	0.1047	0.0504	0.0078	0.0236	0.4612	0.0891	0.0000	
Lymphocyte	0.0988	0.0000	0.0988	0.2222	0.1481	0.0123	0.0000	0.1605	0.2593	
LV	Cardiomyocytes	Adipocytes	Macrophages	Fibroblasts	Smooth muscle cells	Endothelial cells	Pericytes	T memory cells	unassigned	Myofibroblasts
Cardiomyocyte	0.9705	0.0000	0.0026	0.0105	0.0008	0.0078	0.0013	0.0003	0.0063	0.0001
Adipocyte	0.0848	0.5960	0.0093	0.0967	0.0013	0.0066	0.0040	0.0000	0.2013	0.0000
Macrophage	0.0623	0.0004	0.6916	0.1493	0.0000	0.0553	0.0288	0.0067	0.0306	0.0000
Fibroblast	0.0152	0.0002	0.0046	0.8670	0.0002	0.0339	0.0115	0.0008	0.0666	0.0001
Vascular Smooth Muscle	0.0363	0.0000	0.0078	0.0285	0.7202	0.0518	0.1244	0.0026	0.0181	0.0104
Pericyte	0.0304	0.0001	0.0031	0.0583	0.0217	0.0438	0.7709	0.0012	0.0668	0.0037
Endothelium	0.0445	0.0000	0.0105	0.0391	0.0001	0.8697	0.0084	0.0013	0.0263	0.0000
Neuronal	0.0858	0.0000	0.0858	0.3431	0.0021	0.1046	0.0711	0.0293	0.2782	0.0000
Lymphocyte	0.0295	0.0000	0.1020	0.0268	0.0000	0.0161	0.0040	0.7691	0.0523	0.0000

**Supplementary Table S7b.** Confusion matrix between author-reported label and MACA's annotation in 4 human-chamber sing-nuclei RNA-seq datasets. RA: right atrium; LA: left atrium; RV: right ventricle; LV: left ventricle. From Supplementary Table S7a, each entry is divided by the sum of row, which the total number of cells of each cell-type by author's annotation. Supplementary Table S7b shows percentages of each cell-type in author's annotation but in all other cell-types in MACA's annotation.