Table S1. Summary of existing gene-expression deconvolution methods utilizing single-cell sequencing data. The comparisons include single-cell input format, marker-gene requirement, and handling of batch effects.

Method	**Single-cell input**	**Marker-gene requirement**	**Adjusting for technical variation between bulk and single-cell expression.**	**Batch effect correction**
MuSiC (Wang et al., Nature Communications, 2019)	Read count matrix (ExpressionSet); RPKM if estimates of cell-type specific total RNA abundance can be provided.	Optional.	No.	No.
BseqSC (Baron et al., Cell Systems, 2016)	Read count matrix (ExpressionSet).	Yes	No.	No.
Bisque (Jew et al., bioRxiv, 2019)	Read count matrix, (ExpressionSet, Seurat Object).	Optional (required when scRNA-seq is not available).	Gene-specific transformations of bulk expression levels using observed bulk samples and *in silico* generated pseudo-bulk samples to account for biases in sequencing technologies.	No.
CIBERSORTx (Newman et al., Nature Biotechnology, 2019)	TPM, RPKM, FPKM, raw read counts. Regardless of the input, the signature matrix and mixture files should be represented in the same normalization space (i.e., if single cell data input is TPM format, the bulk data input should also be TPM format).	No.	Bulk-mode batch correction (B-mode): removes technical differences between a signature matrix derived from purified/sorted bulk samples and an input set of mixture samples; Single-cell batch correction (S-mode): directly adjusts the signature matrix rather than the mixture matrix.	No.
DWLS (Tsoucas et al., Nature Communications, 2019)	Read count matrix.	Automatically selects marker genes according to input scRNA-seq data.	No.	No.
SCDC (this paper)	Read count matrix (ExpressionSet).	Optional.	Optional via a pre-processing step (same as Bisque and CIBERSORTx).	Implicitly addressed via ENSEMBLE.

Table S2. Benchmark of deconvolution results using simulated pseudo bulk samples of human pancreatic islets with only one single-cell reference. The pseudo bulk samples were constructed by summing up raw read counts across all single cells from (A) Baron et al., (B) Segerstolpe et al., and (C) Xin et al.. The performance of deconvolution was assessed using measurements on the deconvolved and true cell-type proportions. Clustering quality control (CQC) procedure resulted in improved deconvolution accuracy except for the cases involving Xin et al., which has a limited number of single cells per subject. Since only one single-cell reference was adopted in each run, SCDC was applied without ENSEMBLE but was still among the methods with top performance.

(A)

Pseudo -Bulk	Single-cell reference (number of cell types used)	Methods	mAD Before CQC	mAD After CQC	Pearson R Before CQC	Pearson R After CQC	Δ Pearson R
Baron et al.	Baron et al. (6 types)	MuSiC	0.027	0.026	0.97	0.966	0.00
		Bseq-SC	0.056	0.057	0.878	0.879	0.00
		Bisque	0.021	0.021	0.982	0.98	0.00
		CIBERSORTx	0.041	0.041	0.966	0.958	-0.01
		CIBERSORTx (B mode)	0.035	0.039	0.96	0.949	-0.01
		CIBERSORTx (S mode)	0.053	0.044	0.955	0.966	0.01
		DWLS	0.022	0.025	0.980	0.975	-0.01
		SCDC	**0.029**	**0.03**	**0.961**	**0.96**	**0.00**
	Segerstolpe et al. (6 types)	MuSiC	0.056	0.049	0.892	0.9	0.01
		Bseq-SC	0.073	0.08	0.808	0.789	-0.02
		Bisque	0.087	0.092	0.671	0.65	-0.02
		CIBERSORTx	0.084	0.076	0.852	0.87	0.02
		CIBERSORTx (B mode)	0.094	0.087	0.815	0.834	0.02
		CIBERSORTx (S mode)	0.065	0.044	0.972	0.973	0.00
		DWLS	0.127	0.115	0.365	0.486	0.12
		SCDC	**0.05**	**0.045**	**0.912**	**0.924**	**0.01**
	Xin et al. (4 types)	MuSiC	0.147	0.117	0.541	0.717	0.18
		Bseq-SC	0.144	0.154	0.434	0.435	0.00
		Bisque*	0.099	0.101	0.813	0.814	0.00
		CIBERSORTx	0.190	0.198	0.324	0.230	-0.09
		CIBERSORTx (B mode)	0.182	0.191	0.26	0.186	-0.07
		CIBERSORTx (S mode)	0.176	0.137	0.684	0.75	0.07
		DWLS	0.153	0.150	0.512	0.558	0.05
		SCDC	**0.06**	**0.085**	**0.964**	**0.973**	**0.01**

*Bisque only used part of the subjects due to the encountered error when including subjects with a missing cell-type.

(B)

Pseudo -Bulk	Single-cell reference (number of cell types used)	Methods	mAD Before CQC	mAD After CQC	Pearson R Before CQC	Pearson R After CQC	Δ Pearson R
Segerstolpe et al.	Baron et al. (6 types)	MuSiC	0.083	0.082	0.915	0.918	0.00
		Bseq-SC	0.089	0.085	0.895	0.899	0.00
		Bisque	0.094	0.096	0.642	0.627	-0.02
		CIBERSORTx	0.085	0.081	0.955	0.955	0.00
		CIBERSORTx (B mode)	0.09	0.084	0.953	0.955	0.00
		CIBERSORTx (S mode)	0.091	0.082	0.913	0.911	0.00
		DWLS	0.074	0.083	0.745	0.702	0.074
		SCDC w/o ENSEMBLE	**0.078**	**0.078**	**0.922**	**0.924**	**0.00**
	Segerstolpe et al. (6 types)	MuSiC	0.029	0.027	0.968	0.975	0.01
		Bseq-SC	0.064	0.06	0.875	0.898	0.02
		Bisque	0.035	0.035	0.963	0.962	0.00
		CIBERSORTx	0.06	0.061	0.949	0.951	0.00
		CIBERSORTx (B mode)	0.056	0.055	0.962	0.963	0.00
		CIBERSORTx (S mode)	0.053	0.063	0.953	0.932	-0.02
		DWLS	0.033	0.050	0.971	0.905	-0.07
		SCDC w/o ENSEMBLE	**0.032**	**0.031**	**0.961**	**0.969**	**0.01**
	Xin et al. (4 types)	MuSiC	0.096	0.091	0.933	0.94	0.01
		Bseq-SC	0.075	0.076	0.92	0.927	0.01
		Bisque*	0.063	0.064	0.944	0.943	0.00
		CIBERSORTx	0.083	0.084	0.910	0.901	-0.01
		CIBERSORTx (B mode)	0.07	0.073	0.927	0.92	-0.01
		CIBERSORTx (S mode)	0.139	0.195	0.582	0.294	-0.29
		DWLS	0.057	0.066	0.938	0.920	-0.02
		SCDC w/o ENSEMBLE	**0.113**	**0.133**	**0.97**	**0.965**	**-0.01**

*Bisque only used part of the subjects due to the encountered error when including subjects with a missing cell-type.

(C)

Pseudo -Bulk	Single-cell reference (number of cell types used)	Methods	mAD Before CQC	mAD After CQC	Pearson R Before CQC	Pearson R After CQC	Δ Pearson R
Xin et al.	Baron et al. (4 types)	MuSiC	0.189	0.188	0.71	0.711	0.00
		Bseq-SC	0.185	0.182	0.72	0.723	0.00
		Bisque	0.091	0.098	0.874	0.855	-0.02
		CIBERSORTx	0.180	0.177	0.733	0.736	0.00
		CIBERSORTx (B mode)	0.184	0.18	0.722	0.724	0.00
		CIBERSORTx (S mode)	0.169	0.153	0.765	0.781	0.02
		DWLS	0.093	0.086	0.874	0.875	0.00
		SCDC w/o ENSEMBLE	**0.188**	**0.187**	**0.71**	**0.711**	**0.00**
	Segerstolpe et al. (4 types)	MuSiC	0.071	0.075	0.92	0.915	-0.01
		Bseq-SC	0.154	0.153	0.757	0.758	0.00
		Bisque	0.078	0.076	0.902	0.897	-0.01
		CIBERSORTx	0.161	0.163	0.756	0.751	-0.01
		CIBERSORTx (B mode)	0.167	0.17	0.74	0.736	0.00
		CIBERSORTx (S mode)	0.089	0.063	0.903	0.94	0.04
		DWLS	0.093	0.067	0.824	0.927	0.10
		SCDC w/o ENSEMBLE	**0.067**	**0.072**	**0.929**	**0.925**	**0.00**
	Xin et al. (4 types)	MuSiC	0.035	0.035	0.976	0.976	0.00
		Bseq-SC	0.066	0.062	0.894	0.906	0.01
		Bisque*	0.066	0.066	0.924	0.925	0.00
		CIBERSORTx	0.050	0.051	0.959	0.957	0.00
		CIBERSORTx (B mode)	0.055	0.053	0.956	0.957	0.00
		CIBERSORTx (S mode)	0.054	0.056	0.957	0.958	0.00
		DWLS	0.023	0.026	0.991	0.990	0.00
		SCDC w/o ENSEMBLE	**0.072**	**0.081**	**0.938**	**0.925**	**-0.01**

*Bisque only used part of the subjects due to the encountered error when including subjects with a missing cell-type.

Table S3. Benchmark of deconvolution results using simulated pseudo bulk samples of human pancreatic islets with multiple single-cell references. The pseudo bulk samples were constructed by summing up raw read counts across all single cells from (A) Baron et al., (B) Segerstolpe et al., and (C) Xin et al.. Multiple single-cell references (from Baron et al., Segerstolpe et al., and Xin et al.) were adopted to aid deconvolution in three ways: (i) single-cell data from different sources were naively pooled and were used as input; (ii) single-cell data from different sources were first corrected for batch effect using MNN and then used as input; and (iii) ENSEMBLE framework was adopted to integrate deconvolution results from different single-cell references. The performance of deconvolution was assessed using measurements on the deconvolved and true cell-type proportions. Pooling single cells with and without batch correction generally resulted in unstable performance, while the ENSEMBLE method by SCDC resulted in Pearson correlations above 0.9 consistently across all three scenarios, highlighting its performance stability.

(A)

Pseudo-Bulk	Single-cell reference (number of cell types used)	Method adopted	mAD	Pearson R
Baron et al.	Pooled single cells from three studies (6 types)	MuSiC	0.123	0.418
		Bseq-SC	0.108	0.698
		Bisque	0.061	0.841
		CIBERSORTx	0.038	0.963
		CIBERSORTx (B mode)	0.053	0.922
		CIBERSORTx (S mode)	0.092	0.843
		DWLS	0.181	0.756
	MNN corrected single cells from three studies (6 types)	MuSiC	0.082	0.776
		Bseq-SC	0.098	0.707
		Bisque	0.050	0.878
		CIBERSORTx	0.032	0.986
		CIBERSORTx (B mode)	0.036	0.994
		CIBERSORTx (S mode)	0.046	0.915
		DWLS	0.187	-0.089
	ENSEMBLE from three studies (6 types)	SCDC (PearsonR on P): $w_1 = 1, w_2 = 0$	0.031	**0.952**
		SCDC (SpearmanR on Y): $w_1 = 1, w_2 = 0$	0.032	**0.949**

(B)

Pseudo-Bulk	Single-cell reference (number of cell types used)	Method adopted	mAD	Pearson R
Segerstolpe et al.	Pooled single cells from three studies (6 types)	MuSiC	0.056	0.897
		Bseq-SC	0.083	0.694
		Bisque	0.059	0.876
		CIBERSORTx	0.069	0.968
		CIBERSORTx (B mode)	0.071	0.961
		CIBERSORTx (S mode)	0.087	0.911
		DWLS	0.078	0.795
	MNN corrected single cells from three studies (6 types)	MuSiC	0.096	0.931
		Bseq-SC	0.131	0.537
		Bisque	0.047	0.948
		CIBERSORTx	0.102	0.838
		CIBERSORTx (B mode)	0.087	0.868
		CIBERSORTx (S mode)	0.058	0.951
		DWLS	0.189	-0.084
	ENSEMBLE from three studies (6 types)	SCDC (PearsonR on P): $w_1 = 0.3, w_2 = 0.7$	0.029	**0.977**
		SCDC (SpearmanR on Y): $w_1 = 0, w_2 = 1$	0.029	**0.971**

(C)

Pseudo-Bulk	Single-cell reference (number of cell types used)	Method adopted	mAD	Pearson R
Xin et al.	Pooled single cells from three studies (4 types)	MuSiC	0.068	0.906
		Bseq-SC	0.117	0.828
		Bisque	0.063	0.938
		CIBERSORTx	0.173	0.740
		CIBERSORTx (B mode)	0.178	0.727
		CIBERSORTx (S mode)	0.066	0.909
		DWLS	0.088	0.878
	MNN corrected single cells from three studies (4 types)	MuSiC	0.078	0.939
		Bseq-SC	0.199	0.691
		Bisque	0.067	0.943
		CIBERSORTx	0.084	0.891
		CIBERSORTx (B mode)	0.089	0.905
		CIBERSORTx (S mode)	0.082	0.917
		DWLS	0.305	-0.324
	ENSEMBLE from three studies (4 types)	SCDC (PearsonR on P): $w_1 = 0.15, w_2 = 0.15, w_3 = 0.7$	0.055	**0.960**
		SCDC (SpearmanR on Y): $w_1 = 0, w_2 = 0, w_3 = 1$	0.073	**0.934**

Table S4. ENSEMBLE weight selection results for the human pancreatic islet bulk samples. The weights are presented for 51 healthy donors and 26 diabetic donors separately. SCDC selects weights via the method of least absolute deviation (LAD) regression, grid search to minimize the mAD of Y and $\hat{Y}$, and grid search to maximize the Spearman correlation of Y and $\hat{Y}$.

	Metrics	**Weight for Baron et al.**	**Weight for Segerstolpe et al.**	**RMSD (Y)**	**mAD (Y)**	**Spearman R (Y)**
51 Healthy samples	LAD	0.17	0.83	12.630	5.834	0.649
	mAD $(Y, \hat{Y})$	0.24	0.76	12.684	5.829	0.652
	Spearman R $(Y, \hat{Y})$	0.40	0.60	13.020	5.851	0.655
26 Diabetic samples	LAD	0.33	0.67	11.509	5.494	0.687
	mAD $(Y, \hat{Y})$	0.38	0.62	11.562	5.492	0.688
	Spearman R $(Y, \hat{Y})$	0.48	0.52	11.755	5.501	0.689

Table S5. Associating cell-type proportions with HbA1c levels in human pancreatic islet samples. A linear regression model (deconvolved cell-type proportion ~ HbA1c + age + BMI + sex) is adopted for each cell type separately. SCDC through ENSEMBLE (weights are derived via grid search to maximize Spearman correlation of Y and $\hat{Y}$, shown in Table S4) derived a p-value of 0.0018 for the association between the HbA1c levels and the beta cell proportions, more significant than those from deconvolution without ENSEMBLE.

Cell type (% As Outcome)		Estimate	Std. Error	p-value using ENSEMBLE	p-value using Baron et al.	p-value using Segerstolpe et al.
alpha	(Intercept)	0.7782	0.2124	0.0005	0.0001	0.0398
	HbA1c	-0.0066	0.0284	0.8159	0.4346	0.6449
	age	-0.0010	0.0020	0.6119	0.7861	0.9251
	BMI	-0.0104	0.0081	0.2040	0.2033	0.7888
	sexFemale	0.0463	0.0438	0.2940	0.4485	0.1597
beta	(Intercept)	0.4430	0.1127	0.0002	0.0316	0.0000
	HbA1c	**-0.0488**	**0.0150**	**0.0018**	**0.0380**	**0.0310**
	age	0.0021	0.0010	0.0483	0.2291	0.2332
	BMI	-0.0045	0.0043	0.3019	0.6331	0.0590
	sexFemale	-0.0679	0.0233	0.0047	0.1952	0.0040
delta	(Intercept)	0.0537	0.0098	0.0000	0.0000	0.0000
	HbA1c	-0.0011	0.0013	0.4214	0.4243	0.4527
	age	-0.0001	0.0001	0.1134	0.1344	0.1256
	BMI	-0.0011	0.0004	0.0051	0.0058	0.0059
	sexFemale	0.0000	0.0020	0.9948	0.8213	0.8473
gamma	(Intercept)	0.0108	0.0122	0.3763	0.3556	0.3827
	HbA1c	0.0012	0.0016	0.4749	0.5547	0.4622
	age	0.0001	0.0001	0.3369	0.2919	0.3508
	BMI	-0.0007	0.0005	0.1370	0.1281	0.1467
	sexFemale	-0.0018	0.0025	0.4680	0.4248	0.4588
acinar	(Intercept)	-0.0067	0.0804	0.9335	0.8045	0.4762
	HbA1c	0.0205	0.0107	0.0598	0.0283	0.0517
	age	-0.0015	0.0007	0.0466	0.1660	0.0476
	BMI	0.0017	0.0031	0.5842	0.7244	0.2173
	sexFemale	0.0308	0.0166	0.0673	0.0516	0.0574
ductal	(Intercept)	-0.2790	0.1735	0.1122	0.2490	0.1091
	HbA1c	0.0348	0.0232	0.1380	0.2929	0.1213
	age	0.0005	0.0016	0.7802	0.8249	0.5345
	BMI	0.0150	0.0066	0.0270	0.0567	0.0513
	sexFemale	-0.0074	0.0358	0.8361	0.4099	0.9686

Table S6. ENSEMBLE weight selection results for the mouse mammary gland bulk samples. Single-cell reference dataset from the same source as the bulk samples gains more weight.

	Metrics	**Weight for Tabula Muris**	**Weight for Perou**	**RMSD (Y)**	**mAD (Y)**	**Spearman R (Y)**
Pooled single cells after suspension	LAD	0.18	0.82	18.116	7.658	0.634
	mAD $(Y, \hat{Y})$	0.16	0.84	18.064	7.657	0.633
	Spearman R $(Y, \hat{Y})$	0.32	0.68	18.794	7.690	0.636
Fresh-frozen	LAD	0.15	0.85	25.102	7.959	0.532
	mAD $(Y, \hat{Y})$	0.16	0.84	25.108	7.958	0.532
	Spearman R $(Y, \hat{Y})$	0.44	0.56	26.148	8.030	0.539