
Revealing a coherent cell-state landscape across single-cell datasets with CONCORD

In the format provided by the
authors and unedited

Supplementary Tables

cluster (full overlap, 1:1)

	Graph connectivity	Batch correction				PCR comparison	Silhouette batch	cLSI	Isolated labels	Bio conservation				Linear state accuracy	Geometry		Topology			Aggregate score		
		ILISI	KBET							Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy		Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology
Method																						
scvi	0.985	0.641	0.991	0.993	0.955	1.000	0.769	1.000	1.000	0.770	1.000	1.000	0.827	0.922			0.913	0.934	0.875		0.907	
concord_hd	0.996	0.873	0.718	0.966	0.911	1.000	0.717	1.000	1.000	0.716	1.000	1.000	0.841	0.921			0.893	0.919	0.881		0.898	
concord_knn	0.994	0.316	0.315	0.998	0.755	1.000	0.945	0.930	0.957	0.944	1.000	0.995	0.983	0.923			0.675	0.967	0.953		0.865	
liger	0.915	0.766	0.896	0.979	0.939	1.000	0.671	0.919	0.938	0.670	1.000	0.490	0.810	0.921			0.897	0.890	0.866		0.854	
harmony	0.479	0.000	0.001	0.230	0.528	1.000	0.715	0.614	0.823	0.715	1.000	1.000	0.795	0.922			0.248	0.838	0.858		0.648	
contrastive	0.494	0.000	0.010	0.610	0.633	1.000	0.713	0.613	0.821	0.713	1.000	1.000	0.670	0.921			0.230	0.837	0.795		0.621	
unintegrated	0.480	0.000	0.001	0.000	0.436	1.000	0.694	0.614	0.823	0.694	1.000	1.000	0.637	0.922			0.183	0.832	0.779		0.598	
scanorama	0.481	0.000	0.000	0.000	0.405	1.000	0.702	0.614	0.823	0.702	1.000	0.993	0.637	0.922			0.177	0.833	0.780		0.597	

cluster (full overlap, 5:1)

	Graph connectivity	Batch correction			PCR comparison	Silhouette batch	cLSI	Isolated labels	Bio conservation		Silhouette label	KNN state accuracy	Linear state accuracy	Geometry		Topology		Aggregate score			
		ILISI	KBET						Leiden ARI	Leiden NMI				Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology
Method																					
harmony	0.842	0.166	0.806	0.983	0.935	1.000	0.792	1.000	1.000	0.792	1.000	1.000	0.983	0.923				0.747	0.941	0.953	0.880
concord_knn	0.993	0.072	0.503	0.998	0.714	1.000	0.963	1.000	1.000	0.963	1.000	1.000	0.995	0.924				0.656	0.989	0.960	0.868
concord_hd	0.997	0.050	0.599	0.955	0.776	1.000	0.772	1.000	1.000	0.772	1.000	1.000	0.884	0.924				0.675	0.935	0.904	0.838
scvi	0.923	0.000	0.000	0.966	0.756	1.000	0.860	0.807	0.878	0.860	1.000	0.934	0.924					0.529	0.915	0.928	0.791
liger	0.963	0.230	0.476	0.988	0.946	1.000	0.635	0.990	0.990	0.635	1.000	0.439	0.723	0.921				0.721	0.813	0.822	0.785
contrastive	0.825	0.000	0.000	0.618	0.630	1.000	0.816	0.805	0.877	0.816	1.000	1.000	0.844	0.912				0.295	0.902	0.878	0.692
scanorama	0.751	0.000	0.000	0.000	0.385	1.000	0.751	0.806	0.877	0.751	1.000	0.982	0.906	0.920				0.227	0.881	0.913	0.674
unintegrated	0.717	0.000	0.000	0.000	0.412	1.000	0.741	0.806	0.877	0.741	1.000	1.000	0.904	0.920				0.226	0.881	0.912	0.673

cluster (full overlap, 20:1)

	Graph connectivity	Batch correction			PCR comparison	Silhouette batch	cLSI	Isolated labels	Bio conservation			Linear state accuracy	Geometry		Topology		Aggregate score			
		ILISI	KBET						Leiden ARI	Leiden NMI	Silhouette label		KNN state accuracy	Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry
Method																				
concord_knn	0.983	0.001	0.939	0.984	0.604	1.000	0.969	0.965	0.973	0.969	1.000	1.000	0.993	0.923			0.703	0.982	0.958	0.881
harmony	0.915	0.000	0.970	0.905	0.845	1.000	0.792	1.000	1.000	0.792	1.000	0.980	0.980	0.921			0.731	0.940	0.951	0.874
concord_hd	0.988	0.000	0.925	0.938	0.692	1.000	0.810	1.000	1.000	0.811	1.000	1.000	0.920	0.923			0.709	0.946	0.922	0.859
scvi	0.990	0.012	0.891	0.970	0.811	1.000	0.720	1.000	1.000	0.720	1.000	0.987	0.817	0.921			0.735	0.918	0.869	0.841
unintegrated	0.870	0.000	0.953	0.000	0.409	1.000	0.776	0.941	0.944	0.776	1.000	1.000	0.977	0.919			0.446	0.920	0.948	0.771
scanorama	0.890	0.000	0.953	0.000	0.380	1.000	0.789	0.941	0.944	0.789	1.000	0.930	0.979	0.921			0.445	0.911	0.950	0.768
liger	0.940	0.000	0.901	0.958	0.699	1.000	0.626	0.998	0.998	0.626	0.993	0.983	0.913	0.923			0.623	0.982	0.978	0.738
contrastive	0.846	0.000	0.000	0.244	0.055	1.000	0.913	0.880	0.910	0.913	1.000	0.972	0.949	0.909			0.249	0.941	0.929	0.707

cluster (partial overlap, batch number=2)

	Graph connectivity	Batch correction				PCR comparison	Silhouette batch	cLSI	Isolated labels	Bio conservation				Linear state accuracy	Geometry		Topology			Aggregate score		
		ILISI	KBET							Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy		Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology
Method																						
concord_knn	0.991	0.015	0.232	0.845	0.746	1.000	0.886	1.000	1.000	0.908	1.000	0.935	0.913	0.941				0.566	0.961	0.927		0.818
liger	0.984	0.367	0.575	0.770	0.992	1.000	0.736	1.000	1.000	0.682	1.000	0.236	0.856	0.941				0.738	0.808	0.898		0.815
concord_hd	0.830	0.046	0.001	0.477	0.832	1.000	0.787	1.000	1.000	0.755	1.000	0.943	0.939					0.487	0.935	0.941		0.778
scvi	0.981	0.032	0.182	0.731	0.822	1.000	0.745	1.000	0.999	0.740	0.996	0.993	0.703	0.940				0.552	0.925	0.821		0.766
unintegrated	0.724	0.268	0.124	0.974	0.874	0.994	0.596	0.979	0.974	0.511	0.983	0.970	0.566	0.902				0.593	0.858	0.734		0.729
contrastive	0.806	0.000	0.000	0.000	0.696	1.000	0.788	0.793	0.921	0.724	1.000	1.000	0.846	0.942				0.301	0.889	0.894		0.695
scanorama	0.856	0.000	0.000	0.174	0.074	1.000	0.966	0.793	0.921	0.853	1.000	1.000	0.829	0.942				0.221	0.933	0.886		0.680
liger	0.786	0.000	0.000	0.000	0.657	1.000	0.794	0.793	0.921	0.738	1.000	0.531	0.852	0.941				0.289	0.825	0.897		0.670

cluster (partial overlap, batch number=10)

	Graph connectivity	Batch correction				PCR comparison	Silhouette batch	cLSI	Isolated labels	Bio conservation		Silhouette label	KNN state accuracy	Linear state accuracy	Geometry		Topology		Aggregate score				
		ILISI	KBET							Leiden ARI	Leiden NMI				Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology	Average	
Method																							
harmony	0.759	0.165	0.972	0.552	0.959	1.000	0.789	1.000	1.000	0.772	1.000	1.000	0.967	0.939				0.681	0.937	0.953		0.857	
concord_knn	0.992	0.068	0.246	0.803	0.812	1.000	0.851	0.997	0.997	0.823	1.000	0.996	0.827	0.937				0.584	0.952	0.882		0.806	
scvi	0.680	0.189	0.802	0.848	0.881	1.000	0.612	0.993	0.990	0.557	0.999	0.990	0.653	0.922				0.700	0.877	0.788		0.788	
concord_hd	0.992	0.084	0.273	0.698	0.849	1.000	0.756	0.994	0.993	0.719	0.999	0.996	0.633	0.933				0.579	0.922	0.783		0.762	
liger	0.942	0.127	0.475	0.575	0.884	1.000	0.689	0.845	0.915	0.662	0.988	0.983	0.223	0.771	0.927				0.603	0.764	0.849		0.739
unintegrated	0.526	0.000	0.069	0.000	0.714	1.000	0.723	0.637	0.836	0.705	1.000	1.000	0.907	0.938				0.262	0.843	0.922		0.676	
scanorama	0.816	0.000	0.069	0.000	0.759	1.000	0.725	0.643	0.855	0.713	1.000	0.378	0.915	0.938				0.329	0.759	0.926		0.671	
contrastive	0.446	0.000	0.001	0.000	0.136	1.000	0.689	0.397	0.722	0.631	0.994	0.971	0.565	0.908				0.116	0.772	0.736		0.542	

cluster (partial overlap, batch number=30)

	Graph connectivity	Batch correction				PCR comparison	Silhouette batch	cLSI	Isolated labels	Bio conservation				Linear state accuracy	Geometry		Topology			Aggregate score		
		ILISI	KBET							Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy		Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology
Method																						
harmony	0.795	0.058	0.937	0.799	0.856	1.000	0.820	1.000	1.000	0.768	1.000	1.000	0.971	0.909				0.589	0.941	0.940		0.823
concord_knn	0.965	0.051	0.771	0.777	0.832	1.000	0.883	0.967	0.853	0.982	0.866	0.932	0.907	0.878				0.578	0.913	0.904		0.804
sct	0.721	0.080	0.902	0.922	0.837	0.993	0.687	0.882	0.864	0.551	0.939	0.973	0.581	0.843				0.692	0.879	0.712		0.745
concord_hd	0.973	0.031	0.656	0.199	0.818	1.000	0.782	0.977	0.954	0.664	0.991	0.973	0.596	0.899				0.535	0.906	0.747		0.730
unintegrated	0.000	0.000	0.325	0.000	0.798	1.000	0.768	0.705	0.877	0.712	1.000	0.930	0.910				0.381	0.866	0.920		0.722	
scanorama	0.917	0.001	0.324	0.000	0.833	1.000	0.730	0.726	0.890	0.705	0.917	0.945	0.917	0.915				0.415	0.907	0.786		0.760
contrastive	0.455	0.000	0.331	0.000	0.338	1.000	0.739	0.444	0.692	0.611	0.984	0.981	0.592	0.897				0.223	0.779	0.744		0.588

loop (full overlap)

	Batch correction					Bio conservation					Geometry		Topology		Aggregate score						
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLSI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology	Average
Method																					
concord_knn	0.934	0.937	0.903	0.994	0.819	0.969	0.588	0.541	0.782	0.593	0.780	0.340	0.851	0.988	0.337	1.000	0.917	0.656	0.920	0.867	0.840
concord_hsl	0.971	0.953	0.965	0.989	0.889	0.967	0.581	0.515	0.775	0.581	0.755	0.665	0.690	0.987	0.233	1.000	0.953	0.691	0.838	0.847	0.832
scvi	0.490	0.016	0.119	0.906	0.474	0.954	0.537	0.347	0.676	0.536	0.705	0.355	0.925	0.978	0.236	1.000	0.401	0.587	0.952	0.847	0.697
liger	0.453	0.783	0.991	0.972	0.943	0.887	0.481	0.188	0.494	0.481	0.450	0.105	0.565	0.895	0.010	0.500	0.829	0.441	0.730	0.402	0.600
contrastive	0.495	0.000	0.000	0.600	0.155	0.956	0.509	0.346	0.698	0.508	0.745	0.335	0.667	0.973	0.390	0.333	0.250	0.585	0.820	0.345	0.500
harmony	0.471	0.000	0.000	0.039	0.510	0.917	0.501	0.290	0.627	0.501	0.540	0.315	0.674	0.957	0.004	0.500	0.204	0.527	0.816	0.401	0.487
unintegrated	0.472	0.000	0.000	0.000	0.470	0.917	0.501	0.286	0.620	0.501	0.500	0.385	0.599	0.957	0.004	0.500	0.188	0.523	0.778	0.401	0.472
scanorama	0.462	0.000	0.000	0.000	0.441	0.923	0.497	0.296	0.630	0.498	0.440	0.300	0.600	0.961	0.010	0.500	0.181	0.483	0.781	0.402	0.462

loop (partial overlap)

	Batch correction					Bio conservation					Geometry		Topology		Aggregate score						
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLSI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology	Average
Method																					
concord_knn	0.943	0.113	0.353	0.841	0.768	0.968	0.637	0.477	0.744	0.578	0.736	0.659	0.715	0.981	0.212	1.000	0.604	0.685	0.848	0.842	0.745
concord_hsl	0.686	0.000	0.029	0.601	0.141	0.972	0.707	0.443	0.744	0.564	0.779	0.437	0.685	0.982	0.370	1.000	0.291	0.663	0.833	0.874	0.665
concord_knn	0.909	0.160	0.385	0.966	0.682	0.966	0.622	0.521	0.759	0.580	0.764	0.455	0.755	0.978	0.318	0.500	0.620	0.667	0.866	0.464	0.654
liger	0.665	0.209	0.879	0.736	0.904	0.904	0.506	0.240	0.538	0.477	0.511	0.050	0.625	0.908	0.009	0.500	0.679	0.463	0.766	0.402	0.577
scanorama	0.670	0.000	0.036	0.000	0.428	0.944	0.563	0.346	0.669	0.488	0.629	0.227	0.697	0.964	0.016	0.500	0.227	0.552	0.830	0.403	0.503
scvi	0.470	0.703	1.000	0.985	0.978	0.867	0.509	0.149	0.458	0.480	0.486	0.337	0.526	0.822	0.003	0.001	0.827	0.469	0.674	0.001	0.493
harmony	0.655	0.000	0.047	0.000	0.478	0.941	0.574	0.340	0.667	0.487	0.571	0.495	0.702	0.959	0.005	0.333	0.236	0.582	0.830	0.268	0.479
unintegrated	0.655	0.000	0.039	0.000	0.473	0.941	0.574	0.336	0.664	0.485	0.629	0.437	0.695	0.959	0.005	0.333	0.233	0.581	0.827	0.268	0.477

loop (connected)

	Batch correction					Bio conservation					Geometry		Topology		Aggregate score						
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLSI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology	Average
Method																					
concord_hsl	0.945	0.000	0.406	0.763	0.423	0.970	0.610	0.538	0.783	0.611	0.805	0.635	0.722	0.984	0.270	1.000	0.597	0.707	0.853	0.854	0.730
concord_knn	0.988	0.000	0.387	0.916	0.326	0.973	0.612	0.573	0.791	0.604	0.863	0.435	0.672	0.977	0.245	1.000	0.504	0.690	0.824	0.869	0.722
scvi	0.908	0.000	0.000	0.950	0.151	0.972	0.648	0.590	0.793	0.610	0.800	0.555	0.733	0.991	0.278	0.500	0.402	0.710	0.862	0.456	0.607
contrastive	0.932	0.000	0.000	0.657	0.055	0.974	0.673	0.556	0.785	0.623	0.845	0.435	0.741	0.992	0.384	0.333	0.329	0.699	0.866	0.343	0.559
harmony	0.907	0.000	0.000	0.000	0.429	0.948	0.541	0.522	0.768	0.523	0.750	0.345	0.898	0.981	0.003	0.500	0.267	0.628	0.940	0.401	0.559
unintegrated	0.960	0.000	0.000	0.000	0.368	0.987	0.540	0.519	0.766	0.518	0.740	0.395	0.860	0.981	0.003	0.333	0.215	0.635	0.925	0.267	0.320
scanorama	0.923	0.000	0.000	0.000	0.357	0.948	0.541	0.515	0.764	0.517	0.745	0.185	0.871	0.983	0.004	0.333	0.256	0.602	0.927	0.268	0.513
liger	0.766	0.004	0.656	0.783	0.845	0.892	0.493	0.306	0.544	0.491	0.490	0.085	0.645	0.899	0.003	0.001	0.611	0.472	0.772	0.001	0.644

tree (full overlap)

	Batch correction					Bio conservation					Geometry		Topology		Aggregate score						
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLSI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology	Average
Method																					
concord_knn	0.907	0.937	0.867	0.998	0.954	1.000	0.637	0.397	0.681	0.637	0.923	0.826	0.695	0.991	0.434	1.000	0.933	0.729	0.843	0.887	0.848
concord_hsl	0.806	0.831	0.740	0.909	0.881	0.995	0.610	0.447	0.675	0.609	0.891	0.871	0.947	0.984	0.025	1.000	0.814	0.728	0.965	0.805	0.833
concord_hsl	0.939	0.830	0.973	0.992	0.953	1.000	0.621	0.432	0.694	0.621	0.877	0.866	0.431	0.992	0.308	1.000	0.938	0.730	0.712	0.862	0.810
liger	0.879	0.835	0.682	0.968	0.910	0.998	0.588	0.505	0.699	0.589	0.917	0.314	0.509	0.969	0.211	1.000	0.855	0.659	0.739	0.842	0.734
scvi	0.956	0.831	1.000	0.998	0.984	0.971	0.589	0.556	0.732	0.590	0.906	0.829	0.701	0.964	0.022	0.500	0.954	0.739	0.833	0.404	0.732
contrastive	0.479	0.000	0.014	0.764	0.494	0.998	0.581	0.338	0.626	0.581	0.900	0.834	0.540	0.983	0.451	1.000	0.350	0.694	0.762	0.890	0.674
scanorama	0.484	0.000	0.000	0.000	0.345	0.983	0.546	0.333	0.632	0.546	0.869	0.743	0.589	0.978	0.080	1.000	0.178	0.666	0.784	0.816	0.611
unintegrated	0.481	0.000	0.060	0.000	0.371	0.981	0.565	0.283	0.635	0.564	0.900	0.880	0.645	0.977	0.040	0.500	0.182	0.686	0.810	0.410	0.522

tree (partial overlap)

	Batch correction					Bio conservation					Geometry		Topology		Aggregate score						
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLSI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology	Average
Method																					
liger	0.897	0.658	0.482	0.952	0.871	0.998	0.482	0.577	0.720	0.586	0.894	0.367	0.730	0.964	0.039	1.000	0.772	0.661	0.847	0.808	0.722
concord_knn	0.906	0.077	0.420	0.978	0.794	1.000	0.631	0.342	0.636	0.613	0.878	0.865	0.627	0.988	0.377	1.000	0.635	0.709	0.807	0.875	0.757
concord_hsl	0.968	0.165	0.614	0.935	0.797	1.000	0.612	0.375	0.662	0.608	0.902	0.872	0.400	0.989	0.322	1.000	0.696	0.719	0.695	0.864	0.743
scvi	0.946	0.778	0.661	0.994	0.828	0.974	0.501	0.516	0.722	0.565	0.900	0.889	0.611	0.952	0.005	0.333	0.861	0.717	0.782	0.268	0.657
contrastive	0.537	0.000	0.034	0.681	0.488	0.977	0.657	0.248	0.576	0.556	0.802	0.731	0.514	0.957	0.432	0.500	0.348	0.650	0.735	0.486	0.555
harmony	0.630	0.000	0.039	0.002	0.336	0.997	0.783	0.330	0.634	0.565	0.870	0.857	0.711	0.974	0.034	0.500	0.201	0.720	0.843	0.407	0.543
unintegrated	0.630	0.000	0.039	0.000	0.330	0.997	0.783	0.300	0.624	0.563	0.854	0.870	0.697	0.974	0.033	0.500	0.200	0.713	0.836	0.407	0.539
scanorama	0.630	0.000	0.038	0.000	0.308	0.998	0.723	0.322	0.621	0.556	0.876	0.739	0.681	0.976	0.040	0.500	0.195	0.691	0.828	0.408	0.531

tree (connected)

	Batch correction					Bio conservation					Geometry		Topology		Aggregate score						
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLSI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Cell distance correlation	Trustworthiness	Betti curve stability	Betti number accuracy	Batch correction	Bio conservation	Geometry	Topology	Average
Method																					
concord_knn	0.950	0.000	0.948	0.973	0.841	1.000	0.552	0.423	0.707	0.571	0.926	0.743	0.629	0.987	0.432	1.000	0.742	0.703	0.808	0.886	0.785
scvi	0.969	0.058	0.924	0.982	0.859	0.971	0.587	0.539	0.717	0.559	0.910	0.700	0.628	0.933	0.011	1.000	0.758	0.712	0.781	0.862	0.763
concord_his	0.950	0.000	0.974	0.748	0.588	1.000	0.612	0.531	0.741	0.614	0.926	0.863	0.537	0.993	0.329	0.500	0.552	0.736	0.765	0.466	0.635
contrast	0.000	0.000	0.789	0.463	0.342	0.643	0.343	0.286	0.710	0.346	0.906	0.662	0.396	0.993	0.043	0.489	0.292	0.449	0.273	0.358	0.358
figer	0.851	0.000	0.000	0.829	0.633	0.976	0.538	0.524	0.700	0.582	0.869	0.657	0.516	0.943	0.056	0.500	0.468	0.621	0.829	0.411	0.431
scanorama	0.946	0.046	0.000	0.000	0.366	1.000	0.654	0.590	0.735	0.636	0.934	0.889	0.793	0.986	0.020	0.333	0.254	0.774	0.890	0.271	0.547
sc3	0.000	0.000	0.000	0.000	0.126	0.000	0.554	0.519	0.686	0.584	0.883	0.794	0.686	0.920	0.024	0.333	0.254	0.774	0.890	0.271	0.547
unscrambled	0.949	0.049	0.000	0.000	0.333	1.000	0.648	0.564	0.753	0.634	0.920	0.854	0.764	0.987	0.041	0.333	0.256	0.768	0.876	0.275	0.544

Lung atlas

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Batch correction	Bio conservation	Total
concord_hcl	0.915	0.116	0.369	0.875	0.836	0.997	0.710	0.618	0.722	0.551	0.868	0.868	0.622	0.762	0.692
scvi	0.898	0.118	0.341	0.927	0.905	0.995	0.640	0.572	0.742	0.539	0.878	0.878	0.813	0.638	0.689
concord_knn	0.892	0.109	0.411	0.938	0.752	0.998	0.859	0.450	0.679	0.565	0.886	0.859	0.621	0.757	0.689
seurat_rpca	0.745	0.133	0.457	0.848	0.817	0.993	0.734	0.528	0.699	0.578	0.829	0.821	0.600	0.740	0.670
harmony	0.799	0.093	0.393	0.900	0.997	0.997	0.577	0.528	0.717	0.555	0.867	0.812	0.549	0.722	0.636
seurat_cca	0.545	0.172	0.431	0.902	0.849	0.975	0.647	0.456	0.616	0.530	0.778	0.753	0.580	0.679	0.630
contrastive	0.782	0.015	0.262	0.557	0.716	1.000	0.777	0.401	0.721	0.612	0.937	0.927	0.467	0.768	0.617
liger	0.753	0.129	0.374	0.839	0.861	0.995	0.641	0.477	0.671	0.560	0.824	0.255	0.591	0.632	0.611
scanorama	0.760	0.049	0.275	0.309	0.911	1.000	0.647	0.491	0.725	0.549	0.934	0.878	0.461	0.746	0.603
unintegrated	0.735	0.001	0.196	0.189	0.863	1.000	0.687	0.386	0.708	0.547	0.952	0.944	0.397	0.746	0.572

GTEx v9

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	KNN state accuracy	Leiden ARI	Leiden NMI	Linear state accuracy	Silhouette label	Batch correction	Bio conservation	Total
scvi	0.878	0.024	0.388	0.448	0.914	1.000	0.633	0.931	0.717	0.836	0.919	0.528	0.530	0.795	0.663
concord_knn	0.819	0.031	0.466	0.708	0.746	1.000	0.562	0.901	0.628	0.769	0.905	0.496	0.554	0.751	0.653
concord_hcl	0.834	0.032	0.488	0.553	0.828	0.999	0.592	0.888	0.622	0.754	0.894	0.503	0.547	0.750	0.649
scanorama	0.792	0.019	0.399	0.000	0.906	1.000	0.621	0.941	0.574	0.804	0.935	0.514	0.423	0.770	0.597
unintegrated	0.723	0.015	0.352	0.000	0.946	1.000	0.571	0.940	0.606	0.821	0.943	0.517	0.407	0.771	0.589
harmony	0.622	0.052	0.502	0.687	0.884	0.991	0.545	0.729	0.412	0.594	0.646	0.481	0.549	0.628	0.589
contrastive	0.760	0.020	0.474	0.000	0.692	1.000	0.579	0.934	0.411	0.773	0.932	0.552	0.389	0.740	0.565

HypoMap

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	KNN state accuracy	Leiden ARI	Leiden NMI	Linear state accuracy	Silhouette label	Batch correction	Bio conservation	Total
concord_hcl	0.842	0.055	0.192	0.825	0.800	1.000	0.760	0.928	0.614	0.787	0.920	0.526	0.542	0.791	0.667
scvi	0.888	0.026	0.135	0.719	0.909	1.000	0.706	0.942	0.641	0.786	0.919	0.554	0.535	0.792	0.664
concord_knn	0.821	0.046	0.188	0.871	0.749	1.000	0.482	0.918	0.524	0.764	0.908	0.585	0.535	0.740	0.638
harmony	0.671	0.055	0.158	0.799	0.897	0.998	0.540	0.870	0.437	0.674	0.790	0.552	0.516	0.695	0.605
liger	0.467	0.156	0.386	0.941	0.743	0.979	0.265	0.704	0.453	0.548	0.590	0.510	0.539	0.578	0.559
scanorama	0.483	0.012	0.066	0.000	0.862	1.000	0.595	0.897	0.230	0.575	0.873	0.525	0.284	0.671	0.478
unintegrated	0.315	0.002	0.111	0.039	0.913	1.000	0.592	0.910	0.222	0.571	0.923	0.509	0.276	0.675	0.476
contrastive	0.429	0.003	0.035	0.061	0.633	1.000	0.592	0.819	0.193	0.535	0.811	0.509	0.232	0.637	0.435

Immune cell atlas

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	KNN state accuracy	Leiden ARI	Leiden NMI	Linear state accuracy	Silhouette label	Batch correction	Bio conservation	Total
concord_hcl	0.939	0.088	0.212	0.000	0.853	1.000	0.665	0.921	0.682	0.788	0.918	0.564	0.418	0.791	0.605
concord_knn	0.914	0.049	0.204	0.150	0.802	1.000	0.724	0.919	0.557	0.756	0.916	0.620	0.424	0.785	0.604
scvi	0.909	0.033	0.144	0.000	0.916	0.999	0.578	0.909	0.646	0.757	0.914	0.541	0.400	0.763	0.582
harmony	0.804	0.091	0.175	0.000	0.920	0.997	0.564	0.868	0.607	0.715	0.859	0.534	0.398	0.735	0.566
scanorama	0.863	0.037	0.138	0.000	0.905	0.998	0.548	0.899	0.527	0.710	0.890	0.551	0.389	0.732	0.560
liger	0.713	0.218	0.487	0.000	0.854	0.993	0.396	0.791	0.487	0.626	0.466	0.544	0.455	0.615	0.535
contrastive	0.747	0.000	0.137	0.000	0.653	1.000	0.626	0.904	0.253	0.646	0.902	0.557	0.307	0.698	0.503
unintegrated	0.468	0.001	0.101	0.000	0.927	0.995	0.548	0.858	0.307	0.589	0.910	0.525	0.299	0.676	0.488

Mouse pancreatic islet

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	KNN state accuracy	Leiden ARI	Leiden NMI	Linear state accuracy	Silhouette label	Batch correction	Bio conservation	Total
concord_knn	0.892	0.047	0.301	0.757	0.698	1.000	0.874	0.969	0.550	0.732	0.965	0.585	0.539	0.811	0.675
concord_hcl	0.913	0.047	0.294	0.426	0.774	1.000	0.869	0.969	0.781	0.822	0.965	0.525	0.491	0.847	0.669
scvi	0.942	0.031	0.203	0.416	0.888	1.000	0.688	0.974	0.836	0.833	0.971	0.566	0.496	0.838	0.667
harmony	0.789	0.059	0.316	0.333	0.843	1.000	0.748	0.953	0.775	0.763	0.844	0.609	0.468	0.813	0.641
liger	0.648	0.102	0.570	0.806	0.673	0.997	0.661	0.894	0.540	0.646	0.577	0.526	0.560	0.692	0.626
scanorama	0.798	0.028	0.131	0.000	0.904	1.000	0.766	0.974	0.313	0.652	0.972	0.608	0.372	0.756	0.564
contrastive	0.810	0.019	0.124	0.000	0.751	1.000	0.865	0.969	0.189	0.607	0.965	0.521	0.341	0.731	0.536
unintegrated	0.484	0.018	0.089	0.000	0.932	1.000	0.713	0.969	0.342	0.678	0.973	0.584	0.304	0.751	0.528

Tabula Sapiens

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	KNN state accuracy	Leiden ARI	Leiden NMI	Linear state accuracy	Silhouette label	Batch correction	Bio conservation	Total
scvi	0.748	0.008	0.253	0.186	0.817	1.000	0.572	0.896	0.592	0.787	0.862	0.530	0.402	0.749	0.575
harmony	0.494	0.008	0.191	0.786	0.879	1.000	0.566	0.862	0.593	0.758	0.823	0.497	0.472	0.728	0.600
unintegrated	0.587	0.001	0.232	0.181	0.887	1.000	0.571	0.906	0.578	0.778	0.883	0.518	0.378	0.748	0.563
concord_knn	0.689	0.018	0.255	0.919	0.740	1.000	0.573	0.872	0.574	0.777	0.869	0.482	0.524	0.735	0.630
concord_hcl	0.716	0.014	0.249	0.851	0.798	1.000	0.551	0.883	0.583	0.771	0.881	0.476	0.526	0.735	0.630
contrastive	0.696	0.010	0.271	0.000	0.690	1.000	0.594	0.886	0.458	0.759	0.875	0.513	0.333	0.726	0.530

C.elegans cell type

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	ILISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Batch correction	Bio conservation	Total
concord_hcl	0.856	0.268	0.731	0.969	0.860	1.000	0.636	0.500	0.820	0.542	0.890	0.900	0.737	0.755	0.746
concord_knn	0.823	0.265	0.699	0.963	0.823	1.000	0.636	0.506	0.794	0.541	0.863	0.886	0.715	0.747	0.731
seurat_cca	0.725	0.279	0.731	0.729	0.871	1.000	0.567	0.565	0.832	0.528	0.893	0.897	0.667	0.754	0.711
seurat_rpca	0.764	0.319	0.674	0.675	0.870	1.000	0.575	0.534	0.835	0.528	0.890	0.902	0.660	0.752	0.706
harmony	0.589	0.152	0.633	0.815	0.938	0.998	0.567	0.463	0.778	0.513	0.825	0.889	0.625	0.719	0.672
scanorama	0.736	0.195	0.612	0.497	0.916	0.998	0.566	0.460	0.779	0.501	0.850	0.861	0.591	0.717	0.654
scvi	0.866	0.190	0.522	0.392	0.867	0.998	0.580	0.495	0.793	0.528	0.839	0.858	0.568	0.727	0.647
liger	0.654	0.238	0.698	0.951	0.817	0.998	0.554	0.458	0.750	0.484	0.779	0.107	0.672	0.590	0.631
unintegrated	0.637	0.141	0.438	0.287	0.917	0.998	0.567	0.441	0.764	0.508	0.823	0.879	0.484	0.711	0.598
contrastive	0.590	0.134	0.312	0.597	0.770	0.996	0.625	0.339	0.700	0.520	0.731	0.810	0.480	0.674	0.577

C.elegans lineage

	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	iLISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Batch correction	Bio conservation	Total
Method															
concord_hcl	0.823	0.229	0.857	0.950	0.877	0.999	0.561	0.412	0.802	0.562	0.817	0.843	0.747	0.714	0.731
concord_knn	0.798	0.238	0.829	0.939	0.829	0.999	0.594	0.370	0.781	0.544	0.768	0.810	0.726	0.695	0.711
scanorama	0.642	0.189	0.857	0.615	0.945	0.998	0.479	0.282	0.767	0.511	0.811	0.628	0.650	0.668	0.659

harmony	0.501	0.167	0.810	0.873	0.948	0.997	0.485	0.236	0.733	0.505	0.756	0.847	0.660	0.651	0.656
scvi	0.805	0.193	0.777	0.522	0.899	0.998	0.525	0.250	0.731	0.523	0.750	0.791	0.639	0.653	0.646
seurat_cca	0.496	0.237	0.745	0.725	0.894	0.998	0.490	0.287	0.753	0.519	0.757	0.823	0.619	0.661	0.640
seurat_rpca	0.619	0.195	0.637	0.639	0.921	0.998	0.524	0.300	0.757	0.519	0.770	0.824	0.602	0.670	0.636
unintegrated	0.536	0.154	0.692	0.354	0.944	0.997	0.482	0.242	0.736	0.509	0.770	0.852	0.536	0.655	0.596
liger	0.442	0.224	0.765	0.874	0.878	0.995	0.451	0.151	0.643	0.467	0.557	0.026	0.637	0.470	0.553
contrastive	0.540	0.141	0.561	0.259	0.799	0.996	0.430	0.242	0.672	0.487	0.561	0.711	0.460	0.586	0.523

PBMC scATAC

Method	Batch correction					Bio conservation							Aggregate score		
	Graph connectivity	iLISI	KBET	PCR comparison	Silhouette batch	cLISI	Isolated labels	Leiden ARI	Leiden NMI	Silhouette label	KNN state accuracy	Linear state accuracy	Batch correction	Bio conservation	Total
concord_hci	0.924	0.130	0.192	0.948	0.881	0.994	0.553	0.606	0.648	0.516	0.877	0.868	0.615	0.723	0.669
concord_knn	0.915	0.152	0.355	0.986	0.774	0.993	0.640	0.353	0.563	0.565	0.871	0.862	0.636	0.692	0.664
liger	0.775	0.167	0.550	0.984	0.821	0.976	0.553	0.378	0.500	0.557	0.841	0.700	0.659	0.643	0.651
scvi	0.934	0.108	0.135	0.801	0.936	0.990	0.555	0.486	0.610	0.536	0.888	0.874	0.583	0.705	0.644
contrastive	0.923	0.074	0.092	0.807	0.876	0.997	0.554	0.304	0.538	0.513	0.891	0.881	0.554	0.668	0.611
scanorama	0.730	0.089	0.140	0.218	0.868	0.971	0.500	0.308	0.459	0.488	0.858	0.837	0.409	0.632	0.520
unintegrated	0.942	0.021	0.002	0.000	0.910	0.998	0.524	0.194	0.400	0.509	0.929	0.908	0.375	0.637	0.506
harmony	0.501	0.073	0.201	0.163	0.814	0.966	0.478	0.385	0.489	0.457	0.835	0.835	0.350	0.635	0.493

Table S2. Benchmarking data-integration methods across real-world datasets. For each dataset, we computed batch-correction and biological-label conservation metrics using the scIB package and probing classifiers. The table reports per-method metric values and summary scores.