Aperformance comparison with iCpSc to confirm the superiority of our PeakMatcth method.

We believe that the estimates by PeakMatch are accurate and reliable because they are generated based on the following empirical observations. That is,
(1) a peak on the bulk side can be explained somehow by a peak on the single cell side; and
(2) the orders of pseudo times should be preserved when they are assigned actual times.

Concretely, PeakMatch achieves (1) by giving a large weight to an edge connecting two nodes that are identified as peaks in many genes, and achieves (2) by estimating the actual time based on a non-crossing matching.

To support its "accuracy", we compared PeakMatch (PM) with iCpSc in terms of coefficient of determination (R^2), which is a typical criterion in the regression problem. We use two data sets:

- PM data: the data set used in our manuscript (2,217 genes); and
- iCpSc data: the data set used in the original paper of iCpSc (11,137 genes).

The regression model requires "correct time" when a pseudo time point really happens. PM data does not have such correct time, whereas iCpSc data has correct time, which is described in iCpSc/2CpToScTime/cptime.txt of the file collection. For PM data, we regard the sampling time as a tentative correct time for each pseudo time point. The program of PeakMatch outputs matching edges in terms of node index. We scale the bulk node index into [-24, 84] for PM data (resp., [0, 216] for iCpSc data) because the starting/finishing sampling times are -24 and 84 (resp., 0 and 216), and use the scaled time as the estimate.

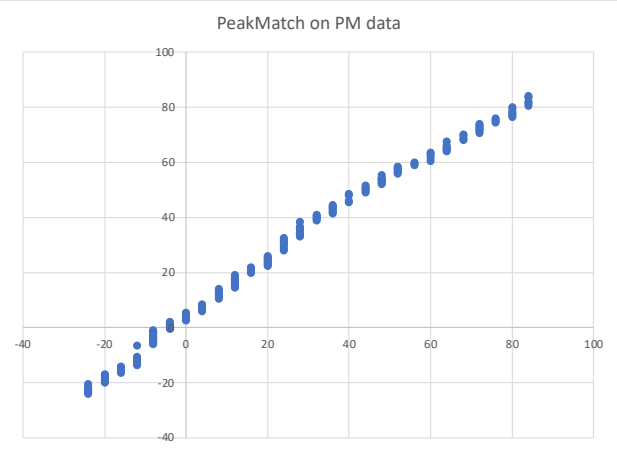
We show the results in Table 1*. We ran PeakMatch without order constraints for consideration. Apparently, PeakMatch with standard settings outperforms the others in the sense that the R^2 values are closer to 1.0 than others. In Figures 1 and 2, we present scatter plots for PM data and iCpSc data respectively, where each point corresponds to a pseudo time point, the x-axis indicates the correct time, and the y-axis indicates the estimated time. It is no wonder that Figure 2(c) (i.e., iCpSc on iCpSc data) is similar to Figure 2b of the iCpSc paper (Sun et al., 2017) since they should represent the same result.

*Note
The coefficient of determination can take a value less than -1 if the predicted and measured values are far off. In this case, the prediction of the actual time of iCpsc was so far off that the value

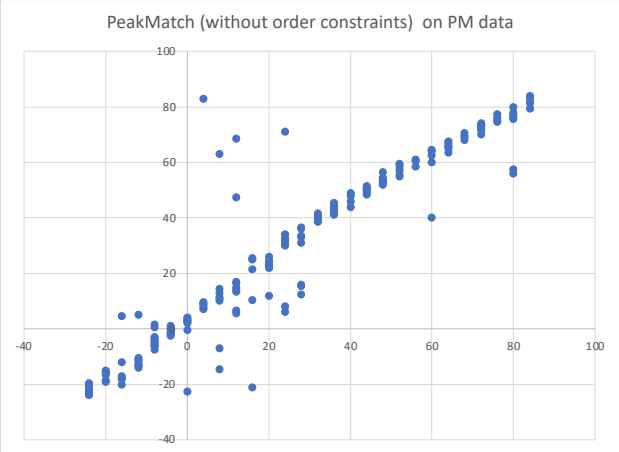
Table 1. R^2 values

	PeakMatch	PeakMatch (without order constraints)	iCpSc
PM data	0.978	0.864	-18.413
iCpSc data	0.651	0.349	-2.636

(a)



(b)



(c)

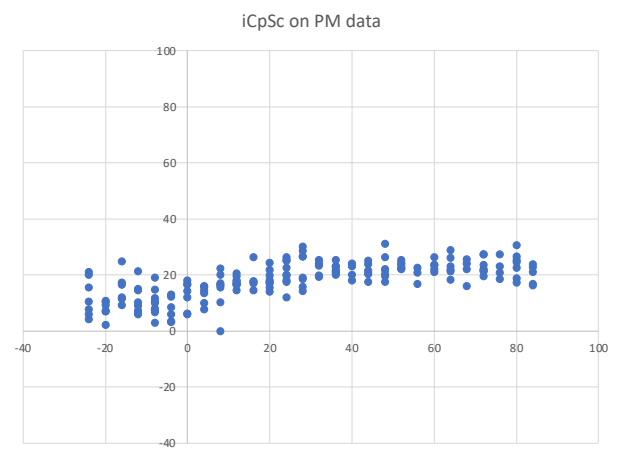
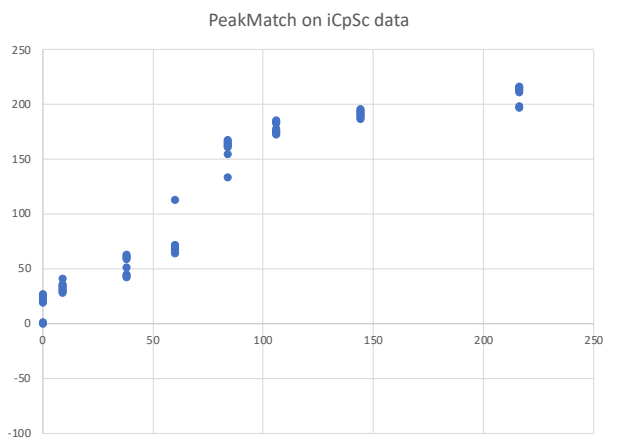
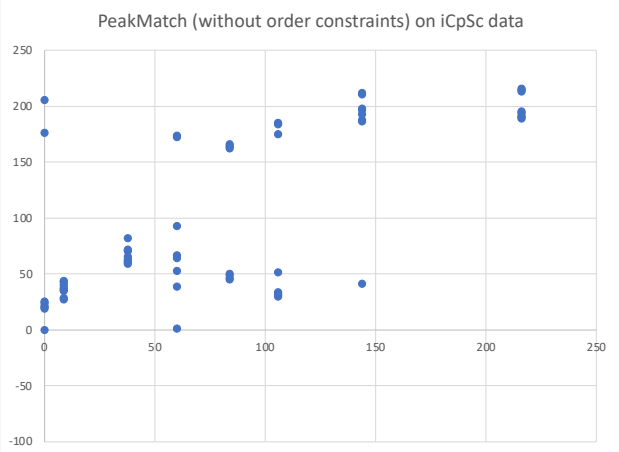


Figure 1. Scatter plots for PM data: (a) PeakMatch; (b) PeakMatch without order constraints; (c) iCpSc

(a)



(b)



(c)

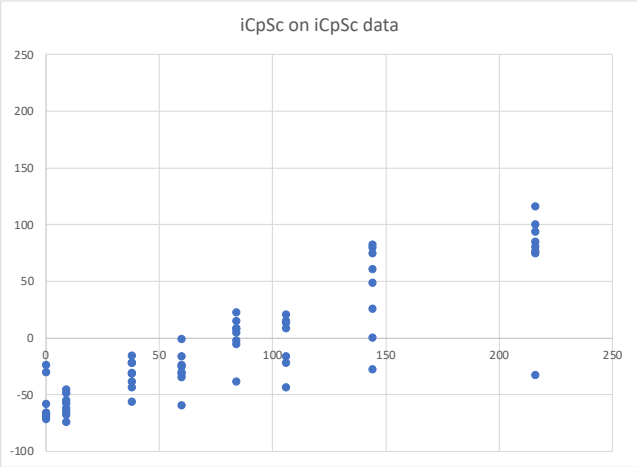


Figure 2. Scatter plots for iCpSc data: (a) PeakMatch; (b) PeakMatch without order constraints; (c) iCpSc

Here we report results on robustness to cell sampling and on effects of parameters T and inter.

Cell sampling

Let us see how the evaluation values change with respect to the gene set. We sample a fixed sized subset of 2,217 genes randomly, compute a noncrossing matching by PeakMatch and observe the evaluation values, say R^2, complete ratio*, and coverage ratio, where the definitions of the latter two are mentioned in the document at our github site (https://github.com/endo-lab/PeakMatch). Roughly speaking, complete ratio represents the ratio of genes such that all peaks on the single cell side are matched to peaks on the bulk side, and coverage ratio represents the ratio of peaks on the single cell side that are matched to peaks on the bulk side over all genes.

We show the results in Figure 3. For each size, we sampled a gene set 100 times. As can be seen, all evaluation values are somewhat saturated when the gene set size is no less than 200. This indicates that the matching should not change significantly, at least in the sense of these three evaluation functions, when the gene set size is no less than 200. Thousands of genes may not be needed for the task of estimating real time of pseudo time series accurately.

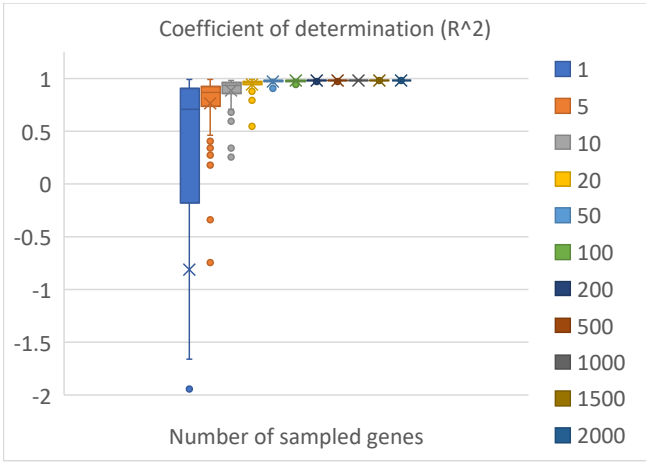
* Note

Both the complete ratio and the coverage ratio are indices to measure the accuracy of estimating the real time of the pseudotime by matching.

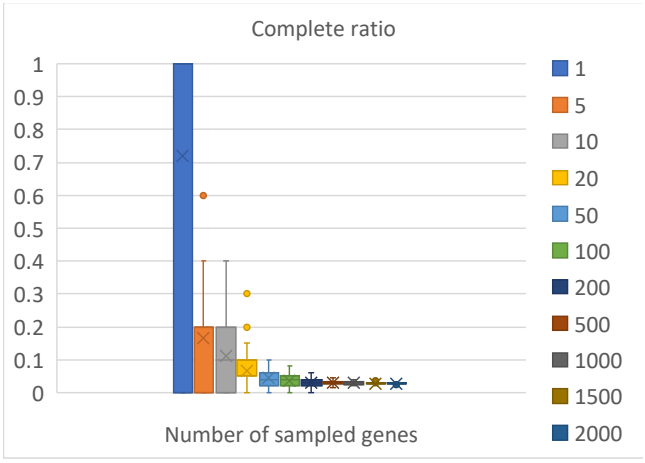
The former is the ratio of how many genes match all the pseudotime peaks to the real time peaks.

The latter is the average over all genes of the following values: the percentage of pseudotime peaks that match the real time peaks.

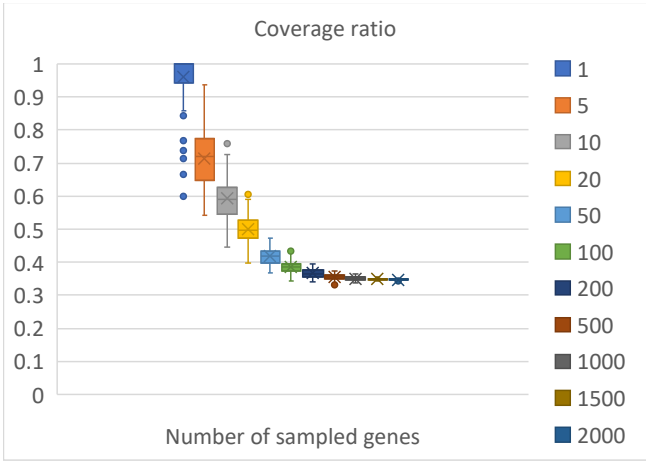
The former is considered the "tougher" measure, since all pseudo-temporal peaks are required to match.



(a)



(b)



(c)

Figure 3. Change of evaluation values with respect to the gene set size; (a) R^2; (b) complete ratio; (c) coverage ratio

Parameters T and inter

We observe how parameters T and inter affects the results.

First, let us observe how the number of peaks in the bipartite graph changes with respect to the two parameters. See Table 2. Parameter T represents the strictness of the definition of peak. We see that the number of peaks decreases as T gets larger, both for the single side and the bulk side. Parameter inter is the number of points that are inserted between consecutive bulk times for interpolation. This parameter is concerned with only the bulk side. We see that the number of peaks on the bulk side increases as inter gets larger.

Table 2. Change of the number of points near a peak in the bipartite graph with respect to parameters T and inter

		Parameter inter									
		1	3	5	7	9	11	13	15	17	19
Parameter T	(# Points near a peak on single side)	(# Points near a peak on bulk side)									
0.5	25.0	10.2	19.6	29.0	38.3	47.5	56.5	65.3	73.9	82.3	90.3
0.75	19.4	8.2	15.3	22.5	29.5	36.3	43.0	49.4	55.8	61.8	67.6
1	15.5	6.7	12.3	17.8	23.2	28.4	33.5	38.4	43.2	47.8	52.2
1.25	12.7	5.5	10.0	14.3	18.6	22.6	26.6	30.4	34.1	37.7	41.2
1.5	10.4	4.5	8.2	11.6	15.0	18.2	21.4	24.4	27.4	30.1	32.8
1.75	8.7	3.7	6.7	9.5	12.2	14.8	17.3	19.7	22.0	24.3	26.4
2	7.2	3.0	5.5	7.7	9.9	12.0	14.0	15.9	17.8	19.6	21.5
2.25	6.1	2.4	4.4	6.3	8.0	9.7	11.3	12.9	14.5	16.0	17.5
2.5	5.1	1.9	3.6	5.1	6.6	7.9	9.2	10.5	11.8	13.0	14.3
2.75	4.4	1.5	2.9	4.1	5.3	6.4	7.5	8.6	9.6	10.7	11.7
3	3.8	1.2	2.3	3.3	4.3	5.1	6.1	7.0	7.9	8.7	9.6

Next, we observe how the three evaluation values (i.e., R^2, complete ratio, and coverage ratio) change with respect to T and inter. See Table 3. In this experiment, we used a full set of 2,217 genes.

- R^2 does not change much with respect to change of the parameters, in comparison with the other two evaluation functions.
- Complete ratio increases as both parameters get larger, whereas the coverage ratio increases as parameter T gets smaller and parameter inter gets larger. The point is that each gene contributes either 0 or 1 to the complete ratio, whereas it contributes a real value between 0 and 1 to the coverage ratio.
- When parameter T is smaller, there are more peaks on both sides. By the definition of the complete ratio, it is likely that, in many genes, "ALL" peaks on the single side are not matched to peaks on the bulk side. Such genes contribute nothing to the complete ratio but may count some non-zero values to the coverage ratio.
- On the other hand, when inter is small (e.g., 1 or 3), we see that the number of peaks on the bulk side is smaller than the number of peaks in the single side on average, regardless of parameter T. In many genes,

all peaks on the single side cannot be matched to peaks on the bulk side, which must be the reason why the complete ratio is small. Similarly, such genes cannot contribute a large value (i.e., close to 1) to the coverage ratio.

Table 3. Change of evaluation values with respect to parameters T and inter; (a) R^2; (b) complete ratio; (c) coverage ratio

(a)

	Parameter inter									
Parameter T	1	3	5	7	9	11	13	15	17	19
0.5	0.976	0.981	0.978	0.982	0.981	0.980	0.980	0.981	0.981	0.980
0.75	0.976	0.977	0.978	0.980	0.980	0.981	0.980	0.980	0.980	0.979
1	0.974	0.977	0.977	0.978	0.979	0.979	0.979	0.979	0.978	0.979
1.25	0.975	0.978	0.978	0.976	0.977	0.978	0.978	0.978	0.977	0.978
1.5	0.977	0.978	0.978	0.975	0.977	0.979	0.978	0.978	0.978	0.979
1.75	0.977	0.978	0.979	0.977	0.977	0.978	0.977	0.979	0.979	0.978
2	0.977	0.978	0.978	0.978	0.978	0.978	0.978	0.979	0.979	0.978
2.25	0.978	0.978	0.979	0.979	0.980	0.980	0.979	0.979	0.979	0.980
2.5	0.978	0.977	0.979	0.979	0.980	0.981	0.981	0.981	0.980	0.980
2.75	0.977	0.978	0.978	0.979	0.979	0.979	0.980	0.979	0.979	0.980
3	0.972	0.976	0.975	0.974	0.978	0.975	0.979	0.980	0.978	0.980

(b)

	Parameter inter									
Parameter T	1	3	5	7	9	11	13	15	17	19
0.5	0.000	0.007	0.018	0.023	0.021	0.020	0.022	0.022	0.022	0.020
0.75	0.000	0.006	0.022	0.025	0.026	0.026	0.025	0.024	0.024	0.022
1	0.001	0.008	0.022	0.028	0.029	0.028	0.028	0.024	0.026	0.025
1.25	0.001	0.008	0.023	0.032	0.030	0.030	0.031	0.028	0.029	0.027
1.5	0.001	0.008	0.023	0.028	0.033	0.033	0.030	0.031	0.031	0.031
1.75	0.004	0.013	0.028	0.031	0.037	0.040	0.040	0.041	0.041	0.042
2	0.001	0.017	0.037	0.040	0.047	0.048	0.051	0.052	0.054	0.055
2.25	0.002	0.018	0.041	0.047	0.049	0.052	0.055	0.059	0.061	0.060
2.5	0.001	0.021	0.032	0.041	0.051	0.052	0.052	0.055	0.059	0.062
2.75	0.003	0.019	0.033	0.042	0.046	0.055	0.060	0.068	0.071	0.074
3	0.005	0.022	0.034	0.047	0.052	0.063	0.065	0.068	0.073	0.077

(c)

		Parameter inter									
Parameter T		1	3	5	7	9	11	13	15	17	19
0.5		0.164	0.274	0.353	0.391	0.407	0.418	0.425	0.431	0.437	0.442
0.75		0.161	0.265	0.337	0.371	0.385	0.394	0.401	0.409	0.414	0.420
1		0.157	0.253	0.319	0.348	0.362	0.372	0.380	0.387	0.394	0.401
1.25		0.151	0.243	0.301	0.328	0.343	0.353	0.363	0.371	0.377	0.383
1.5		0.146	0.231	0.284	0.309	0.327	0.338	0.346	0.355	0.362	0.369
1.75		0.137	0.217	0.266	0.291	0.309	0.319	0.329	0.338	0.347	0.354
2		0.128	0.207	0.251	0.275	0.292	0.304	0.314	0.322	0.330	0.336
2.25		0.113	0.186	0.232	0.255	0.272	0.284	0.294	0.305	0.312	0.318
2.5		0.099	0.170	0.209	0.234	0.250	0.262	0.273	0.282	0.292	0.299
2.75		0.085	0.151	0.187	0.211	0.228	0.242	0.252	0.265	0.273	0.280
3		0.077	0.136	0.171	0.194	0.209	0.226	0.233	0.243	0.252	0.260

We compare the noncrossing matching obtained by PeakMatch with randomly generated ones, in terms of the above-mentioned three evaluation values (i.e., R^2, complete ratio, and coverage ratio). With a full set of 2,217 genes and parameters T=1 and inter=7, we generate 100 random noncrossing matchings and take the average and the standard deviation of the evaluation values. The result is summarized in the following table. Note that the result of PeakMatch is taken from Table 3.

	R^2	Complete ratio	Coverage ratio
PeakMatch	0.978	0.028	0.348
Random noncrossing matchings	0.972 ± 0.011	0.0206 ± 0.003	0.3028 ± 0.0151

As shown, the two matchings are not different that much in terms of R^2, whereas PeakMatch's complete ratio and coverage ratio are significantly larger than those of random matchings. We can say that it is not a trivial task to match many peaks on both sides, and that PeakMatch succeeds in doing this task to some extent.