

1, We first need to calculate similarities between the sections as described in the text, and from these we obtain a similarity table as shown below using only 5 sections.

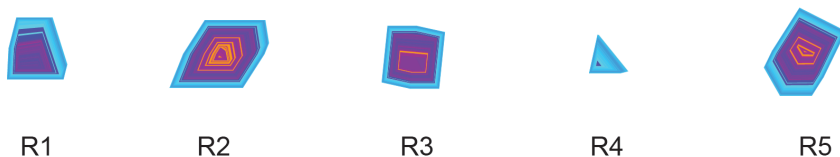


Figure S1 2D images

Similarity Table S1

	R1	R2	R3	R4	R5
R1	100	60	61	33	71
R2	60	100	76	24	83
R3	61	76	100	23	81
R4	33	24	23	100	28
R5	71	83	81	28	100

2, For each section we do a similarity histogram, see TableS2 and histogram FigS2 (Compare to Fig 8a).

Similarity Table S2

	R1	R2	R3	R4	R5
0-10%	0	0	0	0	0
10-20%	0	0	0	0	0
20-30%	0	1	1	3	1
30-40%	1	0	0	1	0
40-50%	0	0	0	0	0
50-60%	0	0	0	0	0
60-70%	2	1	1	0	0
70-80%	1	1	1	0	1
80-90%	0	1	1	0	2
90-100%	1	1	1	1	1

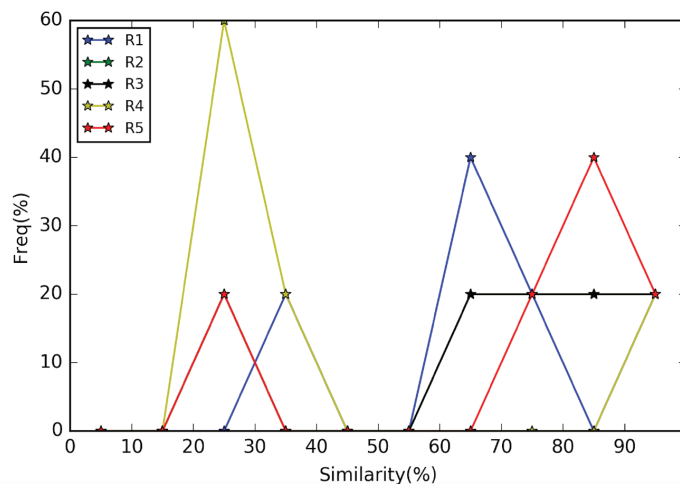


Figure S1 Similarity distributions

3, We calculate the cumulative histogram for each section for similarities =70-100 %.
Tables S3 and FigS3 (Compare to Fig8b and c).

Similarity Table S3

	R1	R2	R3	R4	R5
70-100%	2	3	3	1	4

From this analysis it is apparent that it is possible to identify the reference and ideal sections starting from a series of random sections.

4, We start with the same dataset shown in TableS1 and S2

5, We calculate the cumulative histogram for increasing similarity values starting from 50 until 100 as shown in TableS4.

Table S4

	R1	R2	R3	R4	R5
50	4	4	4	1	4
51	4	4	4	1	4
...					
80	1	2	2	1	3
...					
99	1	1	1	1	1
100	1	1	1	1	1

6, We calculate the similarity between the section with the highest values of occurrence for a given similarity threshold (starting from 50 until 100) and the RA, as shown in TableS5 and FigS4 (compare to figure 10). The distribution of similarities shows that we can indentify the REF and ideal sections (with similarities > 90 5) by choosing similarity thresholds of about 80 %.

Table S5

Threshold	50	51	...	80	...	99	100
S_RA*RI	76	76		89		76	76

