FR3D/BGSU Representative Set Benchmark (Length ≤ 400, non-rRNA) — Improved Pipeline

January 12, 2026

1 Benchmark definition

This report evaluates secondary-structure prediction quality on the FR3D/BGSU representative set under the following constraints:

- Truth is extracted from 3D structures using Barnaba canonical base-pair annotations (WC and GU).
- Targets are filtered to truth length ≤ 400 nucleotides and at least 5 canonical base pairs (to avoid degenerate empty-truth metrics).
- Ribosomal RNAs (rRNAs) are excluded (protein-partner confounders).
- Metrics are reported for the top-1 prediction and the best structure within the top-K list, with $K = 100$.

2 Reference run

The previous report is `docs/fr3d_full/main.pdf`. This report (`docs/fr3d_full_improved/main.pdf`) uses the same manifest/truth set, but predictions were regenerated with the updated predictor pipeline.

3 Pipeline changes

Changes relative to the previous run are concentrated in the predictor input handling and the selection/seeding strategy for weak-evidence targets:

1. **FASTA sanitization: preserve ambiguity.** The benchmark truth sequences can include `X` (unknown/modified base) and `&` (Barnaba fragment separator). The new pipeline preserves `X` and only replaces tool-invalid separators like `&` with `N` for downstream tools (while keeping length stable).
    - Predictor sanitization: `benchmark_runner/src/ssbench/predict/cacofold_mcmc_pipeline.py`
    - CLI FASTA writing: `benchmark_runner/src/ssbench/cli.py`
2. **Seed-boost gating: only when needed.** "Seed boost" (injecting extra thermo-derived seeds) now triggers only when sanitization had to replace illegal characters (e.g. `&`) or when covariation/model evidence is missing, rather than triggering on benign ambiguity (`X`).

3. **No-Rfam-hit selection balance.** For `no_rfam_hit` targets (fallback MFE scaffold), the selector preserves a larger prefix of refined candidates while still reserving budget for diverse thermo seeds, and the thermo seed pick is spread across a mid-energy band (rather than only the earliest AllSub candidates).

4 Results

Table 1: Comparison to previous run (`docs/fr3d_full/`). Positive Δ indicates improvement in this report (`docs/fr3d_full_improved/`).

Metric	Previous	Improved	Δ
F1@1 mean	0.636	0.640	+0.004
F1@100 mean	0.784	0.794	+0.010
F1@1 median	0.701	0.704	+0.003
F1@100 median	0.872	0.875	+0.003
MCC@1 mean	0.644	0.647	+0.004
MCC@100 mean	0.790	0.800	+0.010
Frac(F1@100 ≥ 0.7)	73.13%	74.63%	+1.49%
Frac(F1@100 ≥ 0.9)	43.03%	42.79%	-0.25%
Median rank(best@100)	29.0	30.0	+1.0
P90 rank(best@100)	85.0	87.0	+2.0
Frac(rank(best@100) ≤ 10)	30.47%	30.22%	-0.25%
ΔF1 mean	0.148	0.155	+0.006
Count(F1@1 $= 0$)	37	36	-1
Count(F1@100 $= 0$)	10	3	-7
Count(F1@100 ≤ 0.2)	23	14	-9

Table 2: Overall performance on the FR3D/BGSU representative benchmark (N=804, best-of-100).

Metric	Top-1	Best-of-100
Mean Precision	0.593	0.764
Mean Recall	0.731	0.854
Mean F1	0.640	0.794
Median F1	0.704	0.875
Mean MCC	0.647	0.800
Mean ΔF1	0.155	
Min #preds	10	
Median rank(best@100)	30.0	
Frac(rank(best@100)≤ 10)	30.22%	
Frac(ΔF1≥ 0.3)	14.18%	
Frac(best-of-100 F1≥ 0.7)	74.63%	
Frac(best-of-100 F1≥ 0.9)	42.79%	

Table 3: Performance by truth length bucket (means).

Bucket	N	P@1	R@1	F1@1	MCC@1	P@100	R@100	F1@100	MCC@100	ΔF1
30-80	587	0.616	0.759	0.665	0.673	0.804	0.891	0.834	0.839	0.169
81-150	149	0.586	0.698	0.627	0.632	0.730	0.811	0.758	0.762	0.131
151-300	53	0.399	0.531	0.441	0.451	0.498	0.620	0.536	0.546	0.095
301-400	15	0.403	0.638	0.468	0.491	0.486	0.669	0.537	0.555	0.069

Table 4: Performance split by whether Infernal found an Rfam CM hit (means).

Group	N	P@1	R@1	F1@1	MCC@1	P@100	R@100	F1@100	MCC@100	ΔF1
rfam_hit	716	0.598	0.743	0.648	0.656	0.763	0.856	0.795	0.801	0.147
no_rfam_hit	88	0.546	0.629	0.570	0.575	0.773	0.836	0.792	0.796	0.222

5 Distributions (CDF)

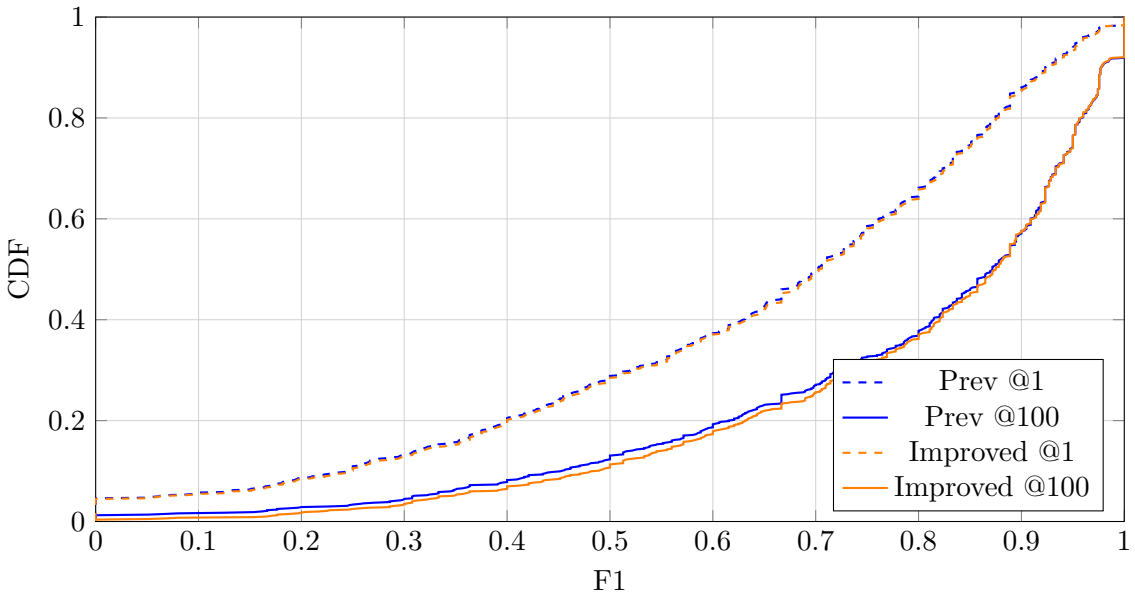

Figure 1: CDF of per-target F1 for top-1 and best-of-100 predictions, comparing `docs/fr3d_full/` (previous) vs `docs/fr3d_full_improved/` (improved).

6 Appendix: by-family breakdown

Table 5: Top Rfam IDs by frequency in this benchmark (means).

Rfam	N	F1@1	F1@100	ΔF1
RF00005	306	0.672	0.820	0.149
no_rfam_hit	88	0.570	0.792	0.222
RF00167	14	0.814	0.893	0.078
RF01852	11	0.718	0.871	0.153
RF00458	9	0.274	0.373	0.098
RF00386	9	0.818	0.895	0.077
RF00023	8	0.427	0.522	0.095
RF00027	8	0.695	0.757	0.062
RF00003	8	0.657	0.906	0.249
RF01684	7	0.792	0.895	0.104
RF03072	6	0.348	0.494	0.146
RF00017	6	0.488	0.576	0.088
RF00028	6	0.619	0.668	0.049
RF00026	6	0.564	0.831	0.267
RF00036	6	0.858	0.897	0.038

Table 8: Performance by Rfam ID using the rank-1 prediction (means).

Rfam	N	P@1	R@1	F1@1	MCC@1
RF00005	306	0.602	0.802	0.672	0.684
no_rfam_hit	88	0.546	0.629	0.570	0.575
RF00167	14	0.785	0.858	0.814	0.816
RF01852	11	0.708	0.751	0.718	0.721
RF00458	9	0.203	0.506	0.274	0.308
RF00386	9	0.806	0.854	0.818	0.822
RF00023	8	0.312	0.731	0.427	0.469
RF00027	8	0.595	0.888	0.695	0.715
RF00003	8	0.617	0.714	0.657	0.658
RF01684	7	0.779	0.824	0.792	0.794
RF03072	6	0.330	0.379	0.348	0.348
RF00017	6	0.459	0.534	0.488	0.490
RF00028	6	0.586	0.658	0.619	0.620
RF00026	6	0.569	0.592	0.564	0.569
RF00036	6	0.803	0.936	0.858	0.863
RF00061	5	0.416	0.436	0.423	0.422
RF00020	5	0.594	0.915	0.710	0.730
RF00442	5	0.540	0.610	0.569	0.569
RF00059	5	0.453	0.508	0.475	0.474
RF01750	5	0.877	0.915	0.885	0.889
RF02925	4	0.397	0.508	0.442	0.446
RF00379	4	0.396	0.460	0.425	0.424
RF02340	4	0.570	0.782	0.652	0.662

Rfam	N	P@1	R@1	F1@1	MCC@1
RF00162	4	0.547	0.678	0.597	0.602
RF00024	4	0.483	0.636	0.527	0.538
RF03013	4	0.593	0.562	0.572	0.570
RF01689	3	0.289	0.393	0.333	0.333
RF00009	3	0.595	0.738	0.637	0.650
RF00010	3	0.619	0.714	0.657	0.661
RF00011	3	0.532	0.599	0.563	0.563
RF01857	3	0.545	0.689	0.602	0.608
RF01051	3	0.660	0.618	0.637	0.635
RF01415	3	0.512	0.523	0.514	0.512
RF00168	3	0.788	0.754	0.770	0.769
RF00174	3	0.675	0.763	0.713	0.715
RF01716	3	0.628	0.785	0.695	0.698
RF01834	3	0.875	0.905	0.875	0.881
RF01739	3	0.707	0.938	0.802	0.811
RF00507	3	0.838	0.955	0.888	0.891
RF03165	3	0.762	0.784	0.772	0.770
RF00173	3	0.767	0.926	0.838	0.840
RF00080	3	0.767	0.818	0.790	0.790
RF00100	3	0.795	0.898	0.842	0.842
RF00169	3	0.504	0.546	0.523	0.520
RF00094	2	0.000	0.000	0.000	-0.009
RF01988	2	0.000	0.000	0.000	-0.008
RF00310	2	0.302	0.680	0.418	0.452
RF02477	2	0.340	0.590	0.431	0.446
RF01807	2	0.372	0.428	0.397	0.397
RF01854	2	0.320	0.473	0.371	0.380
RF03045	2	0.286	0.292	0.287	0.283
RF02348	2	0.272	0.326	0.296	0.294
RF02266	2	0.588	0.749	0.659	0.661
RF03016	2	0.676	0.593	0.631	0.631
RF03160	2	0.452	0.375	0.410	0.405
RF02136	2	0.499	0.655	0.565	0.570
RF01764	2	0.609	0.673	0.638	0.637
RF00233	2	0.463	0.519	0.489	0.487
RF00230	2	0.310	0.421	0.356	0.355
RF01831	2	0.657	0.750	0.697	0.698
RF01084	2	0.737	0.866	0.796	0.798
RF01792	2	0.400	0.500	0.444	0.443
RF00254	2	0.837	0.819	0.828	0.827
RF01734	2	0.717	0.867	0.774	0.781
RF00622	2	0.767	1.000	0.868	0.875
RF00015	2	0.683	0.938	0.790	0.798
RF01054	2	0.720	0.976	0.829	0.837
RF01510	2	0.859	0.805	0.830	0.829
RF03093	2	0.667	0.750	0.706	0.702

Rfam	N	P@1	R@1	F1@1	MCC@1
RF00008	2	0.893	0.913	0.899	0.900
RF02553	2	0.771	0.774	0.772	0.770
RF00170	2	0.607	1.000	0.750	0.774
RF01767	2	0.721	0.800	0.755	0.755
RF02339	2	0.562	0.690	0.619	0.618
RF00065	2	0.845	0.933	0.885	0.885
RF02977	2	1.000	0.971	0.985	0.985
RF01850	2	0.958	1.000	0.978	0.978
RF02803	2	0.900	1.000	0.947	0.948
RF02805	1	0.000	0.000	0.000	-0.006
RF01666	1	0.074	0.400	0.125	0.171
RF01316	1	0.043	0.067	0.053	0.048
RF02975	1	0.000	0.000	0.000	-0.007
RF02359	1	0.099	0.500	0.165	0.221
RF00806	1	0.000	0.000	0.000	-0.004
RF02295	1	0.150	0.429	0.222	0.250
RF02584	1	0.208	0.647	0.314	0.366
RF01554	1	0.211	0.500	0.296	0.319
RF01826	1	0.000	0.000	0.000	-0.010
RF00634	1	0.275	0.275	0.275	0.271
RF01057	1	0.154	0.222	0.182	0.178
RF02561	1	0.130	0.214	0.162	0.162
RF00030	1	0.287	0.360	0.320	0.320
RF03166	1	0.205	0.600	0.305	0.349
RF00172	1	0.417	0.556	0.476	0.478
RF03036	1	0.085	0.111	0.096	0.094
RF01713	1	0.303	0.909	0.455	0.523
RF01696	1	0.385	1.000	0.556	0.619
RF00051	1	0.150	0.182	0.164	0.163
RF00950	1	0.409	0.900	0.563	0.605
RF01379	1	0.391	0.600	0.474	0.482
RF01762	1	0.375	1.000	0.545	0.607
RF00649	1	0.422	0.514	0.463	0.464
RF00050	1	0.517	0.600	0.556	0.555
RF01838	1	0.000	0.000	0.000	-0.016
RF02056	1	0.533	0.727	0.615	0.622
RF00502	1	0.455	0.714	0.556	0.564
RF02922	1	0.562	0.600	0.581	0.578
RF03030	1	0.500	0.333	0.400	0.402
RF02032	1	0.394	0.410	0.402	0.401
RF01725	1	0.000	0.000	0.000	-0.007
RF01753	1	0.120	0.300	0.171	0.184
RF01357	1	0.500	0.857	0.632	0.652
RF01786	1	0.522	0.500	0.511	0.506
RF00166	1	0.486	1.000	0.654	0.696
RF00295	1	0.000	0.000	0.000	-0.017

Rfam	N	P@1	R@1	F1@1	MCC@1
RF02505	1	0.385	1.000	0.556	0.615
RF02802	1	0.543	0.826	0.655	0.668
RF03498	1	0.615	0.640	0.627	0.625
RF00488	1	0.645	0.772	0.703	0.705
RF02926	1	0.385	0.333	0.357	0.350
RF03542	1	0.400	0.353	0.375	0.368
RF02680	1	0.607	0.630	0.618	0.616
RF01406	1	0.500	0.636	0.560	0.561
RF03509	1	0.625	0.909	0.741	0.752
RF00175	1	0.727	0.615	0.667	0.665
RF01317	1	0.357	1.000	0.526	0.596
RF02914	1	0.846	0.688	0.759	0.760
RF00210	1	0.611	0.667	0.638	0.636
RF02519	1	0.500	1.000	0.667	0.703
RF01356	1	0.700	0.875	0.778	0.781
RF03977	1	0.534	0.574	0.554	0.552
RF01232	1	0.417	0.357	0.385	0.381
RF02681	1	0.609	0.609	0.609	0.604
RF00198	1	0.375	0.429	0.400	0.391
RF01773	1	0.636	0.933	0.757	0.770
RF00521	1	0.769	0.909	0.833	0.835
RF03085	1	0.714	0.833	0.769	0.769
RF00228	1	0.629	0.786	0.698	0.701
RF03570	1	0.000	0.000	0.000	-0.022
RF01519	1	0.643	0.818	0.720	0.723
RF00176	1	0.694	0.962	0.806	0.816
RF01763	1	0.385	0.417	0.400	0.391
RF01333	1	0.632	0.667	0.649	0.643
RF03404	1	0.789	0.909	0.845	0.846
RF00029	1	0.708	0.850	0.773	0.774
RF02001	1	0.937	0.832	0.881	0.883
RF01363	1	0.657	0.742	0.697	0.696
RF03081	1	0.778	0.875	0.824	0.824
RF03059	1	0.583	0.778	0.667	0.669
RF00785	1	1.000	0.600	0.750	0.774
RF01365	1	0.250	0.600	0.353	0.381
RF02869	1	0.421	1.000	0.593	0.647
RF03054	1	0.500	0.462	0.480	0.474
RF03918	1	0.889	0.571	0.696	0.706
RF03167	1	0.778	0.560	0.651	0.658
RF00898	1	0.800	0.800	0.800	0.798
RF00163	1	0.767	0.958	0.852	0.857
RF01836	1	0.235	0.267	0.250	0.240
RF00623	1	0.692	0.818	0.750	0.748
RF03736	1	0.714	1.000	0.833	0.842
RF02683	1	0.645	0.690	0.667	0.665

Rfam	N	P@1	R@1	F1@1	MCC@1
RF01541	1	0.970	0.865	0.914	0.915
RF01622	1	1.000	0.846	0.917	0.919
RF01544	1	0.200	0.222	0.211	0.201
RF02521	1	0.906	0.879	0.892	0.891
RF00897	1	1.000	0.850	0.919	0.921
RF02678	1	0.783	0.947	0.857	0.860
RF00044	1	0.636	0.737	0.683	0.682
RF01344	1	0.733	0.815	0.772	0.772
RF02447	1	0.800	0.762	0.780	0.779
RF01472	1	0.607	0.773	0.680	0.682
RF01050	1	0.778	0.875	0.824	0.822
RF02538	1	0.875	1.000	0.933	0.935
RF02837	1	0.778	1.000	0.875	0.881
RF02984	1	0.800	0.857	0.828	0.825
RF01083	1	0.889	0.889	0.889	0.888
RF02888	1	0.727	0.889	0.800	0.802
RF01082	1	0.385	0.556	0.455	0.457
RF03906	1	0.921	0.921	0.921	0.920
RF00419	1	0.857	0.857	0.857	0.856
RF01071	1	0.739	0.895	0.810	0.811
RF01107	1	0.909	1.000	0.952	0.953
RF04067	1	0.909	0.952	0.930	0.930
RF00102	1	0.915	1.000	0.956	0.956
RF00606	1	0.625	0.455	0.526	0.528
RF00504	1	0.920	0.920	0.920	0.919
RF02066	1	0.923	1.000	0.960	0.960
RF03863	1	0.852	0.885	0.868	0.867
RF00164	1	0.786	0.733	0.759	0.756
RF02871	1	0.743	0.897	0.812	0.815
RF00013	1	0.857	1.000	0.923	0.925
RF01241	1	0.882	1.000	0.938	0.939
RF02885	1	0.933	0.933	0.933	0.933
RF00380	1	0.902	0.920	0.911	0.911
RF00220	1	0.941	1.000	0.970	0.970
RF03900	1	1.000	0.889	0.941	0.941
RF03120	1	0.938	0.857	0.896	0.896
RF03790	1	1.000	0.947	0.973	0.973
RF00032	1	1.000	1.000	1.000	1.000
RF00234	1	0.857	1.000	0.923	0.925
RF00250	1	1.000	0.909	0.952	0.953
RF00381	1	0.000	0.000	0.000	-0.011
RF00455	1	0.944	1.000	0.971	0.971
RF00464	1	1.000	1.000	1.000	1.000
RF00480	1	0.846	0.917	0.880	0.879
RF02695	1	0.867	0.765	0.812	0.812
RF03031	1	1.000	0.933	0.966	0.965

Rfam	N	P@1	R@1	F1@1	MCC@1
RF03055	1	1.000	0.933	0.966	0.966
RF03339	1	0.895	1.000	0.944	0.945
RF03394	1	0.885	0.920	0.902	0.901
RF03493	1	1.000	0.833	0.909	0.912
RF04151	1	1.000	1.000	1.000	1.000
RF04217	1	1.000	0.923	0.960	0.960

Table 9: Performance by Rfam ID using best-of-100 (means).

Rfam	N	P@100	R@100	F1@100	MCC@100
RF00005	306	0.781	0.895	0.820	0.828
no_rfam_hit	88	0.773	0.836	0.792	0.796
RF00167	14	0.892	0.897	0.893	0.893
RF01852	11	0.867	0.887	0.871	0.873
RF00458	9	0.290	0.610	0.373	0.406
RF00386	9	0.861	0.941	0.895	0.897
RF00023	8	0.403	0.789	0.522	0.556
RF00027	8	0.659	0.915	0.757	0.770
RF00003	8	0.895	0.931	0.906	0.909
RF01684	7	0.898	0.898	0.895	0.895
RF03072	6	0.483	0.521	0.494	0.497
RF00017	6	0.570	0.588	0.576	0.576
RF00028	6	0.660	0.680	0.668	0.668
RF00026	6	0.777	0.913	0.831	0.837
RF00036	6	0.861	0.948	0.897	0.900
RF00061	5	0.640	0.617	0.625	0.625
RF00020	5	0.655	0.915	0.752	0.767
RF00442	5	0.741	0.777	0.754	0.755
RF00059	5	0.780	0.812	0.793	0.793
RF01750	5	0.941	0.970	0.955	0.955
RF02925	4	0.457	0.549	0.496	0.498
RF00379	4	0.515	0.559	0.533	0.533
RF02340	4	0.645	0.841	0.720	0.729
RF00162	4	0.705	0.822	0.749	0.755
RF00024	4	0.711	0.851	0.754	0.765
RF03013	4	0.909	0.888	0.894	0.895
RF01689	3	0.564	0.664	0.605	0.608
RF00009	3	0.643	0.702	0.666	0.669
RF00010	3	0.683	0.682	0.679	0.680
RF00011	3	0.668	0.707	0.687	0.686
RF01857	3	0.672	0.812	0.729	0.734
RF01051	3	0.784	0.747	0.762	0.762
RF01415	3	0.737	0.794	0.764	0.763
RF00168	3	0.804	0.819	0.811	0.810

Rfam	N	P@100	R@100	F1@100	MCC@100
RF00174	3	0.806	0.897	0.843	0.847
RF01716	3	0.848	0.897	0.858	0.864
RF01834	3	0.875	0.905	0.875	0.881
RF01739	3	0.836	0.979	0.897	0.901
RF00507	3	0.865	0.968	0.908	0.912
RF03165	3	0.921	0.909	0.915	0.914
RF00173	3	0.921	0.926	0.922	0.922
RF00080	3	0.928	0.942	0.934	0.934
RF00100	3	0.933	0.972	0.952	0.952
RF00169	3	0.912	1.000	0.953	0.954
RF00094	2	0.130	0.150	0.139	0.133
RF01988	2	0.293	0.700	0.411	0.448
RF00310	2	0.349	0.780	0.482	0.520
RF02477	2	0.407	0.618	0.486	0.498
RF01807	2	0.505	0.534	0.519	0.518
RF01854	2	0.498	0.580	0.521	0.528
RF03045	2	0.500	0.646	0.561	0.563
RF02348	2	0.590	0.708	0.635	0.640
RF02266	2	0.625	0.749	0.681	0.682
RF03016	2	0.745	0.660	0.698	0.699
RF03160	2	0.750	0.659	0.698	0.698
RF02136	2	0.657	0.770	0.705	0.708
RF01764	2	0.776	0.743	0.759	0.758
RF00233	2	0.815	0.722	0.764	0.765
RF00230	2	0.718	0.868	0.781	0.786
RF01831	2	0.839	0.819	0.828	0.827
RF01084	2	0.779	0.907	0.838	0.840
RF01792	2	0.944	0.800	0.846	0.858
RF00254	2	0.886	0.888	0.886	0.886
RF01734	2	0.875	0.917	0.892	0.893
RF00622	2	0.842	0.969	0.901	0.902
RF00015	2	0.838	1.000	0.911	0.914
RF01054	2	0.865	1.000	0.927	0.929
RF01510	2	0.952	0.912	0.932	0.931
RF03093	2	0.889	1.000	0.941	0.942
RF00008	2	0.929	0.977	0.950	0.951
RF02553	2	0.958	0.958	0.958	0.958
RF00170	2	0.929	1.000	0.962	0.963
RF01767	2	0.929	1.000	0.962	0.962
RF02339	2	0.929	1.000	0.962	0.963
RF00065	2	0.967	1.000	0.983	0.983
RF02977	2	1.000	0.971	0.985	0.985
RF01850	2	1.000	1.000	1.000	1.000
RF02803	2	1.000	1.000	1.000	1.000
RF02805	1	0.000	0.000	0.000	-0.006
RF01666	1	0.105	0.400	0.167	0.204

Rfam	N	P@100	R@100	F1@100	MCC@100
RF01316	1	0.158	0.200	0.176	0.173
RF02975	1	0.188	0.188	0.188	0.183
RF02359	1	0.123	0.562	0.202	0.263
RF00806	1	0.229	0.320	0.267	0.267
RF02295	1	0.231	0.429	0.300	0.312
RF02584	1	0.226	0.618	0.331	0.373
RF01554	1	0.286	0.500	0.364	0.374
RF01826	1	0.308	0.444	0.364	0.363
RF00634	1	0.390	0.400	0.395	0.392
RF01057	1	0.333	0.556	0.417	0.425
RF02561	1	0.400	0.571	0.471	0.475
RF00030	1	0.448	0.520	0.481	0.482
RF03166	1	0.342	0.867	0.491	0.544
RF00172	1	0.455	0.556	0.500	0.500
RF03036	1	0.447	0.583	0.506	0.509
RF01713	1	0.357	0.909	0.513	0.569
RF01696	1	0.385	1.000	0.556	0.619
RF00051	1	0.553	0.636	0.592	0.592
RF00950	1	0.450	0.900	0.600	0.634
RF01379	1	0.600	0.600	0.600	0.599
RF01762	1	0.429	1.000	0.600	0.650
RF00649	1	0.656	0.568	0.609	0.609
RF00050	1	0.625	0.600	0.612	0.611
RF01838	1	1.000	0.444	0.615	0.663
RF02056	1	0.533	0.727	0.615	0.622
RF00502	1	0.556	0.714	0.625	0.625
RF02922	1	0.655	0.633	0.644	0.642
RF03030	1	0.688	0.611	0.647	0.644
RF02032	1	0.685	0.630	0.656	0.656
RF01725	1	0.655	0.679	0.667	0.665
RF01753	1	0.526	1.000	0.690	0.724
RF01357	1	0.600	0.857	0.706	0.715
RF01786	1	0.667	0.750	0.706	0.704
RF00166	1	0.548	1.000	0.708	0.740
RF00295	1	0.625	0.833	0.714	0.717
RF02505	1	0.556	1.000	0.714	0.742
RF02802	1	0.600	0.913	0.724	0.739
RF03498	1	0.842	0.640	0.727	0.732
RF00488	1	0.692	0.783	0.735	0.736
RF02926	1	0.833	0.667	0.741	0.742
RF03542	1	0.722	0.765	0.743	0.740
RF02680	1	0.750	0.778	0.764	0.762
RF01406	1	0.667	0.909	0.769	0.777
RF03509	1	0.667	0.909	0.769	0.777
RF00175	1	0.769	0.769	0.769	0.767
RF01317	1	0.625	1.000	0.769	0.790

Rfam	N	P@100	R@100	F1@100	MCC@100
RF02914	1	0.917	0.688	0.786	0.791
RF00210	1	0.737	0.848	0.789	0.789
RF02519	1	0.667	1.000	0.800	0.814
RF01356	1	0.697	0.958	0.807	0.816
RF03977	1	0.780	0.852	0.814	0.814
RF01232	1	0.846	0.786	0.815	0.814
RF02681	1	1.000	0.696	0.821	0.832
RF00198	1	0.700	1.000	0.824	0.834
RF01773	1	0.737	0.933	0.824	0.829
RF00521	1	0.769	0.909	0.833	0.835
RF03085	1	0.833	0.833	0.833	0.832
RF00228	1	0.828	0.857	0.842	0.841
RF03570	1	0.889	0.800	0.842	0.840
RF01519	1	0.826	0.864	0.844	0.844
RF00176	1	0.743	1.000	0.852	0.861
RF01763	1	1.000	0.750	0.857	0.864
RF01333	1	0.842	0.889	0.865	0.863
RF03404	1	0.816	0.939	0.873	0.874
RF00029	1	0.857	0.900	0.878	0.877
RF02001	1	0.937	0.832	0.881	0.883
RF01363	1	0.929	0.839	0.881	0.882
RF03081	1	0.833	0.938	0.882	0.883
RF03059	1	0.889	0.889	0.889	0.887
RF00785	1	1.000	0.800	0.889	0.894
RF01365	1	1.000	0.800	0.889	0.894
RF02869	1	0.800	1.000	0.889	0.894
RF03054	1	0.857	0.923	0.889	0.888
RF03918	1	0.923	0.857	0.889	0.886
RF03167	1	0.955	0.840	0.894	0.895
RF00898	1	0.900	0.900	0.900	0.899
RF00163	1	0.852	0.958	0.902	0.903
RF01836	1	0.875	0.933	0.903	0.902
RF00623	1	0.909	0.909	0.909	0.907
RF03736	1	0.833	1.000	0.909	0.911
RF02683	1	0.929	0.897	0.912	0.912
RF01541	1	0.970	0.865	0.914	0.915
RF01622	1	1.000	0.846	0.917	0.919
RF01544	1	0.895	0.944	0.919	0.918
RF02521	1	0.967	0.879	0.921	0.921
RF00897	1	0.947	0.900	0.923	0.923
RF02678	1	0.900	0.947	0.923	0.923
RF00044	1	0.864	1.000	0.927	0.929
RF01344	1	0.897	0.963	0.929	0.929
RF02447	1	0.909	0.952	0.930	0.930
RF01472	1	0.913	0.955	0.933	0.933
RF01050	1	1.000	0.875	0.933	0.934

Rfam	N	P@100	R@100	F1@100	MCC@100
RF02538	1	0.875	1.000	0.933	0.935
RF02837	1	0.875	1.000	0.933	0.935
RF02984	1	0.875	1.000	0.933	0.934
RF01083	1	1.000	0.889	0.941	0.942
RF02888	1	1.000	0.889	0.941	0.942
RF01082	1	0.900	1.000	0.947	0.948
RF03906	1	0.925	0.974	0.949	0.949
RF00419	1	1.000	0.905	0.950	0.951
RF01071	1	0.905	1.000	0.950	0.951
RF01107	1	0.909	1.000	0.952	0.953
RF04067	1	0.952	0.952	0.952	0.952
RF00102	1	0.915	1.000	0.956	0.956
RF00606	1	0.917	1.000	0.957	0.957
RF00504	1	0.960	0.960	0.960	0.960
RF02066	1	0.923	1.000	0.960	0.960
RF03863	1	1.000	0.923	0.960	0.960
RF00164	1	1.000	0.933	0.966	0.966
RF02871	1	0.935	1.000	0.967	0.967
RF00013	1	0.938	1.000	0.968	0.968
RF01241	1	0.938	1.000	0.968	0.968
RF02885	1	0.938	1.000	0.968	0.968
RF00380	1	1.000	0.940	0.969	0.969
RF00220	1	0.941	1.000	0.970	0.970
RF03900	1	1.000	0.944	0.971	0.971
RF03120	1	0.946	1.000	0.972	0.972
RF03790	1	1.000	0.947	0.973	0.973
RF00032	1	1.000	1.000	1.000	1.000
RF00234	1	1.000	1.000	1.000	1.000
RF00250	1	1.000	1.000	1.000	1.000
RF00381	1	1.000	1.000	1.000	1.000
RF00455	1	1.000	1.000	1.000	1.000
RF00464	1	1.000	1.000	1.000	1.000
RF00480	1	1.000	1.000	1.000	1.000
RF02695	1	1.000	1.000	1.000	1.000
RF03031	1	1.000	1.000	1.000	1.000
RF03055	1	1.000	1.000	1.000	1.000
RF03339	1	1.000	1.000	1.000	1.000
RF03394	1	1.000	1.000	1.000	1.000
RF03493	1	1.000	1.000	1.000	1.000
RF04151	1	1.000	1.000	1.000	1.000
RF04217	1	1.000	1.000	1.000	1.000

Table 6: Worst targets by best-of-100 F1 (lower is worse).

Target	Rfam	L	$\lvert P_{ref} \rvert$	F1@1	F1@100	Rank(best)
9FXO|1|L6	no_rfam_hit	71	9	0.000	0.000	77
6ZJ3|1|LL	RF02805	91	24	0.000	0.000	5
2B63|1|R	no_rfam_hit	31	9	0.000	0.000	3
4PRF|1|B	RF00094	74	16	0.000	0.051	73
9BUQ|1|S7	RF00005	74	8	0.000	0.069	93
7NWG|1|33	RF00005	75	6	0.069	0.080	70
7OBQ|1|1	RF00017	165	27	0.132	0.154	91
8A22|1|A1	RF01666	109	5	0.125	0.167	46
7V9A|1|R	RF00024	146	9	0.151	0.170	54
7A01|1|E1	RF00458	153	12	0.000	0.176	18
8ESQ|1|6	RF01316	81	15	0.053	0.176	62
8JDJ|1|B	RF02975	74	16	0.000	0.188	59

Table 7: Targets with the largest improvement from rank-1 to best-of-100 (ranking/sampling sensitivity).

Target	Rfam	L	F1@1	F1@100	ΔF1	Rank(best)
7M5O|1|B	RF00381	37	0.000	1.000	1.000	27
6E8U|1|B	no_rfam_hit	36	0.000	0.923	0.923	37
7JRR|1|A	no_rfam_hit	50	0.000	0.917	0.917	91
8FN6|1|g	no_rfam_hit	46	0.000	0.909	0.909	4
9ERF|1|T	RF00005	48	0.000	0.857	0.857	17
2XXA|1|F	RF00169	102	0.088	0.941	0.853	81
9RVP|1|B	RF03570	30	0.000	0.842	0.842	6
7D7V|1|A	RF03013	57	0.000	0.812	0.812	81
8EUY|1|6	RF01792	56	0.000	0.750	0.750	79
4O26|1|E	RF00024	47	0.194	0.933	0.740	37
8QEQ|1|R	RF00295	30	0.000	0.714	0.714	50
1KUQ|1|B	RF01544	57	0.211	0.919	0.708	81

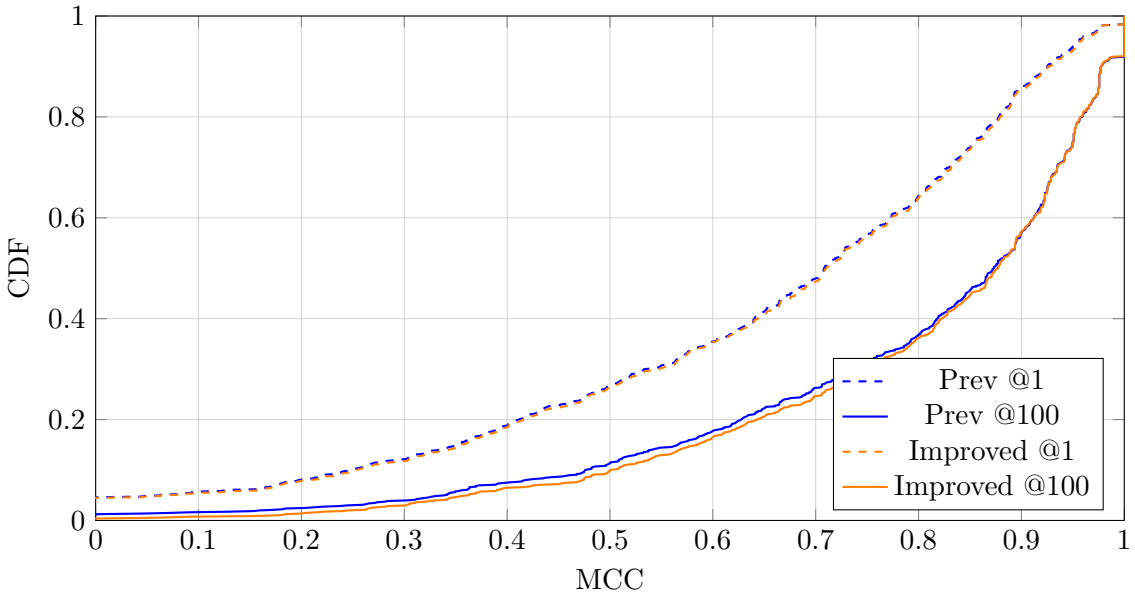

Figure 2: CDF of per-target MCC for top-1 and best-of-100 predictions, comparing `docs/fr3d_full/` (previous) vs `docs/fr3d_full_improved/` (improved).