

Table S2. Summary of probes on GeoChip 5.0M based on the phylogenetic distribution of the functional genes.

Domain	Kingdom	Phylum (order for virus)	No. of class	No. of order	No. of family	No. of genus	No. of species	No. of strains	No. of gene	No. of covered CDS	No. of total probe	No. of group-specific probe	No. of sequence-specific probe	
Archaea	Archaea	Crenarchaeota	2	6	9	22	43	57	142	13921	1312	1133	179	
		Euryarchaeota	10	12	23	69	133	187	229	33323	3992	3625	367	
		Korarchaeota	1	1	1	1	1	1	1	33	3718	35	33	2
		Nanoarchaeota	1	1	1	1	1	1	1	6	12	6	6	0
		Thaumarchaeota	1	4	5	7	9	14	46	5218	95	73	22	
		unclassified	1	1	1	1	1	25	46	1885	288	115	173	
		Archaea Total	16	25	40	101	188	285	502	58077	5728	4985	743	
Bacteria	Bacteria	Acidobacteria	5	4	4	8	13	18	155	20180	745	715	30	
		Actinobacteria	1	8	48	120	351	753	381	80060	15394	13618	1776	
		Aquificae	1	2	4	12	17	22	127	19379	533	498	35	
		Bacteroidetes	7	10	23	107	239	394	267	58970	9661	8790	871	
		Caldiserica	1	1	1	1	1	1	13	38	19	19	0	
		Candidatus poribacteria	1	1	1	1	1	1	2	5	4	1	3	
		Candidatus saccharibacteria	1	1	1	1	1	3	9	22	17	4	13	
		Chlamydiae	2	1	4	7	15	23	74	12652	241	230	11	
		Chlorobi	1	1	1	6	15	20	138	21932	995	944	51	
		Chloroflexi	8	11	12	14	19	28	209	32839	1353	1289	64	
		Chrysiogenetes	1	1	1	2	2	2	33	81	41	40	1	
		Cyanobacteria	2	8	12	57	97	203	237	46219	4727	4209	518	
		Deferribacteres	1	1	1	4	4	4	81	2583	165	154	11	
		Deinococcus-thermus	2	3	4	7	18	27	172	21352	1026	978	48	
		Dictyoglomi	1	1	1	1	2	2	51	9447	99	98	1	
		Elusimicrobia	2	2	2	2	2	2	34	5713	50	49	1	
		Fibrobacteres	2	1	1	1	1	2	50	6671	69	64	5	

		Firmicutes	6	10	42	176	593	1556	429	107993	25601	21921	3680
		Fusobacteria	2	1	2	5	15	38	111	17980	541	493	48
		Gemmatimonadetes	1	1	1	1	1	1	51	8772	71	69	2
		Ignavibacteriae	1	1	2	2	2	2	27	85	43	42	1
		Lentisphaerae	2	3	2	2	2	2	56	4823	93	91	2
		Nitrospirae	2	2	2	4	5	10	114	8777	237	197	40
		Planctomycetes	3	3	4	15	23	32	183	26590	994	765	229
		Poribacteria	1	1	1	1	1	1	13	25	13	12	1
		Proteobacteria	7	47	115	508	1132	2855	823	187705	68319	59749	8570
		Spirochaetes	2	1	3	8	48	80	172	11471	1007	829	178
		Synergistetes	2	2	2	11	13	14	109	11410	472	392	80
		Tenericutes	1	3	4	7	39	82	56	11489	266	218	48
		Thermodesulfobacteria	1	1	1	2	5	6	50	162	85	77	8
		Thermotogae	1	1	2	8	19	26	93	16870	588	553	35
		Verrucomicrobia	4	6	8	14	17	30	172	29149	947	824	123
		unclassified	1	2	2	7	8	239	233	33937	6737	2583	4154
		Bacteria Total	76	142	314	1122	2721	6479	4725	815381	141153	120515	20638
Eukaryota	Fungi	Ascomycota	10	25	70	164	245	402	201	19120	6995	2471	4524
		Basidiomycota	7	24	61	104	122	175	119	8732	1549	530	1019
		Chytridiomycota	1	2	2	1	2	2	7	10	10	0	10
		Glomeromycota	1	2	2	3	3	3	9	71	12	7	5
		Microsporidia	1	1	7	9	14	16	25	1900	82	47	35
		Neocallimastigomycota	1	1	1	3	3	8	7	136	28	4	24
		unclassified	1	4	10	13	15	20	19	725	180	90	90
		Fungi Total	22	59	153	297	404	626	387	30694	8856	3149	5707
	Protist	Apicomplexa	2	3	6	7	19	47	44	1875	374	123	251
		Bacillariophyta	5	10	14	16	19	31	37	1269	218	96	122
		Chromerida	1	1	1	2	2	2	3	6	5	1	4

		Euglenida	1	4	6	6	6	7	3	17	12	4	8
		Eustigmatophyceae	1	1	1	1	1	1	1	1	1	0	1
		Haplosporidia	1	1	2	4	4	5	1	12	6	3	3
		Phaeophyceae	1	2	2	2	2	2	19	1129	28	6	22
		Pinguiophyceae	1	1	1	1	1	1	1	1	1	0	1
		Xanthophyceae	1	2	2	2	2	2	4	7	5	2	3
		Protist Total	14	25	35	41	56	98	113	4317	650	235	415
	Metazoa	Arthropoda	2	5	5	7	7	7	4	19	13	5	8
		Chordata	2	3	3	3	3	3	3	9	7	2	5
		Echinodermata	1	1	1	1	1	1	1	2	3	1	2
		Nematoda	1	2	3	3	3	4	6	9	8	2	6
		Platyhelminthes	1	1	1	1	1	1	1	5	3	2	1
		Metazoa Total	7	12	13	15	15	16	15	44	34	12	22
	Viridiplantae	Chlorophyta	7	15	23	37	39	57	39	694	444	221	223
		Streptophyta	7	11	12	12	12	13	15	432	31	12	19
		Viridiplantae Total	14	26	35	49	51	70	54	1126	475	233	242
	unclassified	unclassified	24	74	116	178	195	264	165	3148	1401	569	832
Viruses	Viruses	Nidovirales	-	-	1	7	18	75	2	259	143	60	83
		Picornavirales	-	-	6	24	47	275	7	707	518	144	374
		Tymovirales	-	-	1	6	8	46	2	236	111	79	32
		Caudovirales	-	-	4	26	30	156	40	1506	347	279	68
		unclassified	-	-	28	104	208	814	78	3320	1729	768	961
		Virus Total			156	345	506	1630	294	9176	4249	1899	2350
unclassified	-	-	-	-	-	-	-	33	125	2561	816	293	523

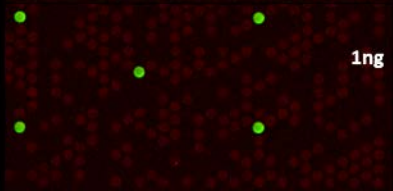
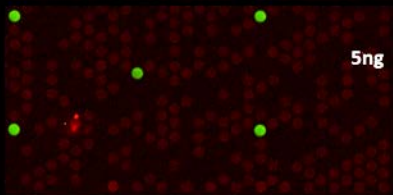
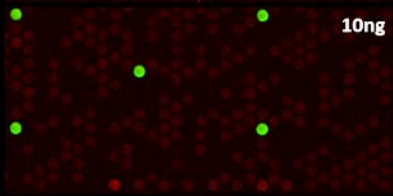
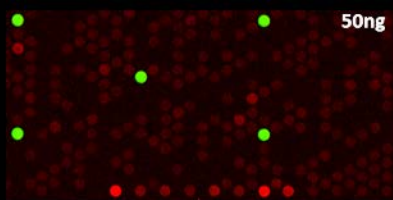
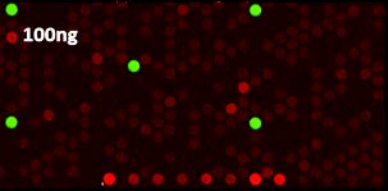
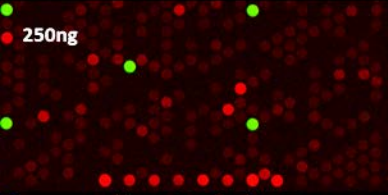
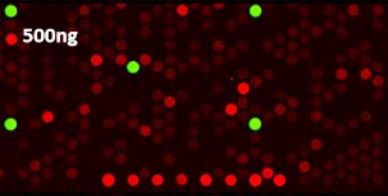
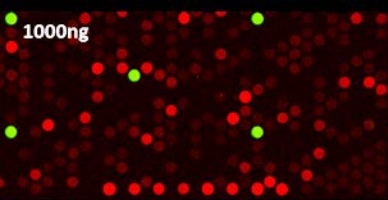
Table S3. Overall differences between GeoChip 4 and 5.

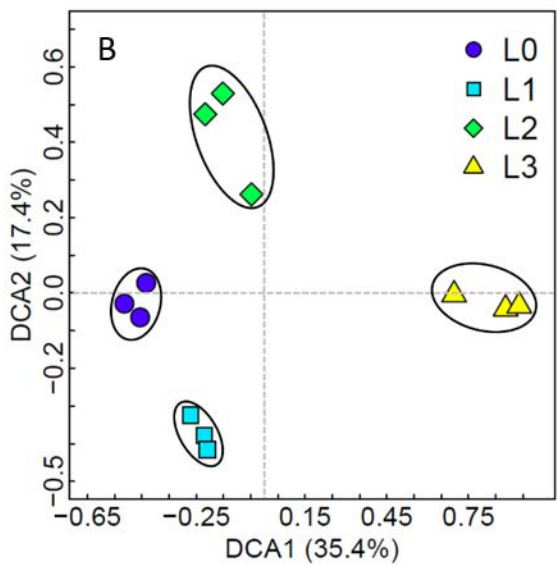
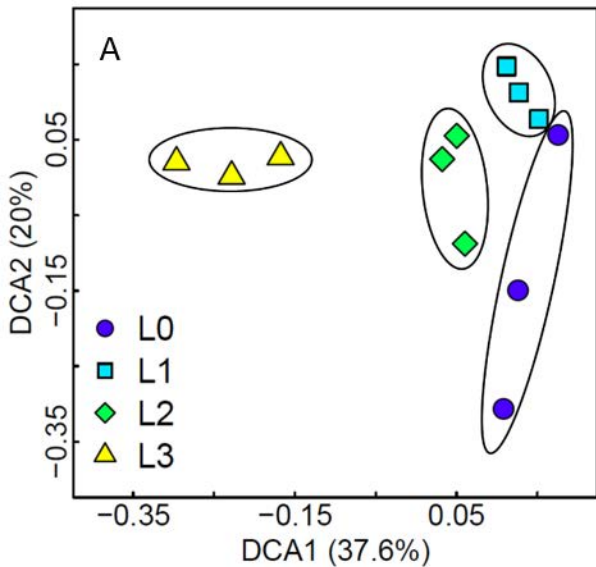
Entry	GeoChip 4	GeoChip 5.0S	GeoChip 5.0M	% increases in GeoChip 5.0M compared with GeoChip 4
Manufacturers	NimbleGen	Agilent		-
Feature shape	Square	Circular		-
Feature size	13 μm \times 13 μm	30 micron (diameter)		-
Maximum features per array	135,000	60,000	180,000	-
No. of arrays per slide	12	8	4	-
No. of genes	410	373	1,447	+253%
No. of probes	82,074	57,498	161,961	+97%
No. of sequence-specific probes	18,098	11,819	30,640	+69%
No. of group-specific probes	63,976	45,679	131,321	+105%
No. of covered CDS	141,995	151,797	365,651	+158%
No. of covered strains	5,247	6,507	8,987	+71%
16S positive controls	640	1,536	5,282	+540%
Controls from thermophiles	1,689	1,130	3,378	+100%
Universal standards	6,000	480	1,360	-78.6%
Agilent negative controls	-	1,565	3,390	-

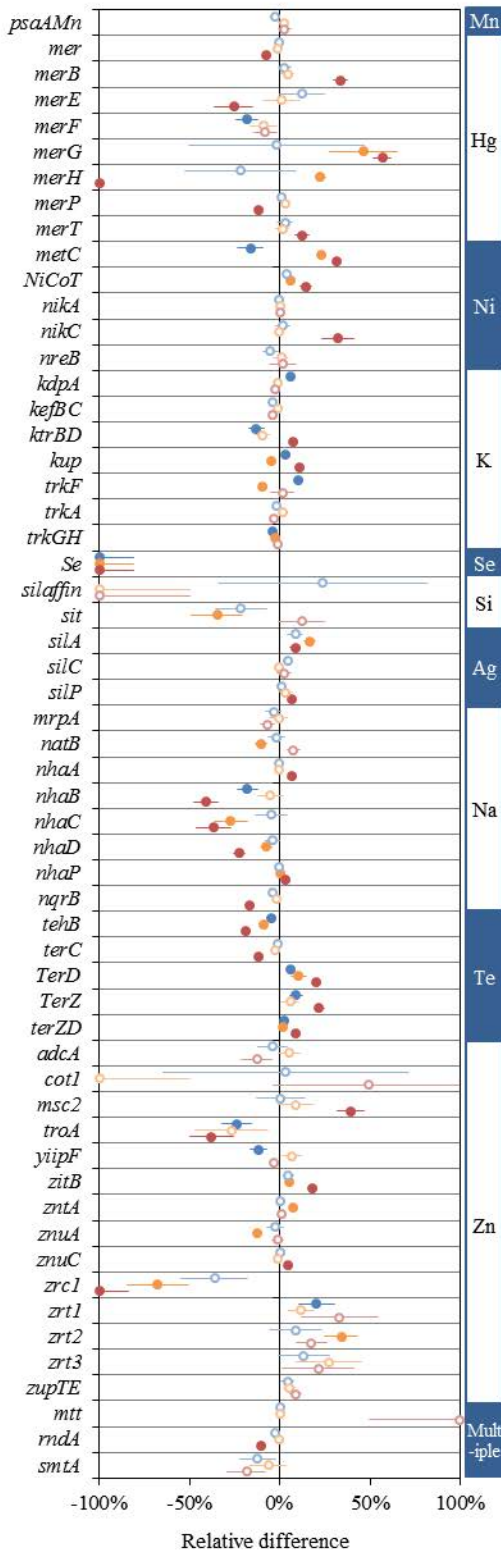
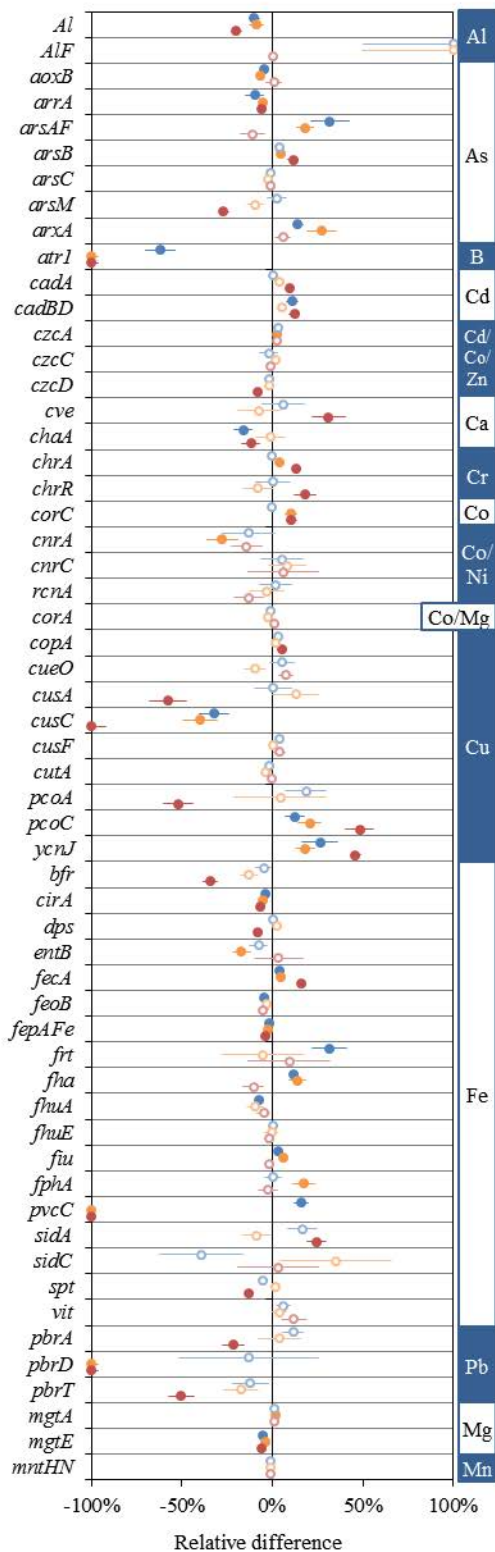
Table S4. Environmental variable measurements from groundwater samples used in the application study. A total of 12 samples were collected from wells representing 4 levels of contamination (L0, L1, L2 and L3). Variables were divided into 5 categories: general environmental parameters (Env. Parameters), gas TCD, dissolved Carbon (C), anion and metal ion.

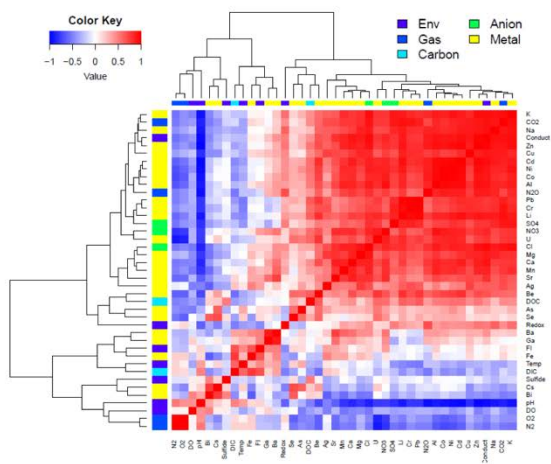
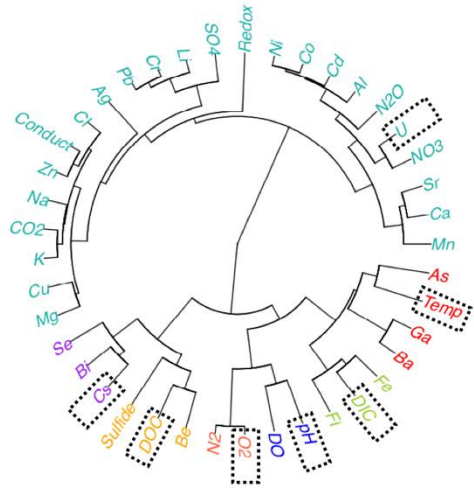
		L0			L1			L2			L3		
		GW-654	GW-199	GW-350	FW-300	FW-301	FW-303	FW-215	DP16D	GW-101	FW-106	FW-126	FW-021
General Env. Parameters	Temp. (°C)	13.01	14.29	17.98	15.54	15.77	15.82	17.88	17.05	18.12	14.84	12.16	16.41
	D.O. (mg/L)	2.35	0.5	0.02	0.28	0.78	0.71	0.13	0.26	0.96	0.18	0.16	0.27
	Conduct. (µS/cm)	269	581.6	545	378.9	334.4	316.4	637.4	661.0	1721.0	7864.0	18770	7967
	Redox (mV)	175	305	137	-129	39	147	43	-50	-72	426	168	387
	pH	7.19	6.53	6.67	6.59	6.68	7.16	6.6	6.67	6.81	3.55	3.04	3.43
	Sulfide (mg/L)	0.003	0.003	0	0.026	0.188	0.041	0	0	0	0.004	0.044	0.004
	F.I. (mg/L)	0.1	0.13	0.38	0.5	0.98	0	0.21	1.46	2.7	1.02	0.13	0.03
Gas TCD (µmol/mL)	N ₂	59.313	104.347	125.995	57.392*	45.931	68.853	31.878	30.766	42.717	28.132	14.475	29.89
	O ₂	12.429	23.330	24.629	11.256*	9.934	12.579	6.006	6.082	8.156	8.331	4.423	6.193
	CO ₂	36.69	172	127.76	54.3*	60.34	48.26	173.47	201.94	283.42	631.07	718.48	288.99
	N ₂ O	0	0	0	0	0	0	0	0	2.652	22.639	17.022	30.117
C (mg/L)	D.I.C	30.98	85.05	67.46	48.13	55.44	40.82	85.25	87.91	115.9	43.18	36.65	22.27
	D.O.C.	0.345	1.335	0.717	44.54	48.65	39.59	1.928	2.326	4.065	47.87	128.2	7.298
Anion (mg/L)	Cl	1.259	5.492	13.527	2.186	3.534	3.328	15.592	23.712	42.244	318.17	373.69	152.313
	NO ₃	0.460	0.169	0.315*	3.66	36.357	3.962	5.503	141.022	1470.9	2692.04	11648.3	4506.9
	SO ₄	15.642	20.966	13.388	6.472	9.139	7.397	75.914	64.844	7.514	2062.69	1460.19	41.945
Metal (mg/L)	Ag	0.007	0.010	0.021	0.014	0.008	0.008	0.011	0.011	0.022	0.022	0.023	0.011
	Al	0.015	0.013	0.038	0.03	0.418	0.017	0.013	3.444	1.129	108.497	558.765	114.501
	As	0.003	0.002	0.005	0.021	0.003	0.003	0.011	0.011	0.005	0.008	0.006	0.011
	Ba	0.04	0.242	0.055	0.093	0.077	0.072	0.104	0.269	2.820	0.077	0.131	2.073
	Be	0.04	0.019	0.038	0.082	0.04	0.04	0.041	0.041	0.039	0.059	0.149	0.079
	Bi	0.038	0.001	0.005	0.018	0.038	0.038	0.009	0.009	0.005	0.005	0.005	0.009
	Ca	16.619	89.37	92.374	67.26	78.81	52.516	120.122	139.769	419.983	273.264	9837.8	3970.12
	Cd	0.003	0.002	0.004	0.006	0.006	0.003	0.003	0.01	0.007	0.132	0.866	0.173
	Co	0.004	0.000	0.002	0.006	0.011	0.004	0.003	0.056	0.009	0.509	1.364	1.225
	Cr	0.004	0.004	0.007	0.010	0.004	0.004	0.005	0.005	0.007	0.384	0.798	0.005
	Cs	0.029	0.008	0.016	0.064	0.029	0.029	0.032	0.032	0.016	0.017	0.016	0.032
	Cu	0.009	0.013	0.025	0.008	0.009	0.009	0.004	0.004	0.222	0.810	1.587	0.118
	Fe	0.011	0.851	0.338	0.067	0.030	0.011	0.016	1.933	4.585	0.038	0.167	0.016
	Ga	0.011	0.010	0.007	0.014	0.011	0.011	0.007	0.011	0.089	0.007	0.009	0.06
	K	0.996	3.291	3.435	1.546	2.372	1.456	5.166	5.962	4.857	216.441	102.491	28.827
	Li	0.038	0.022	0.042	0.096	0.038	0.038	0.048	0.048	0.069	1.946	5.190	0.227
	Mg	16.175	22.308	80.874	16.175	16.175	16.175	32.349	32.349	80.874	80.874	216.229	117.901
	Mn	0.011	4.318	0.291	0.199	2.249	0.043	0.475	9.393	8.348	32.498	134.068	128.452
	Na	39.949	20.806	52.015	10.403	10.403	10.403	20.806	29.262	52.015	865.339	826.325	269.202
	Ni	0.006	0	0.011	0.043	0.048	0.006	0.021	0.242	0.026	7.184	15.352	5.339
	Pb	0.003	0.002	0.003	0.006	0.003	0.003	0.003	0.003	0.003	0.032	0.060	0.004
Se	0.009	0	0.004	0.018	0.009	0.009	0.009	0.009	0.009	0.024	0.005	0.013	
Sr	0.096	0.212	0.119	0.115	0.12	0.168	0.399	0.372	2.219	0.369	2.43	1.373	
U	0.051	0.003	0.006	0.216	0.16	0.081	1.452	0.744	0.417	16.625	55.286	3.751	
Zn	0.02	0.059	0.084	0.051	0.058	0.041	0.040	0.078	0.093	1.099	2.189	0.897	

* missing values were imputed using the mean of values from other two replicates in the same group.

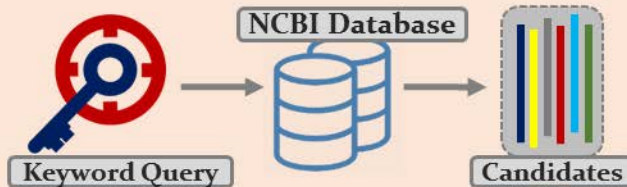




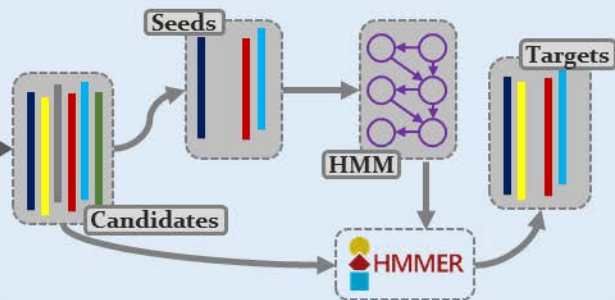


a.**b.**

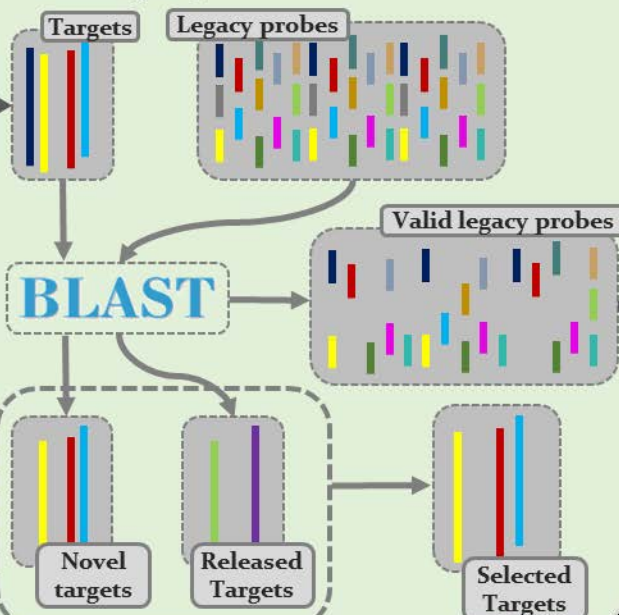
a. Candidate Sequence Retrieval



b. Candidate Sequence Confirmation



c. Target screening and Legacy Probe Validation



d. Probe Design and Selection

