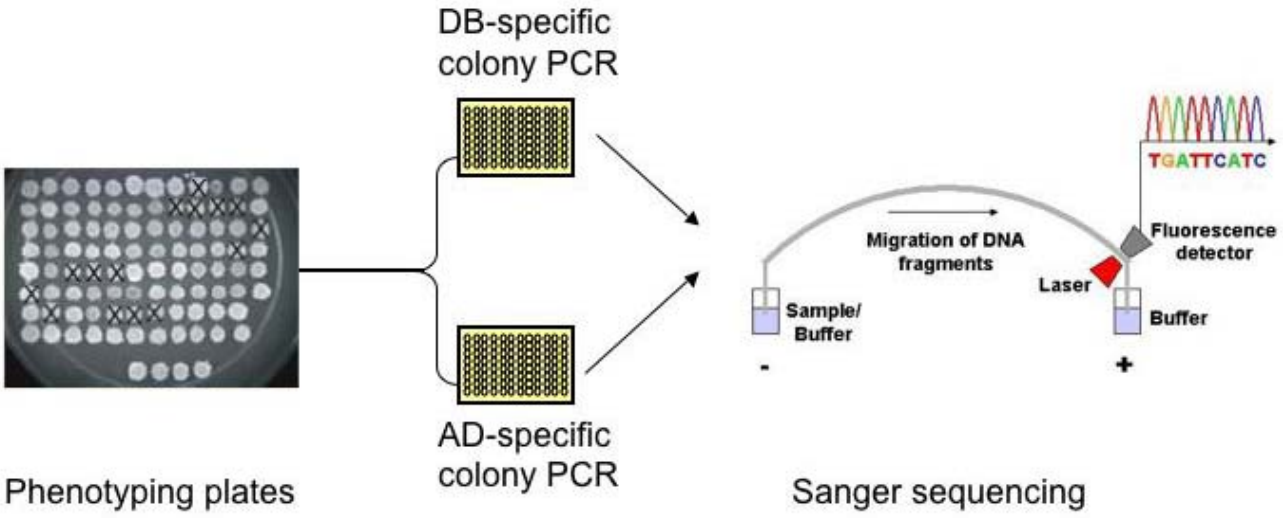
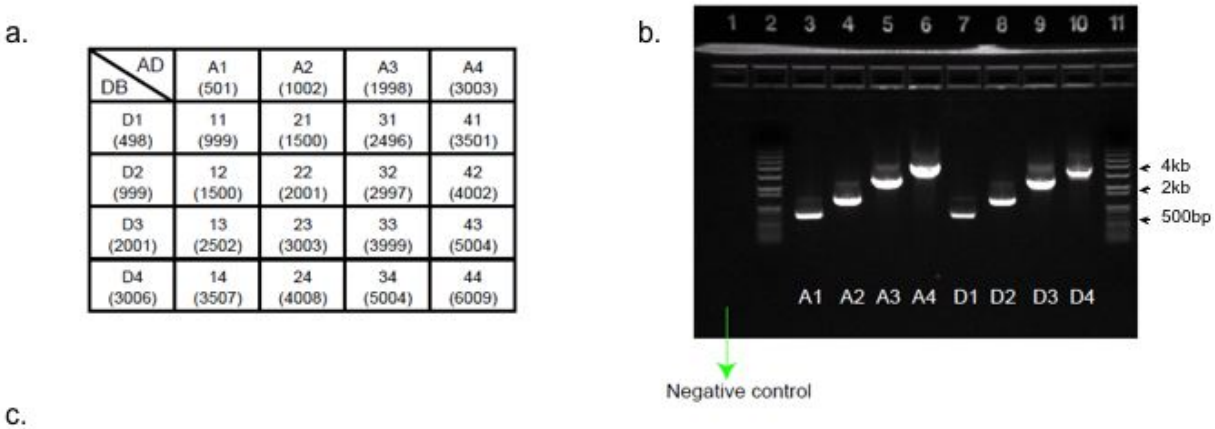
SUPPLEMENTARY FIGURE AND TABLE LEGENDS



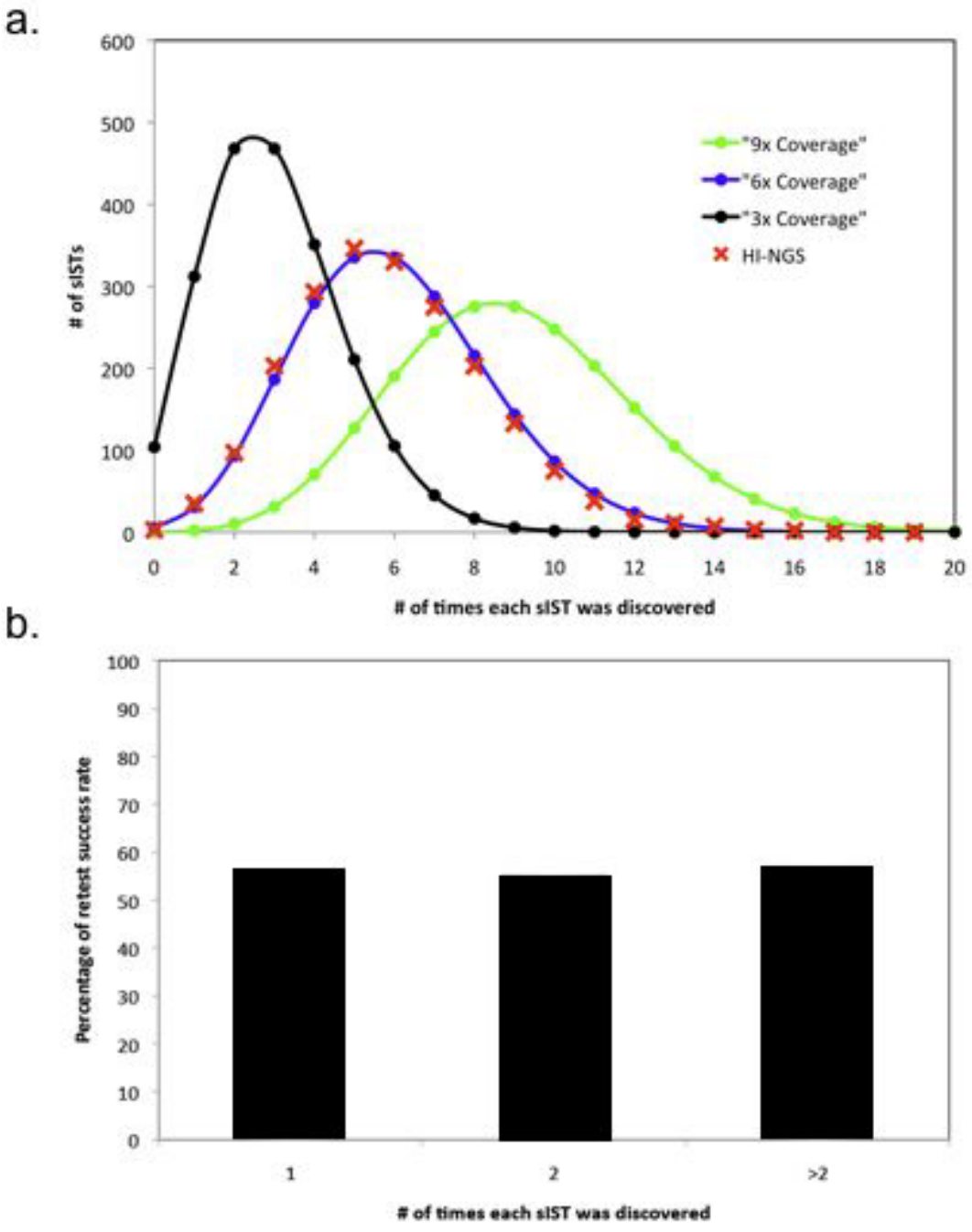
Supplementary Figure 1. The conventional approach for sequencing large numbers of ORF pairs resulting from high-throughput Y2H screens. Within the cells of each colony, there is a pair of DB-X and AD-Y ORFs that encode putatively-interacting proteins. X and Y ORFs are PCR amplified individually using DB- and AD-vector-specific primers, respectively, while keeping the association between them. Each ORF is then sequenced by Sanger sequencing.

a.

AD / DB	A1 (501)	A2 (1002)	A3 (1998)	A4 (3003)
D1 (498)	11 (999)	21 (1500)	31 (2496)	41 (3501)
D2 (999)	12 (1500)	22 (2001)	32 (2997)	42 (4002)
D3 (2001)	13 (2502)	23 (3003)	33 (3999)	43 (5004)
D4 (3006)	14 (3507)	24 (4008)	34 (5004)	44 (6009)



Supplementary Figure 2. Robust stitching PCR amplifications over a wide ORF size range. (**a**) Lengths of different ORFs selected for testing PCR stitching and expected lengths of the stitched PCR products. (**b**) PCR products of the selected ORFs from the first round PCR. (**c**) Conjugated PCR products from the second round PCR. The sizes of all PCR products are as expected.



Supplementary Figure 3. Statistical analysis of the distribution of useful 454 reads for sISTs. (**a**) Poisson distributions of the number of reads for each sIST at different sequencing coverage. (**b**) Retest success rates for sISTs covered with one, two, or many (>2) sISTs are indistinguishable.

A. Traditional Sanger sequencing

Failure rate of a single PCR ≈ 5%

Failure rate of a single sequence run ≈ 5%

Failure rate of two PCRs and two sequence runs

= 1 - (1 - 5%) ^ 4

= 19%

The overall success rate should be about 81%

B. 454 FLX sequencing

Failure rate of a single PCR ≈ 5%

Failure rate of three PCRs = 1 - (1 - 5%) ^ 3

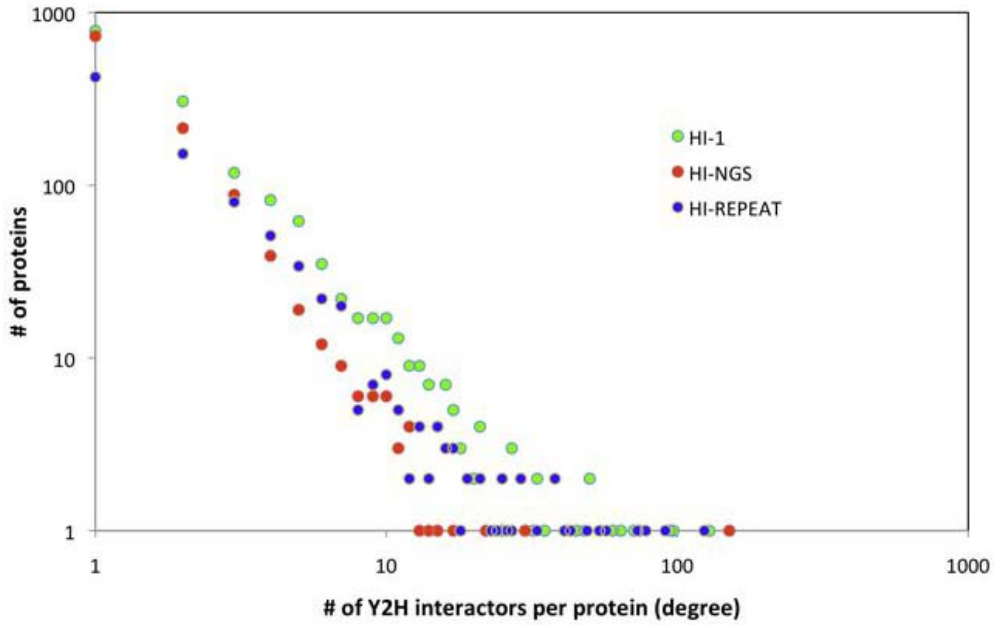
= 14%

The overall success rate should be about 86%

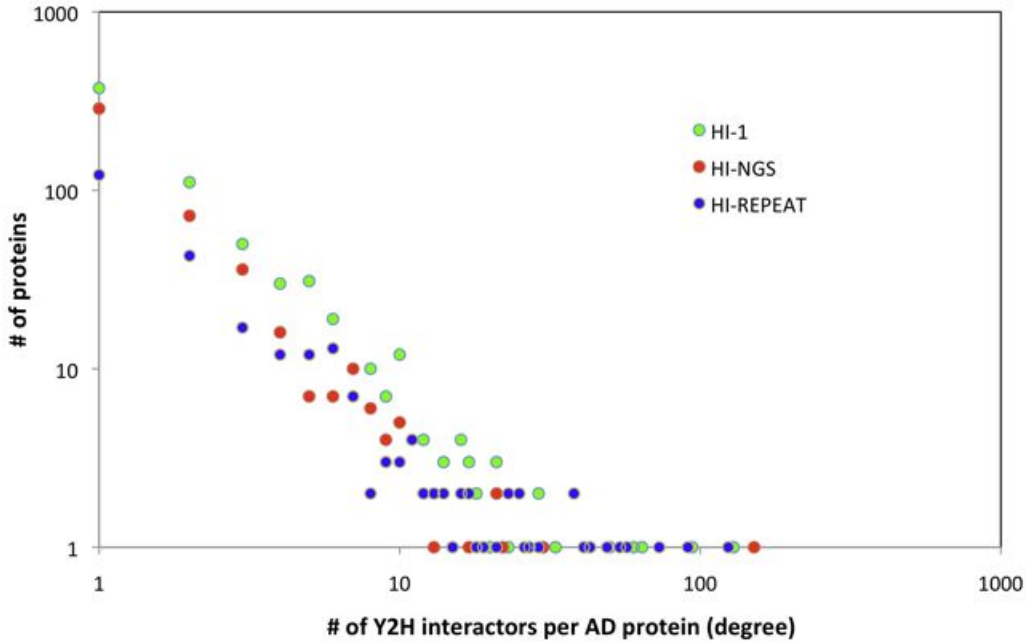
Therefore, the overlap should be about 81% x 86% = 70%

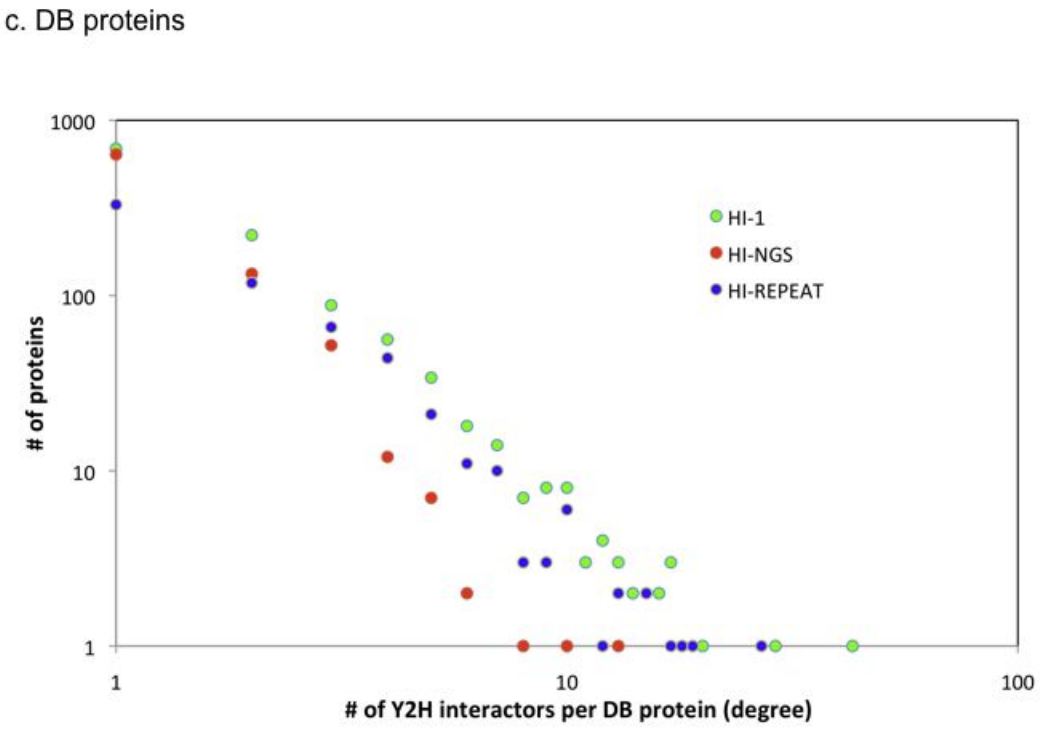
Supplementary Figure 4. Statistical calculation of the overall success rate for PCR stitching with three PCR reactions for each pair of ORFs that encode interacting proteins.

a. Total proteins



b. AD proteins





Supplementary Figure 5. Comparison of degree distributions of HI-NGS, HI1, and HI-REPEAT (unpublished data) datasets. (**a**) Degree distribution for all proteins. (**b**) Degree distribution for all AD proteins. (**c**) Degree distribution for all DB proteins. The scale-free degree distributions show that all three networks contain hubs. For instance, gene 11007 (CCDC85B) has 129 interactions in HI1, and gene 10488 (CREB3) has 124 interactions in HI-REPEAT. The degree distributions of HI-NGS (especially for total proteins and DBs) are the least extreme among the three networks.

a. Traditional PCR and Sanger sequencing

	Unit Price	Subtotal
2 PCR reactions	$1.00	$2.00
2 Sanger sequencing reads	$3.00	$6.00
	Total	$8.00

b. PCR stitching and 454 sequencing

	Unit Price	Subtotal
3 PCR reactions	$1.00	$3.00
$8K for 5.2K pairs	$1.50	$1.50
	Total	$4.50

At current cost our new approach on average saves ~40% for each pair of interactions.

Supplementary Figure 6. Calculation and comparison of costs for each pair of ORFs that encode interacting proteins. (**a**) Traditional PCR and Sanger sequencing. (**b**) PCR stitching and 454 FLX sequencing.

Supplementary Table 1. PCR reaction mix setup and PCR conditions of the three PCR reactions for each pair of ORFs that encode interacting proteins. (**a**) Setup and conditions for first round DB- and AD-specific PCR reactions. (**b**) Setup and conditions for second round PCR reaction.

a. First round of PCR stitching

PCR reaction #1:

Reagent	Volume (ul)
10x Buffer	3
10 mM dNTP	0.6
50 mM $MgSO_4$	1.2
10 uM AD-specific primer	0.6
10 uM B2-linker-forward primer	0.6
Invitrogen HiFi Platinum ™ Taq polymerase	0.12
Yeast cell lysate	3
ddH_2O	20.88

PCR reaction #2:

Reagent	Volume (ul)
10x Buffer	3
10 mM dNTP	0.6
50 mM $MgSO_4$	1.2
10 uM DB-specific primer	0.6
10 uM B2-linker-reverse primer	0.6
Invitrogen HiFi Platinum ™ Taq polymerase	0.12
Yeast cell lysate	3
ddH_2O	20.88

PCR conditions:

Temperature (°C)	Time (minute)	Cycle
94	2:00	1
94	0:45	35
58	0:30	
68	5:00	
68	5:00	1
10	!	1

b. Second round of PCR stitching

PCR reaction:

Reagent	Volume (ul)
10x Buffer	3
10 mM dNTP	0.6
50 mM $MgSO_4$	1.2
10 uM AD-specific primer	0.6
10 uM DB-specific primer	0.6
Invitrogen HiFi Platinum ™ Taq polymerase	0.12
Reaction #1	0.5
Reaction #2	0.5
ddH_2O	22.88

Primer sequences (5'→3'):

AD:

CGCGTTTGGAATCACTACAGGG

DB:

GGCTTCAGTGGAGACTGATATGCCTC

B2-linker-forward:

GGATACCGCCGAGCTGAGAGCCATCAAACCACTTTGTACAAGAAAGTTGGG

B2-linker-reverse:

CTCTCAGCTCGGCGGTATCCCCATCAAACCACTTTGTACAAGAAAGTTGGG

PCR conditions:

Temperature (°C)	Time (minute)	Cycle
94	2:00	1
94	0:45	35
58	0:30	
68	6:00	
68	5:00	1
10	!	1

Supplementary Table 2. 1166 interactions in HI-NGS. Entrez Gene IDs are used as gene identifiers.

Interactions			Interactions			Interactions			Interactions	
Protein A	Protein B		Protein A	Protein B		Protein A	Protein B		Protein A	Protein B
10016	553115		11043	80705		22929	7597		5082	5705
10024	22924		11043	80726		22929	91614		50848	6449
10048	55770		11043	81853		22932	9111		50855	5584
10048	595		11043	9747		22934	26003		5092	5092
10062	6258		11047	51377		22934	5781		5092	9513
10069	10320		1106	90933		22954	7356		51043	92856
10072	9817		11060	11100		22980	6303		51072	6867
10078	3839		11060	4088		23001	4188		51160	6293
10100	10488		11060	7456		23001	56658		51160	6498
10105	54971		11060	9260		23028	56658		51160	6606
10106	83461		11060	92714		23028	9111		51160	79720
10112	80256		11060	9618		23028	9189		51225	7454
10126	9457		11060	9871		23051	51701		51225	9818
10133	10318		11063	154313		23132	29883		51324	57085
10133	222484		11094	57146		23132	8887		51421	54927
10133	51019		11100	63948		23201	26039		51421	5892
10133	8697		11117	66036		23201	54461		51435	9144
10133	9413		11199	80153		23201	78990		5152	7023
10136	10488		11201	56658		23201	84959		51603	55145
10138	6015		1123	8440		23204	55204		51608	5565
10142	2239		11235	8428		23204	6642		51647	9391
10142	84970		11266	28227		23204	80221		51765	5987
10152	1838		11267	84313		23208	6449		51765	79959
10152	339834		11274	28978		23214	23636		5252	7024
10159	10488		11275	51060		23219	2801		5276	54557
10174	5062		11275	59349		23281	23281		5276	7917
1019	896		11277	51142		23281	293		5300	81628
10190	4105		112869	3866		23281	51164		53347	9802
10193	5984		11338	6626		23281	5682		5349	90993
10195	10488		11338	7307		23281	57088		5354	5798
10197	55840		11339	54069		23281	57474		53615	57459
10197	64798		11340	56915		23281	655		539	9240
10197	85403		11340	6628		23281	6687		54033	9802
1021	896		11345	8878		23281	7288		54107	54108
10212	84324		11043	7572		23281	7695		5413	79666
10228	63908		113675	113675		23281	79230		5432	54550
10228	6810		113878	121274		23281	79696		5441	8161
10241	11266		113878	23413		23281	83541		54458	84708
10241	153657		113878	26984		23281	84775		54461	6285

10241	167153		113878	51324		23281	85865		54472	54959
22929	55038		113878	54784		23281	8851		54550	79635
10241	3142		113878	56478		23281	90594		54550	83878
10241	51564		113878	5859		23281	9873		54550	84708
10241	5204		113878	7321		23283	2801		54557	5479
10241	55138		113878	7920		23321	80131		54623	9513
10241	553115		113878	83988		23338	84289		54793	54793
10241	8115		113878	85376		23387	7205		54797	56658
10241	84904		114822	80321		23392	57085		54927	64396
10241	9413		114904	6449		23424	4107		54954	7024
10241	9610		114928	154313		506	91647		54955	57541
10245	10488		114932	120379		23468	54904		54955	5987
10252	3198		114932	375341		23468	6672		54955	9189
10253	5013		114932	5296		23479	54550		54971	6668
10253	53347		114932	55234		23558	4103		54971	7148
10253	80270		114932	5682		23582	29088		54971	92399
10253	81853		114932	705		23582	496		54974	54974
10254	9413		114932	8099		23582	56906		5500	84988
1026	595		114971	4188		23582	5699		55012	80125
10262	2444		114971	8601		23582	57474		55034	55742
10262	9747		114984	51608		23582	6158		55037	79444
10289	1856		115106	23636		23582	7695		55086	79666
10301	57120		115106	51019		23582	84619		55093	55145
10302	9111		1153	3190		23582	90594		55138	9618
10318	29922		1158	140462		23582	9063		55145	55145
10318	29965		116113	9502		23635	5300		55145	7726
10318	55388		116173	148103		23636	284058		55165	7004
10318	6919		116173	2571		23636	51019		55211	6455
10318	7301		116173	3190		23636	51668		55211	79666
10318	79635		116173	51228		23636	53371		55216	6285
10318	81926		116173	5780		23636	55145		55216	7681
10318	900		116173	6642		23636	79635		55284	9051
1032	22806		116173	80221		23636	8473		5530	5534
1032	84970		116173	89882		23636	9818		553115	63948
10320	1488		116225	4665		23647	57085		55324	5987
10320	1641		116225	7126		23650	2801		55536	84619
10320	26240		116225	84445		23710	79036		5563	84456
10320	5892		1164	8548		24144	284058		55646	8575
10320	64396		116442	2801		24144	3603		5565	7186
10320	744		117144	8601		24144	54474		5565	80125
10320	80256		117177	81567		24144	64745		55663	7718
10342	9500		117177	8985		24144	84988		55663	84307
10362	7185		117178	23503		24144	9753		5571	8666
10385	10488		117178	58491		2444	663		55734	64782
10422	1856		117178	60491		245973	2801		55734	90933
10422	23321		117178	84619		246184	8881		55739	6449

10425	80125		117581	6938		246329	79169		55758	7186
10431	10488		118471	29979		246329	84080		55773	6794
10435	23092		118788	8440		246329	85403		5585	9454
10443	5478		1196	1198		2495	2512		55888	9753
10447	10488		1196	6733		253980	7126		55906	58500
10458	148223		1196	79753		254122	6642		55916	56000
10482	23636		120534	79666		2553	4004		5598	63940
10483	133619		121268	2275		2553	7186		56658	56658
10483	3200		121457	7170		256302	54507		56658	5715
10483	9871		121536	55145		2577	64395		56658	7321
10488	10488		122786	6938		2577	64396		56658	7325
10488	10490		124220	2801		25788	7341		56658	9129
10488	1066		124404	1731		25816	79666		5682	60490
10488	10668		124790	124790		25900	84260		5684	5684
10488	10695		124790	58500		25959	2801		5684	5687
10488	10948		124790	63978		25988	339834		56935	665
10488	10990		124790	6938		25988	56658		56952	81853
9513	9883		126070	23582		25988	7157		5701	5711
10488	11070		126070	4188		26003	5998		5705	7138
10488	11094		126070	7185		26003	64518		5706	5715
10488	11161		126308	7050		26003	79656		57085	5780
10488	113178		127557	6133		26003	91703		57085	7329
10488	113452		127557	9129		26036	55007		57085	8050
10488	114569		127557	92822		26036	55038		57085	81853
10488	116173		127703	23281		26036	7329		57085	84191
10488	1186		127703	7170		26039	6628		57088	81926
10488	125875		128977	79959		26151	9521		57088	8862
10488	133022		129807	4188		26228	30011		57088	9404
10488	145957		129807	7205		26234	6500		5709	81853
10488	1521		130502	29781		26273	6500		57326	64395
10488	1534		130502	4656		26353	9531		57456	79959
10488	192683		130540	6820		26508	4188		57506	7186
10488	200185		1329	339834		266740	9702		57559	5862
10488	2030		1329	373		26751	79720		57639	7185
10488	2224		1329	54906		26994	51619		57693	5940
10488	22845		134288	6449		27111	441519		57727	57727
10488	22908		139596	79912		27111	441521		57820	9852
10488	23471		14	55216		27111	51538		5805	5805
10488	23480		140688	1487		27111	55660		58155	9782
10488	25972		140691	4103		27111	84285		58500	7185
10488	27040		140691	516		27232	51564		58500	79666
10488	2706		140691	8434		27238	6455		58516	79959
10488	28978		140735	8631		27246	9802		58516	9111
10488	29058		1411	1414		27258	57819		5877	83593
10488	2993		142679	211		27300	5877		5886	7205
10488	3162		145282	4105		27316	3190		5914	6258

10488	3177		1457	54778		27335	51386		59286	5987
10488	336		145946	2275		27430	4144		5929	9070
10488	3386		145957	89885		2773	63940		5987	60491
10488	340481		1468	4188		2801	284058		5987	7286
10488	347		147687	22981		2801	285172		5987	79025
10488	358		147687	23582		2801	29087		5987	79696
10488	362		147687	24144		2801	3707		5987	81926
10488	3732		147687	5987		2801	51164		5987	8697
10488	375035		147687	80817		2801	58155		5987	8930
10488	3775		147700	147700		2801	6941		5987	9929
10488	394261		147700	285753		2801	79169		5990	84970
10488	441521		148581	2801		2801	8065		60	60
10488	4712		1488	7050		2801	81629		60	71
10488	4818		150483	22806		2801	81926		6015	7703
10488	50859		150771	6449		2801	84460		60490	60490
10488	51006		151194	63920		2801	84934		60490	7775
10488	51024		151871	2257		2801	84991		60491	9240
10488	51090		151871	84285		2801	8796		6271	6285
10488	51107		1522	3846		2801	902		6275	6285
10488	51128		152926	329		2801	9690		6285	84277
10488	51465		153657	4253		2801	9814		6285	84830
10488	51522		153743	5501		2815	3198		6285	92856
10488	51540		154313	27033		28227	81853		6293	79874
10488	51703		154313	84970		283385	4105		6293	9589
10488	51715		155465	23636		284001	9814		6303	84991
10488	527		155465	29979		284058	29911		6342	9319
10488	5349		159163	3190		284058	4188		638	89885
10488	53826		1602	55577		284252	7329		63948	84970
10488	5407		1602	79169		285622	64098		63978	79666
10488	54708		163081	51776		285753	51160		63978	91647
10488	54946		163183	64581		285753	55165		64210	9391
10488	54947		1641	2801		285753	84619		64395	64396
10488	55092		164592	23281		285782	8500		6442	6449
10488	55266		164592	27018		28956	8649		6449	6782
10488	55273		164592	7185		28985	8562		6449	79089
10488	55281		164592	80817		29124	9586		64581	665
10488	55332		167826	7185		2947	2947		64744	9802
10488	55716		170082	79596		2949	2949		64745	7205
10488	55850		170463	5300		29761	6850		64778	9319
10488	55851		170954	5987		29781	7186		64786	80254
10488	55973		1808	1808		29781	83538		64798	8945
10488	55974		1838	80254		29785	54507		64798	90933
10488	56063		1856	23623		29883	51435		64802	64802
10488	56475		1856	3104		29907	57142		64853	79666
10488	56894		1856	407		29907	92840		660	6774
10488	57515		1856	51720		29911	55554		6602	79720

10488	5780		1856	55110		29911	6839		66036	9107
10488	58515		1856	5562		29911	7170		6605	80254
10488	598		1856	64396		29911	84934		6606	6606
10488	6009		1856	7170		29911	92999		6606	9360
10488	60343		1856	7454		29922	8061		664	665
10488	6189		1856	7525		29922	84708		6642	92840
10488	6342		1856	79159		29959	81628		665	665
10488	638		1856	84619		29979	3537		6652	6652
10488	6472		1856	9531		29979	5276		6687	9240
10488	64755		1876	7029		29979	55742		6774	6850
10488	6515		191	79078		29979	60626		6810	79901
10488	664		1912	4004		29979	7857		6810	8417
10488	6890		1912	4005		29979	9240		6919	79036
10488	7102		1912	4733		29994	373		6921	6923
10488	7104		1912	6477		3091	6759		6926	79666
10488	7106		1912	705		9500	9802		7004	9659
10488	7780		1912	84078		5050	5050		705	7205
10488	78988		1912	84080		3104	90933		705	9618
10488	78992		1912	84456		3183	3183		7067	79959
10488	79001		1933	1937		3183	55285		71	71
10488	79152		1936	80325		3185	80125		7138	7168
10488	79157		1937	23276		3198	3198		7170	8379
10488	79161		196513	80153		3219	84445		7185	7189
10488	7920		1974	2801		329	7286		7185	81926
10488	79415		201181	55145		329	80705		7185	85302
10488	79669		201181	5684		329	81926		7185	90874
10488	79844		201255	79036		3329	84445		7186	7189
10488	79901		201895	89885		339047	79876		7186	79025
10488	80221		202559	202559		339834	84619		7186	91544
10488	80704		202559	7297		3434	3437		7186	92610
10488	80777		202559	81853		3437	7170		7188	7189
10488	81539		203068	222484		348487	51440		7188	9111
10488	81853		203068	266740		3603	7205		7189	79666
10488	81855		203068	29781		3608	5987		7205	7414
10488	8260		203228	8892		3611	55742		7205	9747
10488	83460		2071	2801		3612	3612		7286	7791
10488	84065		2079	54906		3646	93487		7314	9802
10488	84102		211	26073		373	389741		7325	9666
10488	8417		211	57634		373	5144		7329	84528
10488	84179		211	9168		373	7329		7329	9111
10488	84277		2130	311		374355	79666		7329	9319
10488	84329		2139	6495		3796	55132		7391	79666
10488	84620		2170	23636		3801	54955		7414	9404
10488	84830		220002	26		3801	5563		744	9319
10488	84975		220082	4292		3801	6605		7454	9322
10488	8611		220108	222389		3801	84619		7456	9454

10488	8760		220108	80125		3846	54778		7566	892
10488	8974		2202	6449		3846	59277		7572	90933
10488	91252		2203	2203		3846	6687		7681	84080
10488	9145		220766	220766		3846	7301		7791	84816
10488	91663		220766	90060		3846	81926		7812	83638
10488	91782		221184	5987		3846	84221		78992	89885
10488	931		222389	4675		3846	9586		79078	9852
10488	94101		222389	55646		3866	9894		79088	80705
10488	94107		222389	55840		387	9750		79095	90423
10488	9445		222389	80022		3887	5013		79098	9456
10488	9482		222389	84285		389741	4188		79173	7965
10488	9524		222389	85403		389741	84445		7918	84445
10488	9554		222484	23089		391257	9063		79230	80321
10488	9783		222484	253827		392	57085		79415	90993
10488	9791		222484	29116		3927	9319		79571	8405
10490	9482		222484	3603		3937	9402		79571	84970
10494	11235		222484	506		3937	9513		79571	92822
10513	147746		222484	51068		3939	3945		79571	9702
10513	84795		222484	51447		4001	9063		79666	8545
10534	902		222484	51540		4004	6660		79760	79833
10538	3726		222484	55037		4005	4670		79763	810
10539	2896		222484	5601		401152	801		7984	80258
10539	57019		222484	5607		402	5147		79959	9465
10539	7571		222484	7186		404550	5530		80125	85403
10541	3841		222484	7448		410	9319		80125	8576
10541	60343		222484	79016		4105	7726		80198	84464
10586	9240		222484	80007		4105	91647		80256	9457
10589	4005		222484	80131		4110	55621		80256	9618
10589	6880		222484	84991		414301	56477		80321	84775
10589	79666		222484	85007		4188	50717		8065	9616
10597	9189		222484	8503		4188	527		80817	8434
10615	201181		222484	9295		4188	57474		81853	83988
10615	23503		222546	6628		4188	60370		81926	8601
10615	84080		2239	81926		4188	7050		81926	9240
10615	84619		2239	892		4188	7695		8266	9618
10640	26098		2274	3055		4188	7766		83755	93185
10640	30000		2274	391257		4188	81926		84080	91544
10687	147687		2274	56243		4188	83538		84108	8892
10687	84080		2274	5991		4188	84527		84108	9240
10687	9025		2274	7597		4188	90594		84283	84283
1072	1073		2274	7690		4188	9139		84283	9025
1073	60		2275	23281		4188	92906		84305	84816
10758	2801		2275	26505		4281	4281		8440	868
10758	7189		2275	4793		4281	7325		84445	84988
10797	55577		2275	51298		4292	55086		84456	84456
10801	23157		2275	5552		4342	9454		84528	9823

10807	8907		2275	7297		439	51608		84661	84726
10865	6660		2275	7965		441521	6285		84661	9070
10900	8907		2275	81853		4615	4615		84959	8502
10919	81926		22806	23641		4618	6938		84975	90993
10928	2874		22806	256302		4665	4739		85013	90993
10928	9702		22806	51765		468	8945		85403	9454
10933	114932		22806	55110		4733	55854		8557	89884
10946	140691		22806	5565		4739	54971		862	9404
10946	6871		22806	55758		4739	5987		862	9454
10966	2275		22806	7329		4739	63948		8636	8636
10979	116173		22806	7341		4739	7024		875	875
11016	53615		22806	91647		4783	51421		8796	9111
11030	145946		22806	9747		4800	84872		8834	89885
11030	146664		22827	22827		4800	9747		8834	90993
11030	166863		22827	29115		4801	56655		8939	9444
11030	23132		22827	5910		4841	4841		8976	9025
11030	2483		22827	6477		4899	57125		89885	9144
11030	57459		22827	80306		4931	7329		90423	9550
11043	1329		22893	84445		4950	89885		9111	9111
11043	1641		22920	4688		5013	54507		91544	9470
11043	285753		22924	4005		5013	84661		93621	9643
11043	342132		22924	5682		5013	8601		95	95
11043	51619		22929	22929						

Supplementary Note 1. Distribution of 454 reads for unique sISTs.

There is a key difference between using next-generation sequencing for determining DNA sequences and using it for identifying sISTs. Determining genomic DNA sequences with next-generation sequencing requires high coverage because the goal is to determine every nucleotide correctly; identifying sISTs does not require high coverage (a few mismatches per read is not problematic). One read of 15 bases on both sides (our definition of usable reads) is enough to identify the sIST from our search space of 6,000 genes. The total number of nucleotides for our search space is ~9 million (6,000 genes times average length of 1,500 bp). As $4^{15} > 1$ billion, the chance of having two 15 mers with the same sequence in our search space at random is less than 1 in 100. In our sequencing results, on average we have at least 50 bases on each side of the usable reads. This is basically how we did Sanger sequencing in the past: only the first 20-50 bases from a Sanger sequencing read are helpful and relevant. Each IST is only sequenced once by Sanger sequencing, whereas shotgun genomic sequencing with Sanger technology requires much higher coverage.

We have 35 sISTs identified once and 97 sISTs identified twice (**Supplementary Fig. 3**). The retest success rates of these sISTs are identical to those identified more often (**Supplementary Fig. 3**). The distribution of number of times sISTs identified fits a Poisson distribution. Based on the Poisson distribution, the fraction of missed sISTs is less than 0.3%. This is why we placed the failure rate for 454 sequencing at 0% (**Supplementary Fig. 4**). If we double the number of sISTs for one 454 run (in other words, decrease the coverage by half), the fraction of missed sISTs is still expected to be less than 5%, which is about the same failure rate as Sanger sequencing.

Supplementary Note 2. Data quality assessment of HI-NGS.

The PRS set was initially chosen from the HI1 search space because at the time only clones from that search space were available. Here we only screened 2,275 genes in the HI1 search space (**Fig. 2a**), therefore, only 21 PRS interactions were screened in HI-NGS search space, 6 of which are found positive in our screen (29% success rate, which is comparable to the Braun *et al.*[1] success rate of 25%). Furthermore, 8 RRS pairs (negative controls) were also screened in HI-NGS space and none of them were found positive.

Supplementary Note 3. Comprehensive comparison of HI1 and HI-NGS.

Overlaps between HI1 and HI-NGS. The screen to produce HI-NGS used a newly-developed HT-Y2H pipeline[2] with yeast strains *MAT***a** Y8800 and *MAT*α Y8930 containing *HIS3* and *ADE2* reporter genes[3]. In contrast, HI1 was produced with yeast strains *MAT***a** MaV103 and *MAT*α MaV203 containing a different set of reporter genes (*HIS3*, *URA3*, and *lacZ*)[4]. The common search space between HI-NGS and HI1 is 2,275 DB-Genes × 2,275 AD-Genes (~ 2.6×10^6 protein pairs). In this common space, HI1 has 689 interactions, 127 of which are recovered in HI-NGS. This overlap is what is expected given the completeness, assay- and sampling sensitivity of the two experiments[5].

The search space of HI-NGS was screened once and not driven to saturation. We have previously quantified the "sampling sensitivity" of our high-throughput screens to ~35-50% of all Y2H-detectable interactions[5]. The assay sensitivities of the two different Y2H implementations are quite different and overlap by about 40%. By combining these values we find that the observed overlap of 127 interactions agrees well with the expected overlap (689 × 50% × 40% = 138; P = 0.2, meaning that the actual overlap of 127 interactions is not significantly different from the theoretical overlap of 138 interactions).

Comparison of search space and number of identified interactions. Our PCR-stitching-into-next-generation-sequencing methodology, or Stitch-Seq, increases throughput and decreases cost, but does not significantly increase the coverage of interaction screens. The overall coverage of interaction screens is determined by "completeness", "assay-sensitivity", and "sampling sensitivity" as thoroughly analyzed in yeast[3] and human[5]. Next-generation sequencing does increase sequencing coverage, as we will be able to detect all sISTs that are successfully PCR stitched (in this manuscript, 454 sequencing identified 1,771 sISTs whereas Sanger sequencing identified 1,602 ISTs). This deficit can be easily remedied by re-sequencing all failed ISTs, as they are easy to identify in the Sanger sequencing results. We did not re-sequence failed ISTs because our aim was simply to demonstrate that the Stitch-Seq method works.

Our search space is 5,619 × 5,619 genes. The number of interactions (820) identified in our screen using Sanger sequencing agrees with our previous results. We can show this by calculating the expected number of interactions from a space of 5,619 × 5,619 with two methods:

1. A previous study from our group[5] concluded that: "the sampling sensitivity per screen is 53%, and that after a saturating number of screens, Y2H-CCSB can identify 118 interactions per million pairs tested." Our search space is 5,619 × 5,618 / 2 = 15.8 million pairs. 15.8 × 118 × 53% = 988, in close agreement with the 979 interactions identified by 454 sequencing. Because we did not re-sequence the failed IST pairs, we would lose about 10% of the ISTs: 988 × 90% = 889 close to the 820 interactions identified by Sanger sequencing in our screen.
2. We previously screened a space of 8,000 × 8,000 and produced 2,754 interactions for HI1[4]. Our space is about 49% of HI1: 2,754 × 49% × 53% = 715. The 820 interactions identified by Sanger sequencing in our screen exceeds the expected number[4].

Supplementary Note 4. Extrapolation to mapping the whole human interactome.

we previously estimated the size of the human interactome to be ~130,000 binary protein-protein interactions[5] (excluding additional complexity expected from splicing isoforms). Our HT-Y2H pipeline is able to determine ~25% of these interactions[1], or 32,500 interactions. In our human ORFeome v3.1 collection, ~97% of ORFs are shorter than 3 kb and >99% of all possible ORF combinations are less than 6 kb, a size range that works well with our Stitch-Seq strategy. For each pair to be successfully identified, all three PCR reactions have to succeed. With a PCR failure rate of 5%, the probability for all

three PCR reactions to be successful is $(100\% - 5\%)^3 = 85\%$. Therefore, we need to PCR stitch (with all three reactions) three times to be able to produce at least one successful sIST for one interacting pair $[100\% - (100\% - 85\%)^3 = 99\%]$. For 32,500 interactions, we would have $32{,}500 \times 3 \times 3 = 292{,}500$ PCR reactions and generate $32{,}500 \times 3 = 97{,}500$ sISTs for next-generation sequencing. Given that one 454 run can cover up to 10,000 sISTs, we would need about ten 454 runs, which would cost ~ $50,000.

Supplementary Note 5. Size cutoff for adapting Stitch-Seq.

If the whole capacity of 454 is dedicated for sequencing sISTs and the Sanger sequencing cost is $3/run, one would need to sequence at least 2,000 sISTs for Stitch-Seq to be cost effective. In our screen of a space of 6,000 × 6,000 ORFs, we generated 5,200 sISTs after one round of screening. Previous studies have shown that any search space needs to be screened at least four times to reach >85% saturation[5]. Therefore, given the current cost structure, the search space of a saturated screen should be at least 2,500 × 2,500 for cost-effective adoption of this new approach.

If the search space is smaller than 2,500 × 2,500 protein pairs, it does not necessarily mean that the Stitch-Seq methodology cannot be used. For smaller screens one can simply combine the sISTs with some other sequencing samples for a single 454 run. Hence there really is no size restriction for adopting this method. The sISTs sequencing will work with almost any other sequencing samples because the 82 bp linker sequence has no matching sequence in GenBank. When combining sISTs with other samples potential complications for assembling these other samples would need to be considered.

Supplementary Note 6. Useful read length and implementation of paired-end sequencing.

That only 5% of all 454 reads in our experiments produced sISTs at the end might seem troublesome, but this limitation is shared to some extent by all current IST or sIST sequencing strategies.

With Sanger technology, a similarly small fraction of each sequence trace is usable for identification of the interaction partners (~20-30 bases are sufficient) and hence >95% of sequence information is discarded.

Paired-end technology faces from the same problem. Due to the short read-length and the necessary fragmentation of the stitched amplicon, many fragments will not have enough sequence to identify both ORFs and hence will not identify interacting pairs. We can calculate that ~10% of paired-end reads would identify the interacting ORF pair. The average length of the sISTs in human ORFeome v3.1 is 3 kb and the average fragment length for sequencing is 500 bp. The 5' end of fragments for useful paired-end reads therefore has to be within a 352 bp region (500 bp less 82 bp linker less ~30 bp to identify each ORF). Therefore, $352/3{,}000 \approx 10\%$.

Our experiments were done with an earlier generation of 454 chemistry (454 FLX). Since then we have evaluated the more advanced 454 Titanium chemistry, which uses an input library of on average 500 bp fragments and yields read-lengths of ~400 bp, of which ~10% are informative for identification of interacting pairs.

Supplementary Note 7. Vector sequences for pDEST-ADCYH and pDEST-DB.

The following pDEST-DB and pDEST-ADCYH vectors have been used in interactome mapping projects for yeast, worm and human[1-9].

pDEST-DB vector sequence

```
LOCUS          pDest_pPC97 (pDEST-DB)  10262 bp ds-DNA  circular  23-FEB-2011
ACCESSION
VERSION
SOURCE
ORGANISM
FEATURES             Location/Qualifiers
     misc_feature    2066..2172
                     /note="attR1 site"
terminator      3983..4449
                     /note=”ADH1 TT”
misc_feature    complement(3795..3919)
                     /note="attR2 site"
rep_origin      7629..8005
                     /note=”ARS209”
misc_feature    3449..3754
                     /note="ccdB gene"
CDS             5805..6899
                     /note=”LEU2”
misc_feature    2449..3107
                     /note="CmR"
primer_bind     complement(4484..4501)
                     /note=”M13-fwd”
CDS             8178..9137
                     /note=”AmpR”
primer_bind     complement(4458..4485)
                     /note=”T7”
promoter        113..1564
                     /note=”ADH1 promoter”
CDS             4572..4640
                     /note=”LacZ alpha”
rep_origin      8015..8043
                     /note=”CEN6”
gene            1573..2032
                     /note=”GAL4 region”
CDS             1592..2032
                     /note=”GAL4-DB”
gene            5157..7376
                     /note=”LEU2 region”
rep_origin      complement(4646..5086)
                     /note=”F1 ori”
ORIGIN
        1 agcggataac aatttcacac aggaaacagc tatgaccatg attacgccaa gctcggaatt
       61 aaccctcact aaagggaaca aaagctgggt accgggcccc ccctcgagat ccgggatcga
      121 agaaatgatg gtaaatgaaa taggaaatca aggagcatga aggcaaaaga caaatataag
      181 ggtcgaacga aaaataaagt gaaaagtgtt gatatgatgt atttggcttt gcggcgccga
      241 aaaaacgagt ttacgcaatt gcacaatcat gctgactctg tggcggaccc gcgctcttgc
      301 cggcccggcg ataacgctgg gcgtgaggct gtgcccggcg gagttttttg cgcctgcatt
      361 ttccaaggtt taccctgcgc taaggggcga gattggagaa gcaataagaa tgccggttgg
      421 ggttgcgatg atgacgacca cgacaactgg tgtcattatt taagttgccg aaagaacctg
      481 agtgcatttg caacatgagt atactagaag aatgagccaa gacttgcgag acgcgagttt
      541 gccggtggtg cgaacaatag agcgaccatg accttgaagg tgagacgcgc ataaccgcta
      601 gagtactttg aagaggaaac agcaataggg ttgctaccag tataaataga caggtacata
      661 caacactgga aatggttgtc tgtttgagta cgctttcaat tcatttgggt gtgcacttta
      721 ttatgttaca atatggaagg gaactttaca cttctcctat gcacatatat taattaaagt
```

```
 781 ccaatgctag tagagaaggg gggtaacacc cctccgcgct cttttccgat ttttttctaa
 841 accgtggaat atttcggata tccttttgtt gtttccgggt gtacaatatg gacttcctct
 901 tttctggcaa ccaaacccat acatcgggat tcctataata ccttcgttgg tctccctaac
 961 atgtaggtgg cggaggggag atatacaata gaacagatac cagacaagac ataatgggct
1021 aaacaagact acaccaatta cactgcctca ttgatggtgg tacataacga actaatactg
1081 tagccctaga cttgatagcc atcatcatat cgaagtttca ctacccttt tccatttgcc
1141 atctattgaa gtaataatag gcgcatgcaa cttcttttct ttttttttct tttctctctc
1201 ccccgttgtt gtctcaccat atccgcaatg acaaaaaaaa tgatggaaga cactaaagga
1261 aaaaattaac gacaaagaca gcaccaacag atgtcgttgt tccagagctg atgaggggta
1321 tcttcgaaca cacgaaactt tttccttcct tcattcacgc acactactct ctaatgagca
1381 acggtatacg gccttccttc cagttacttg aatttgaaat aaaaaaagtt tgccgctttg
1441 ctatcaagta taaatagacc tgcaattatt aatcttttgt ttcctcgtca ttgttctcgt
1501 tccctttctt ccttgtTTCT TTTTCTGCAC AATATTTCAA GCTATACCAA GCATACAATC
1561 AACTCCAAGC TTGAAGCAAG CCTCCTGAAA GATGAAGCTA CTGTCTTCTA TCGAACAAGC
1621 ATGCGATATT TGCCGACTTA AAAAGCTCAA GTGCTCCAAA GAAAAACCGA AGTGCGCCAA
1681 GTGTCTGAAG AACAACTGGG AGTGTCGCTA CTCTCCCAAA ACCAAAAGGT CTCCGCTGAC
1741 TAGGGCACAT CTGACAGAAG TGGAATCAAG GCTAGAAAGA CTGGAACAGC TATTTCTACT
1801 GATTTTTCCT CGAGAAGACC TTGACATGAT TTTGAAAATG GATTCTTTAC AGGATATAAA
1861 AGCATTGTTA ACAGGATTAT TTGTACAAGA TAATGTGAAT AAAGATGCCG TCACAGATAG
1921 ATTGGCTTCA GTGGAGACTG ATATGCCTCT AACATTGAGA CAGCATAGAA TAAGTGCGAC
1981 ATCATCATCG GAAGAGAGTA GTAACAAAGG TCAAAGACAG TTGACTGTAT CGtcgagGTC
2041 GAATCAAACA AGTTTGTACA AAAAAGCTGA ACGAGAAACG TAAAATGATA TAAATATCAA
2101 TATATTAAAT TAGATTTTGC ATAAAAAACA GACTACATAA TACTGTAAAA CACAACATAT
2161 CCAGTCACTA TGGCGGCCGC GGGTGATGCT GCCAACTTAG CGGCCGCTAA GTTGGCAGCA
2221 TCACCCGACG CACTTTGCGC CGAATAAATA CCTGTGACGG AAGATCACTT CGCAGAATAA
2281 ATAAATCCTG GTGTCCCTGT TGATACCGGG AAGCCCTGGg CCAACTTTTG GCGAAAATGA
2341 GACGTTGATC GGCACGTAAG AGGTTCCAAC TTTCACCATA ATGAAATAAG ATCACTACCG
2401 GGCGTATTTT TTGAGTCATC GAGATTTTCA GGAGCTAAGG AAGCTAAAAT GGAGAAAAAA
2461 ATCACTGGAT ATACCACCGT TGATATATCC CAATGGCATC GTAAAGAACA TTTTGAGGCA
2521 TTTCAGTCAG TTGCTCAATG TACCTATAAC CAGACCGTTC AGCTGGATAT TACGGCCTTT
2581 TTAAAGACCG TAAAGAAAAA TAAGCACAAG TTTTATCCGG CCTTTATTCA CATTCTTGCC
2641 CGCCTGATGA ATGCTCATCC GGAATTCCGT ATGGCAATGA AAGACGGTGA GCTGGTGATA
2701 TGGGATAGTG TTCACCCTTG TTACACCGTT TTCCATGAGC AAACTGAAAC GTTTTCATCG
2761 CTCTGGAGTG AATACCACGA CGATTTCCGG CAGTTTCTAC ACATATATTC GCAAGATGTG
2821 GCGTGTTACG GTGAAAACCT GGCCTATTTC CCTAAAGGGT TTATTGAGAA TATGTTTTTC
2881 GTCTCAGCCA ATCCCTGGGT GAGTTTCACC AGTTTTGATT TAAACGTGGC CAATATGGAC
2941 AACTTCTTCG CCCCCGTTTT CACCATGGGC AAATATTATA CGCAAGGCGA CAAGGTGCTG
3001 ATGCCGCTGG CGATTCAGGT TCATCATGCC GTTTGTGATG GCTTCCATGT CGGCAGAATG
3061 CTTAATGAAT TACAACAGTA CTGCGATGAG TGGCAGGCGG GGCGTAATCT AGAGGATCCG
3121 GCTTACTAAA AGCCAGATAA CAGTATGCGT ATTTGCGCGC TGATTTTTGC GGTATAAGAA
3181 TATATACTGA TATGTATACC CGAAGTATGT CAAAAAGAGG TATGCTATGA AGCAGCGTAT
3241 TACAGTGACA GTTGACAGCG ACAGCTATCA GTTGCTCAAG GCATATATGA TGTCAATATC
3301 TCCGGTCTGG TAAGCACAAC CATGCAGAAT GAAGCCCGTC GTCTGCGTGC CGAACGCTGG
3361 AAAGCGGAAA ATCAGGAAGG GATGGCTGAG GTCGCCCGGT TTATTGAAAT GAACGGCTCT
3421 TTTGCTGACG AGAACAGGGG CTGGTGAAAT GCAGTTTAAG GTTTACACCT ATAAAAGAGA
3481 GAGCCGTTAT CGTCTGTTTG TGGATGTACA GAGTGATATT ATTGACACGC CCGGGCGACG
3541 GATGGTGATC CCCCTGGCCA GTGCACGTCT GCTGTCAGAT AAAGTCCCCC GTGAACTTTA
3601 CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA TGGCCAGTGT
3661 GCCGGTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA ATGACATCAA
3721 AAACGCCATT AACCTGATGT TCTGGGGAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA
3781 GTCTGCAGGT CGACCATAGT GACTGGATAT GTTGTGTTTT ACAGTATTAT GTAGTCTGTT
3841 TTTTATGCAA AATCTAATTT AATATATTGA TATTTATATC ATTTTACGTT TCTCGTTCAG
3901 CTTTCTTGTA CAAAGTGGTT TGATGGCCGC TAAgtaagta agacgtcgag ctctaagtaa
3961 gtaacggccg ccaccgcggt ggAGCTTTGG ACTTCTTCGC CAGAGGTTTG GTCAAGTCTC
4021 CAATCAAGGT TGTCGGCTTG TCTACCTTGC CAGAAATTTA CGAAAAGATG GAAAAGGGTC
4081 AAATCGTTGG TAGATACGTT GTTGACACTT CTAAATAAGC GAATTTCTTA TGATTTATGA
4141 TTTTTATTAT TAAATAAGTT ATAAAAAAAA TAAGTGTATA CAAATTTTAA AGTGACTCTT
4201 AGGTTTTAAA ACGAAAATTC TTATTCTTGA GTAACTCTTT CCTGTAGGTC AGGTTGCTTT
4261 CTCAGGTATA GCATGAGGTC GCTCTTATTG ACCACACCTC TACCGGCATG CCGAGCAAat
4321 gcctgcaaat cgctccccat ttcacccaat tgtagatatg ctaactccag caatgagttg
4381 atgaatctcg gtgtgtattt tatgtcctca gaggacaaca cctgttgtaa tcgttcttcc
4441 acacggatcc caattcgccc tatagtgagt cgtattacaa ttcactggcc gtcgttttac
4501 aacgtcgtga ctggGAAAAC CCTGGCGTTA CCCAACTTAA TCGCCTTGCA GCACATCCCC
4561 CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGCACCGA TCGCCCTTCC CAACAGTTGC
4621 GCAGCCTGAA TGGCGAATGG ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT
4681 GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT
4741 CTTCCCTTCC TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT
4801 CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTAGGG
4861 TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGGA
4921 GTCCACGTTC TTTAATAGTG GACTCTTGTT CCAAACTGGA ACAACACTCA ACCCTATCTC
4981 GGTCTATTCT TTTGATTTAT AAGGGATTTT GCCGATTTCG GCCTATTGGT TAAAAAATGA
5041 GCTGATTTAA CAAAAATTTA ACGCGAATTT TAACAAAATA TTAACGCTTA CAATTTCCTG
```

```
5101 ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTCAC ACCGCATATC GACCGGTCGA
5161 GGAGAACTTC TAGTATATCT ACATACCTAA TATTATTGCC TTATTAAAAA TGGAATCCCA
5221 ACAATTACAT CAAAATCCAC ATTCTCTTCA AAATCAATTG TCCTGTACTT CCTTGTTCAT
5281 GTGTGTTCAA AAACGTTATA TTTATAGGAT AATTATACTC TATTTCTCAA CAAGTAATTG
5341 GTTGTTTGGC CGAGCGGTCT AAGGCGCCTG ATTCAAGAAA TATCTTGACC Gcagttaact
5401 gtgggaatac tcaggtatcg taagatgcaa gagttcgaat ctcttagcaa ccattatttt
5461 tttcctcaac ataacgagaa cacacagggg cgctatcgca cagaatcaaa ttcgatgact
5521 ggaaattttt tgttaatttc agaggtcgcc tgacgcatat acctttttca actgaaaaat
5581 tgggagaaaa aggaaaggtg agagcgccgg aaccggcttt tcatatagaa tagagaagcg
5641 ttcatgacta aatgcttgca tcacaatact tgaagttgac aatattattt aaggacctat
5701 tgtttttttcc aataggtggt tagcaatcgt cttactttct aacttttctt acctttttaca
5761 tttcagcaat atatatatat atatttcaag gatataccat tctaatgtct gcccctaaga
5821 agatcgtcgt tttgccaggt gaccacgttg gtcaagaaat cacagccgaa gccattaagg
5881 ttcttaaagc tatttctgat gttcgttcca atgtcaagtt cgatttcgaa aatcatttaa
5941 ttggtggtgc tgctatcgat gctacaggtg ttccacttcc agatgaggcg ctggaagcct
6001 ccaagaaggc tgatgccgtt ttgttaggtg ctgtgggtgg tcctaaatgg ggtaccggta
6061 gtgttagacc tgaacaaggt ttactaaaaa tccgtaaaga acttcaattg tacgccaact
6121 taagaccatg taactttgca tccgactctc ttttagactt atctccaatc aagccacaat
6181 ttgctaaagg tactgacttc gttgttgtca gagaattagt gggaggtatt tactttggta
6241 agagaaagga agacgatggt gatggtgtcg cttgggatag tgaacaatac accgttccag
6301 aagtgcaaag aatcacaaga atggccgctt tcatggccct acaacatgag ccaccattgc
6361 ctatttggtc cttggataaa gctaatgttt tggcctcttc aagattatgg agaaaaactg
6421 tggaggaaac catcaagaac gaattcccta cattgaaggt tcaacatcaa ttgattgatt
6481 ctgccgccat gatcctagtt aagaacccaa cccacctaaa tggtattata atcaccagca
6541 acatgtttgg tgatatcatc tccgatgaag cctccgttat cccaggttcc ttgggtttgt
6601 tgccatctgc gtccttggcc tctttgccag acaagaacac cgcatttggt ttgtacgaac
6661 catgccacgg ttctgctcca gatttgccaa agaataaggt caaccctatc gccactatct
6721 tgtctgctgc aatgatgttg aaattgtcat tgaacttgcc tGAAGAAGGT AAGGCCATTG
6781 AAGATGCAGT TAAAAAGGTT TTGGATGCAG GTATCAGAAC TGGTGATTTA GGTGGTTCCA
6841 ACAGTACCAC CGAAGTCGGT GATGCTGTCG CCGAAGAAGT TAAGAAAATC CTTGCTTAAA
6901 AAGATTCTCT TTTTTTATGA TATTTGTACA TAAACTTTAT AAATGAAATT CATAATAGAA
6961 ACGACACGAA ATTACAAAAT GGAATATGTT CATAGGGTAG ACGAAACTAT ATACGCAATC
7021 TACATACATT TATCAAGAAG GAGAAAAAGG AGGATGTAAA GGAATACAGG TAAGCAAATT
7081 GATACTAATG GCTCAACGTG ATAAGGAAAA AGAATTGCAC TTTAACATTA ATATTGACAA
7141 GGAGGAGGGC ACCACACAAA AAGTTAGGTG TAACAGAAAA TCATGAAACT ATGATTCCTA
7201 ATTTATATAT TGGAGGATTT TCTCTAAAAA AAAAAAAATA CAACAAATAA AAAACACTCA
7261 ATGACCTGAC CATTTGATGG AGTTTAAGTC AATACCTTCT TGAACCATTT CCCATAATGG
7321 TGAAAGTTCC CTCAAGAATT TTACTCTGTC AGAAACGGCC TTAACGACGT AGTCGATATG
7381 GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGCCCC GACACCCGCC
7441 AACACCCGCT GACGCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC
7501 TGTGACCGTC TCCGGGAGCT GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC
7561 GAGACGAAAG GGCCTCGTGA TACGCCTATT TTtataggtt aatgtcatga TAATAATGGT
7621 TTCTTAggac ggatcgcttg cctgtaactt acacgcgcct cgtatctttt aatgatggaa
7681 taatttggga atttactctg tgtttattta tttttatgtt ttgtatttgg attttagaaa
7741 gtaaataaag aaggtagaag agttacggaa tgaagaaaaa aaaataaaca aaggtttaaa
7801 aaatttcaac aaaaagcgta ctttacatat atatttatta gacaagaaaa gcagattaaa
7861 tagatataca ttcgattaac gataagtaaa atgtaaaatc acaggatttt cgtgtgtggt
7921 cttctacaca gacaagatga aacaattcgg cattaatacc tgagagcagg aagagcaaga
7981 taaaaggtag tatttgttgg cgatcccccc agagtctttt acatcttcgg aaaacaaaaa
8041 ctatttttc tttaaatttct ttttttactt tctatttta atttatatat ttatattaaa
8101 aaatttaaat tataattatt tttatagcac gtgatgaaaa ggacccaggt ggcacttttc
8161 ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca aatatgtatc
8221 cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga
8281 gtattcaaca tttccgtgtc gcccttattc ccttttttgc ggcattttgc cttcctgttt
8341 ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag
8401 tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag
8461 aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta
8521 ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg
8581 agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca
8641 gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag
8701 gaccgaagga gctaaccgct ttttttcaca acatggggga tcatgtaact cgccttgatc
8761 gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg
8821 tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc
8881 ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg
8941 cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg
9001 gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga
9061 cgggcagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac
9121 tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag attgatttaa
9181 aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca
9241 aaatcccttа acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag
9301 gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac
9361 cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa
```

```
     9421 ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg tagttaggcc
     9481 accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc ctgttaccag
     9541 tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac
     9601 cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc
     9661 gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc
     9721 ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca
     9781 cgagggagct tccagggga aacgcctggt atctttatag tcctgtcggg tttcgccacc
     9841 tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg
     9901 ccagcaacgc ggcctttta cggttcctgg ccttttgctg gccttttgct cacatgttct
     9961 ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata
    10021 ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc
    10081 gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg
    10141 acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca
    10201 ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg
    10261 tg
//
```

pDEST-ADCYH vector sequence

```
LOCUS   pDest-pPC86+CYH  (pDEST-ADCYH) 10408 bp ds-DNA   circular   18-JAN-2006

COMMENT     Uppercase parts were sequenced directly; remainder inferred from
known plasmids.
COMMENT
FEATURES        Location/Qualifiers
Promoter        1573..3026
                     /note="ADH1 promoter"
CDS             3035..3055
                     /note="SV40 NLS"
CDS             3065..3412
                     /note="GAL4 AD"
terminator      5357..5823
                     /note="ADH1 TT"
gene            complement(6567..7344)
                     /note="TRP1"
rep_origin      7773..8291
                     /note="CEN6/ARSH4"
gene            8324..9283
                     /note="AmpR"
misc_feature    3425..3546
                     /note="attR1 site"
misc_feature    complement(5169..5293)
                     /note="attR2 site"
misc_feature    4823..5128
                     /note="ccdB gene"
misc_feature    3823..4481
                     /note="CmR"
misc_feature    552..1551
                     /note="CYH2 region"
CDS             548..997
                     /gene="CYH2"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA96382.1"
                     /db_xref="GI:1256572"
                     /translation="MPSRFTKTRKHRGHVSAGKGRIGKHRKHPGGRGMAGGQHHHRIN
                     MDKYHPGYFGKVGMRYFHKQQAHFWKPVLNLDKLWTLIPEDKRDQYLKSASKETAPVI
                     DTLAAGYGKILGKGRIPNVPVIVKARFVSKLAEEKIRAAGGVVELIA"

BASE COUNT     2991 a      2307 c      2212 g      2898 t
ORIGIN
        1 AGCGGATAAC AATTTCACAC AGGAAACAGC TATGACCATG ATTACGCCAA GCTCGAAATT
       61 AACCCTCACT AAAGGGAACA AAAGCTGGGA TCCCCAGATT CTATTTGCAG TTGCCAATCC
      121 TGTATCTCTT GTGCGGCTTC ATCAAAAGTC ACGCCACGTA GCTTACATAG AGATAACTTA
      181 GCATCGGTGA CTCTTCCATG TGCTAACAAC CATTTTGGAG ATTCATCGAA AAACACAGCA
      241 AAAACAAGAG TACTGTAACC AATGTAACAT CTGTACACCA GGGACCCACA CATTACcaaa
      301 atcaaaatta tttttctaat gccctgttat ttttcctatt ttcctctggc gcgtgaatag
      361 cccgcagaga cgcaaacaat tttcctcgca gtttttcgct tgtttaatgc gtattttccc
      421 agataggttc aaacctttca tctgtatccc gtatatttaa gatggcgttt gctttctccg
      481 ttgatttttt tccttcttag tgatttttt gcattaaatc ccagaacaat catccaacta
      541 atcaagaatg ccttccagat tcactaagac tagaaagcac agaggtcacg tctcagccgg
```

```
 601 taaaggtcgt atcggtaagc acagaaagca ccccggtggt agaggtatgg ccggtggtca
 661 acatcaccac agaattaaca tggataaata ccatccaggt tatttcggta aggttggtat
 721 gagatacttc cacaagcaac aagctcattt ctggaagcca gtcttgaact tggacaaatt
 781 gtggacattg atcccagaag acaagagaga ccaatacttg aaatctgctt ctaaggaaac
 841 tgctccagtt attgacactt tggcagccgg ttacggtaag atcttgggta agggtagaat
 901 cccaaatgtt ccagttatcg tcaaagctag attcgtctcc aagttggctg aagaaaaaat
 961 cagagctgct ggtggtgttg ttgaattgat cgcttaagcg catcaacaaa aagctctatg
1021 tattttccaa taaattatat atcttcagtt taatctaatt caacatctac ttctgtatta
1081 tttctatgaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
1141 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aagtACCAAG ATGGCCTTTG
1201 GTGGGTTGAA GAAGGAAAAA GACAGAAACG ACTTAATTAC CTACTTGAAA AAAGCCTGTG
1261 AGTAAACAGG CCCCTTTTCC TTTGTCGATA TCATGTAATT AGTTATGTCA CGCTTACATT
1321 CACGCCCTCC CCCCACATCC GCTCTAACCG AAAAGGAAGG AGTTAGACAA CCTGAAGTCT
1381 AGGTCCCTAT TTATTTTTTT ATAGTTATGT TAGTATTAAG AACGTTATTT ATATTTCAAA
1441 TTTTTCTTTT TTTTCTGTAC AGACGCGTGT ACGCATGTAA CATTATACTG AAAACCTTGC
1501 TTGAGAAGGT TTTGGGACGC TCGAAGGCTT TAATTTGCAA AGCTCGGGAT CCGGGCCCCC
1561 CCTCGAGATC CGGGATCGAA GAAAtgatgg taaatgaaat aggaaatcaa ggagcatgaa
1621 ggcaaaagac aaatataagg gtcgaacgaa aaataaagtg aaaagtgttg atatgatgta
1681 tttggctttg cggcgccgaa aaaacgagtt tacgcaattg cacaatcatg ctgactctgt
1741 ggcggacccg cgctcttgcc ggcccggcga taacgctggg cgtgaggctg tgcccggcgg
1801 agtttttgc gcctgcattt tccaaggttt accctgcgct aaggggcgag attggagaag
1861 caataagaat gccggttggg gttgcgatga tgacgaccac gacaactggt gtcattattt
1921 aagttgccga aagaacctga gtgcatttgc aacatgagta tactagaaga atgagccaag
1981 acttgcgaga cgcgagtttg ccggtggtgc gaacaataga gcgaccatga ccttgaaggt
2041 gagacgcgca taaccgctag agtactttga agaggaaaca gcaatagggt tgctaccagt
2101 ataaatagac aggtacatac aacactggaa atggttgtct gtttgagtac gctttcaatt
2161 catttgggtg tgcactttat tatgttacaa tatggaaggg aactttacac ttctcctatg
2221 cacatatatt aattaaagtc caatgctagt agagaagggg ggtaacaccc ctccgcgctc
2281 ttttccgatt tttttctaaa ccgtggaata tttcggatat ccttttgttg tttccgggtg
2341 tacaatatgg acttcctctt ttctggcaac caaacccata catcgggatt cctataatac
2401 cttcgttggt ctccctaaca tgtaggtggc ggaggggaga tatacaatag aacagatacc
2461 agacaagaca taatgggcta aacaagacta caccaattac actgcctcat tgatggtggt
2521 acataacgaa ctaatactgt agccctagac ttgatagcca tcatcatatc gaagtttcac
2581 tacccttttt ccatttgcca tctattgaag taataatagg cgcatgcaac ttcttttctt
2641 tttttttctt ttctctctcc cccgttgttg tctcaccata tccgcaatga caaaaaaaat
2701 gatggaagac actaaaggaa aaaattaacg acaaagacag caccaacaga tgtcgttgtt
2761 ccagagctga tgaggggtat cttcgaacac acgaaacttt ttccttcctt cattcacgca
2821 cactactctc taatgagcaa cggtatacgg ccttccttcc agttacttga atttgaaata
2881 aaaaaagttt gccgctttgc tatcaagtat aaatagacct gcaattatta atcttttgtt
2941 tcctcgtcat tgttctcgtt ccctttcttc cttgtttctt tttctgcaca atatttcaag
3001 ctataccaag catacaatca actccaagct tatgcccaag aagaagcgga aggtctcgag
3061 cggcgccaat tttaatcaaa gtgggaatat tgctgatagc tcattgtcct tcactttcac
3121 taacagtagc aacggtccga acctcataac aactcaaaca aattctcaag cgctttcaca
3181 accaattgcc tcctctaacg ttcatgataa cttcatgaat aatgaaatca cggctagtaa
3241 aattgatgat ggtaataatt caaaaccact gtcacctggt tggacggacc aaactgcgta
3301 taacgcgttt ggaatcacta cagggatgtt taataccact acaatggatg atgtATATAA
3361 CTATCTATTC GATGATGAAG ATACCCCACC AAACCCAAAA AAAGAGGGTG GGTCGAATCA
3421 AACAAGTTTG TACAAAAAAG CTGAACGAGA AACGTAAAAT GATATAAATA TCAATATATT
3481 AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAC ATATCCAGTC
3541 ACTATGGCGG CCGCGGGTGA TGCTGCCAAC TTAGCGGCCG CTAAGTTGGC AGCATCACCC
3601 GACGCACTTT GCGCCGAATA AATACCTGTG ACGGAAGATC ACTTCGCAGA ATAAATAAAT
3661 CCTGGTGTCC CTGTTGATAC CGGGAAGCCC TGGgCCAACT TTTGGCGAAA ATGAGACGTT
3721 GATCGGCACG TAAGAGGTTC CAACTTTCAC CATAATGAAA TAAGATCACT ACCGGGCGTA
3781 TTTTTTGAGT CATCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAAATCACT
3841 GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG AACATTTTGA GGCATTTCAG
3901 TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC CTTTTTAAAG
3961 ACCGTAAAGA AAAATAAGCA CAAGTTTTAT CCGGCCTTTA TTCACATTCT TGCCCGCCTG
4021 ATGAATGCTC ATCCGGAATT CCGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT
4081 AGTGTTCACC CTTGTTACAC CGTTTTCCAT GAGCAAACTG AAACGTTTTC ATCGCTCTGG
4141 AGTGAATACC ACGACGATTT CCGGCAGTTT CTACACATAT ATTCGCAAGA TGTGGCGTGT
4201 TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG AGAATATGTT TTTCGTCTCA
4261 GCCAATCCCT GGGTGAGTTT CACCAGTTTT GATTTAAACG TGGCCAATAT GGACAACTTC
4321 TTCGCCCCCG TTTTCACCAT GGGCAAATAT TATACGCAAG GCGACAAGGT GCTGATGCCG
4381 CTGGCGATTC AGGTTCATCA TGCCGTTTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAT
4441 GAATTACAAC AGTACTGCGA TGAGTGGCAG GCGGGGCGTA ATCTAGAGGA TCCGGCTTAC
4501 TAAAAGCCAG ATAACAGTAT GCGTATTTGC GCGCTGATTT TTGCGGTATA AGAATATATA
4561 CTGATATGTA TACCCGAAGT ATGTCAAAAA GAGGTATGCT ATGAAGCAGC GTATTACAGT
4621 GACAGTTGAC AGCGACAGCT ATCAGTTGCT CAAGGCATAT ATGATGTCAA TATCTCCGGT
4681 CTGGTAAGCA CAACCATGCA GAATGAAGCC CGTCGTCTGC GTGCCGAACG CTGGAAAGCG
4741 GAAAATCAGG AAGGGATGGC TGAGGTCGCC CGGTTTATTG AAATGAACGG CTCTTTTGCT
4801 GACGAGAACA GGGGCTGGTG AAATGCAGTT TAAGGTTTAC ACCTATAAAA GAGAGAGCCG
4861 TTATCGTCTG TTTGTGGATG TACAGAGTGA TATTATTGAC ACGCCCGGGC GACGGATGGT
```

4921 GATCCCCCTG GCCAGTGCAC GTCTGCTGTC AGATAAAGTC CCCCGTGAAC TTTACCCGGT
4981 GGTGCATATC GGGGATGAAA GCTGGCGCAT GATGACCACC GATATGGCCA GTGTGCCGGT
5041 CTCCGTTATC GGGGAAGAAG TGGCTGATCT CAGCCACCGC GAAAATGACA TCAAAAACGC
5101 CATTAACCTG ATGTTCTGGG GAATATAAAT GTCAGGCTCC CTTATACACA GCCAGTCTGC
5161 AGGTCGACCA TAGTGACTGG ATATGTTGTG TTTTACAGTA TTATGTAGTC TGTTTTTTAT
5221 GCAAAATCTA ATTTAATATA TTGATATTTA TATCATTTTA CGTTTCTCGT TCAGCTTTCT
5281 TGTACAAAGT GGTTTGATGG CCGCTAAGTA AGTAAGACGT CGAGCTCTAA GTAAGTAACG
5341 GCCGCCACCG CGGTGGAGCT TTGGACTTCT TCGCCAGAGG TTTGGTCAAG TCTCCAATCA
5401 AGGTTGTCGG CTTGTCTACC TTGCCAGAAA TTTACGAAAA GATGGAAAAG GGTCAAATCG
5461 TTGGTAGATA CGTTGTTGAC ACTTCTAAAT AAGCGAATTT CTTATGATTT ATGATTTTTA
5521 TTATTAAATA AGTTATAAAA AAAATAAGTG TATACAAATT TTAAAGTGAC TCTTAGGTTT
5581 TAAAACGAAA ATTCTTgttc ttgagtaact ctttcctgta ggtcaggttg ctttctcagg
5641 tatagcatga ggtcgctctt attgaccaca cctctaccgg catgccgagc aaatgcctgc
5701 aaatcgctcc ccatttcacc caattgtaga tatgctaact ccagcaatga gttgatgaat
5761 ctcggtgtgt attttatgtc ctcagaggac aatacctgtt gtaatcgttc ttccacacgg
5821 atcccaattc gccctatagt gagtcgtatt acaattcact ggccgtcgtt ttacaacgtc
5881 gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat ccccctttcg
5941 ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc
6001 tgaatggcga atggacgcgc cctgtagcgg cgcattaagc gcggcgggtg tggtggttac
6061 gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg ctttcttccc
6121 ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt
6181 agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt agggtgatgg
6241 ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt tggagtccac
6301 gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta tctcggtcta
6361 ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa atgagctgat
6421 ttaacaaaaa tttaacgcga attttaacaa aatattaacg tttacaattt cctgatgcgg
6481 tattttctcc ttacgcatct gtgcggtatt tcacaccgca ggcaagtgca caaacaatac
6541 ttaaataaat actactcagt aataacctat ttcttagcat ttttgacgaa atttgctatt
6601 ttgttagagt cttttacacc atttgtctcc acacctccgc ttacatcaac accaataacg
6661 ccatttaatc taagcgcatc accaacattt tctggcgtca gtccaccagc taacataaaa
6721 tgtaagcttt cggggctctc ttgccttcca acccagtcag aaatcgagtt ccaatccaaa
6781 agttcacctg tcccacctgc ttctgaatca aacaagggaa taaacgaatg aggtttctgt
6841 gaagctgcac tgagtagtat gttgcagtct tttggaaata cgagtctttt aataactggc
6901 aaaccgagga actcttggta ttcttgccac gactcatctc catgcagttg gacgatatca
6961 atgccgtaat cattgaccag agccaaaaca tcctccttag gttgattacg aaacacgcca
7021 accaagtatt tcggagtgcc tgaactattt ttatatgctt ttacaagact tgaaattttc
7081 cttgcaataa ccgggtcaat tgttctcttt ctattgggca cacatataat acccagcaag
7141 tcagcatcgg aatctagagc acattctgcg gcctctgtgc tctgcaagcc gcaaactttc
7201 accaatggac cagaactacc tgtgaaatta ataacagaca tactccaagc tgcctttgtg
7261 tgcttaatca cgtatactca cgtgctcaat agtcaccaat gccctccctc ttggccctct
7321 ccttttcttt tttcgaccga attaattctt aatcggcaaa aaaagaaaag ctccggatca
7381 agattgtacg taaggtgaca agctattttt caataaagaa tatcttccac tactgccatc
7441 tggcgtcata actgcaaagt acacatatat tacgatgctg tctattaaat gcttcctata
7501 ttatatatat agtaatgtcg tttatggtgc actctcagta caatctgctc tgatgccgca
7561 tagttaagcc agccccgaca cccgccaaca cccgctgacg cgccctgacg ggcttgtctg
7621 ctcccggcat ccgcttacag acaagctgtg accgtctccg ggagctgcat gtgtcagagg
7681 ttttcaccgt catcaccgaa acgcgcgaga cgaaagggcc tcgtgatacg cctatttta
7741 taggttaatg tcatgataat aatggtttct taggacggat cgcttgcctg taacttacac
7801 gcgcctcgta tcttttaatg atggaataat ttgggaattt actctgtgtt tatttatttt
7861 tatgttttgt atttggattt tagaaagtaa ataaagaagg tagaagagtt acggaatgaa
7921 gaaaaaaaaa taaacaaagg tttaaaaaat ttcaacaaaa agcgtacttt acatatatat
7981 ttattagaca agaaaagcag attaaataga tatacattcg attaacgata agtaaaatgt
8041 aaaatcacag gattttcgtg tgtggtcttc tacacagaca agatgaaaca attcggcatt
8101 aatacctgag agcaggaaga gcaagataaa aggtagtatt tgttggcgat ccccctagag
8161 tcttttacat cttcggaaaa caaaaactat tttttcttta atttcttttt ttactttcta
8221 tttttaattt atatatttat attaaaaaat ttaaattata attatttttа tagcacgtga
8281 tgaaaaggac ccaggtggca cttttcgggg aaatgtgcgc ggaaccccta tttgtttatt
8341 tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca
8401 ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt
8461 ttttgcggca ttttgccttc ctgtttttgc tcacccagaa acgctggtga aagtaaaaga
8521 tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca acagcggtaa
8581 gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagttct
8641 gctatgtggc gcggtattat cccgtattga cgccgggcaa gagcaactcg gtcgccgcat
8701 acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga
8761 tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata acactgcggc
8821 caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt ttcacaacat
8881 gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa
8941 cgacgagcgt gacaccacga tgcctgtagc aatggcaaca acgttgcgca aactattaac
9001 tggcgaacta cttactctag cttcccggca acaattaata gactggatgg aggcggataa
9061 agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc
9121 tggagccggt gagcgtgggt ctcgcggtat cattgcagca ctggggccag atggtaagcc
9181 ctcccgtatc gtagttatct acacgacggg cagtcaggca actatggatg aacgaaatag

```
     9241 acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag accaagttta
     9301 ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa
     9361 gatccttttt gataatctca tgaccaaaat cccttaacgt gagttttcgt tccactgagc
     9421 gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat
     9481 ctgctgcttg caaacaaaaa aaccaccgct accagcggtg gtttgtttgc cggatcaaga
     9541 gctaccaact ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt
     9601 ccttctagtg tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata
     9661 cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac
     9721 cgggttggac tcaagacgat agttaccgga taaggcgcag cggtcgggct gaacgggggg
     9781 ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg
     9841 tgagcattga gaaagcgcca cgcttcccga agggagaaag gcggacaggt atccggtaag
     9901 cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggggaacg cctggtatct
     9961 ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc
    10021 aggggggccg agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt
    10081 ttgctggcct tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg
    10141 tattaccgcc tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga
    10201 gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa ccGCCTCTCC CCGCGCGTTG
    10261 GCCGATTCAT TAATGCAGCT GGCACGACAG GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG
    10321 CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT
    10381 TCCGGCTCGT ATGTTGTGTG GAATTGTG
//
```

REFERENCES

1. Braun, P. *et al. Nat. Methods* 6, 91-97 (2009).
2. Dreze, M. *et al. Methods Enzymol* 470, 281-315 (2010).
3. Yu, H. *et al. Science* 322, 104-110 (2008).
4. Rual, J. F. *et al. Nature* 437, 1173-1178 (2005).
5. Venkatesan, K. *et al. Nat. Methods* 6, 83-90 (2009).
6. Boxem, M. *et al. Cell* 134, 534-545 (2008).
7. Simonis, N. *et al. Nat. Methods* 6, 47-54 (2009).
8. Dreze, M. *et al. Nat Methods* 6, 843-849 (2009).
9. Zhong, Q. *et al. Mol Syst Biol* 5, 321 (2009).