

Supporting Information

Cadoret *et al.* 10.1073/pnas.0805208105

Methods

Bioinformatic Analyses. For the analysis of conservation patterns in replication origins and in promoters, we retrieved 2991 transcripts corresponding to the set of 443 annotated known genes from the University of California, Santa Cruz (UCSC) ENCODE Website. The position of transcription start sites (TSS) was defined as the 5' end of those transcripts, and promoter regions were defined as the 1500 bp upstream of the TSS. Overlapping regions were grouped together. The final set includes 1035 promoter regions, with an average length of 1997 bp. Conserved regions (CRs) were retrieved from the UCSC ENCODE Website. These CRs correspond to the consensus of 3 methods and 3 different alignments from 28 different vertebrate species [Margulies EH, *et al.* (2007) Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. *Genome Res* 17(6):760–774]. We choose the moderate set, which provides CRs covering about 5% of the ENCODE regions [Margulies EH, *et al.* (2007) Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. *Genome Res* 17(6):760–774]. The conservation of sequences of interest (origins or promoters) was investigated by measuring the frequency of those regions that overlap with CRs. We tested the significance of the overlap by using a bootstrap of 1000 replicates of the same length of the origins or promoters randomly picked in the same ENCODE region

Amplification of SNS. First, we polydT tailed the 3' end of SNS with terminal transferase (New England Biolabs) as described by Liu, *et al.* Then, for second-strand synthesis and incorporation of the T7 promoter sequence, the tailing reaction product (20 μ l) was mixed with 1 μ l 0.15 μ M T7-A18B primer, 2.7 μ l 10X NEB-2 buffer, and 1 μ l 10 mM dNTPs (New England Biolabs). Samples were incubated at 94°C for 2 min, were ramped down at -1 C°/sec to 25°C, were held while T4 DNA polymerase (New England Biolabs) was added to a final concentration of 0.22 U/ μ l, and the volume was made up to 27 μ l with H₂O. The sample was incubated at 37°C for 30 min, and the reaction was stopped by the addition of 5 μ l 0.5 M EDTA pH 8.0. Samples were purified with the MinElute Reaction Cleanup Kit (Qiagen), increasing the elution volume to 20 μ l. Finally, the MegaScript Kit (Ambion) was used for *in vitro* transcription of a concentrated 8- μ l sample as described in Liu, *et al.*, except that the NTP pool was optimized for RNA labeling. Two μ l of a mix consisting of 100 mM ATP, CTP, and GTP, 2 μ l 50 mM aminoallyl-UTP (Fermentas) and 1 μ l UTP 100 mM were added. The product was purified with the RNeasy Minikit (Qiagen), and RNA was quantified by Nanodrop and analyzed by BioAnalyser with Lab-on-a-chip total RNA nano biosizing assay (Agilent Technologies).

A

Chromosome	ENCODE begin	ENCODE end	N-domain border	Closest exp origin	Distance
chr7	89428339	90542763	89871000	89871631	631
chr7	115404471	117281897	115429000	115728548	299548
chr8	118882220	119382220	119195000	119250180	55180
chr11	63940888	64440888	64269000	64251467	17533
chr18	23719231	24219231	24011000	24011725	725
chr21	32668236	34364221	33066000	33143912	77912
chr21	32668236	34364221	33525000	33446859	78141

B

Chromosome	ENCODE begin	ENCODE end	N-domain begin	N-domain end	Nb. origins
chr2	234273824	234773888	234060000	235644000	5
chr7	113527083	114527083	113331000	115429000	4
chr7	125672606	126835803	124166000	126880000	7
chr14	52947075	53447075	52702000	54042000	1
chr16	25780427	26280428	24943000	27477000	0

Fig. S4. Intersection between ENCODE regions and N-domains of replication computationally identified [Huvet M, *et al.* (2007) Human gene organization driven by the coordination of replication and transcription. *Genome Res* 17(9):1278–1285]. (A) Experimental validation of 2 computationally predicted origins. Six ENCODE regions contain 7 N-domain borders (i.e., putative replication origins). Positions of ENCODE regions borders, N-domain borders, closest experimental origin, and distance between each putative origin and its closest experimental origin are shown. Two of 7 are located less than 1 kb from our experimentally validated origin. We tested the significance of this overlap by using a bootstrap of 1000 replicates, of the same length of the origins, picked randomly in the same ENCODE region. The probability that 2 computationally predicted origins fall by chance less than 1 kb from an experimentally characterized origin is only 10^{-4} . (B) ENCODE regions included in N-domains contain 0 to 7 origins. Positions of ENCODE regions borders, N-domain borders, and number of origins found experimentally are shown.

Table S1. Origin segments

ENCODE region	Chromosome	Beginning	End	Origin class
ENr231	chr1	147984909	147986846	
	chr1	148244302	148246331	
	chr1	148258507	148260234	
	chr1	148296747	148298101	
	chr1	148360133	148361364	
ENr112	chr2	51943931	51945315	Isolated
ENr121	chr2	118505924	118507071	
ENr331	chr2	220107251	220108794	
	chr2	220135657	220137337	
	chr2	220165844	220168917	
	chr2	220174966	220176411	
	chr2	220188829	220191429	
	chr2	220198477	220200657	Highly dense
	chr2	220201886	220203102	
	chr2	220228419	220230263	
	chr2	220242014	220245496	
	chr2	220250475	220251903	Highly dense
	chr2	220261756	220263177	
	chr2	220336089	220337258	
	chr2	220338833	220340250	
	chr2	220541427	220543128	
	ENr131	chr2	234486579	234487948
chr2		234656507	234657628	
chr2		234666547	234667905	Highly dense
chr2		234671440	234672777	
ENr221	chr2	234740455	234741601	
	chr5	55873281	55874378	
	chr5	55981348	55982831	Isolated
	chr5	56085418	56086517	
	chr5	56088066	56089478	
	chr5	56119168	56120680	
ENm002	chr5	56147525	56149060	
	chr5	56241199	56242625	
	chr5	131571851	131573644	
	chr5	131620957	131622317	
	chr5	131636239	131637448	
	chr5	131744261	131745522	
	chr5	131801134	131802891	
	chr5	131860038	131861573	
	chr5	132025755	132026955	
	chr5	132048184	132049426	
	chr5	132151529	132152659	
ENr212	chr5	132248525	132250628	
	chr5	141950090	141951462	
	chr5	141999954	142001505	
	chr5	142078249	142081176	
	chr5	142176213	142178174	
	chr5	142199652	142201376	
	chr5	142236258	142237249	
	chr5	142259552	142261293	
ENr334	chr5	142361833	142363780	
	chr6	41518826	41520409	
	chr6	41543060	41544342	
	chr6	41544942	41546022	
	chr6	41578447	41579719	
	chr6	41585098	41586734	
	chr6	41599804	41601172	
	chr6	41622513	41624061	
	chr6	41654120	41655750	

continued

Table S1. Continued

ENCODE region	Chromosome	Beginning	End	Origin class XXXXXXX
	chr6	41719178	41720147	
	chr6	41854596	41855697	
	chr6	41856075	41857739	
ENr223	chr6	74011080	74012121	
	chr6	74108528	74110397	
	chr6	74160713	74161992	
	chr6	74287525	74288636	
ENr323	chr6	108385466	108386797	
	chr6	108547474	108548773	
	chr6	108552017	108553295	
	chr6	108711089	108712249	Isolated
ENr222	chr6	132399844	132401367	Isolated
ENm010	chr7	26920770	26922694	
	chr7	26922861	26924734	Highly dense
	chr7	26926630	26928154	
	chr7	26941783	26943229	
	chr7	26943851	26945045	Highly dense
	chr7	26952545	26954275	
	chr7	26967119	26968225	
	chr7	26982812	26983987	
	chr7	27028731	27030030	
	chr7	27032062	27033765	
	chr7	27047481	27048910	
	chr7	27056729	27057891	
ENm013	chr7	89478648	89479835	
	chr7	89519250	89520459	
	chr7	89523106	89525040	
	chr7	89677540	89679141	Isolated
	chr7	89870257	89871631	Isolated
	chr7	90037675	90039130	
	chr7	90089748	90091313	
	chr7	90240841	90242057	Isolated
ENm012	chr7	113873919	113875052	Isolated
	chr7	114116473	114118176	Isolated
	chr7	114318566	114319629	
	chr7	114351397	114352460	
ENm001	chr7	115728548	115730087	
	chr7	115733502	115735495	Highly dense
	chr7	115739541	115741149	
	chr7	115758143	115759653	
	chr7	115760088	115761361	
	chr7	115792540	115794616	
	chr7	115932781	115934827	
	chr7	116029862	116032737	
	chr7	116096721	116098030	
	chr7	116187758	116189117	
	chr7	116301540	116302519	
	chr7	116398408	116400119	
	chr7	116413422	116414684	
	chr7	116425234	116426588	
	chr7	116556983	116558225	Isolated
	chr7	116713437	116715501	
	chr7	116720877	116722362	
	chr7	116799265	116800441	
	chr7	117025196	117026369	Isolated
ENm014	chr7	126079638	126080850	
	chr7	126112892	126113549	
	chr7	126247472	126248824	Isolated
	chr7	126581134	126582490	
	chr7	126626567	126628059	

continued

Table S1. Continued

ENCODE region	Chromosome	Beginning	End	Origin class
ENr322	chr14	98486278	98487757	Highly dense
	chr14	98696258	98697981	
	chr14	98724314	98725645	
	chr14	98767316	98769158	
	chr14	98771506	98772663	
	chr14	98782269	98783573	
	chr14	98808316	98809543	
	chr14	98881859	98883202	
	chr14	98956703	98958083	
	ENr233	chr15	41735285	
chr15		41850700	41852158	
chr15		42013413	42014920	
ENm008	chr16	47415	48530	
	chr16	128359	129221	
	chr16	144166	145603	
	chr16	154879	156172	
	chr16	176977	177951	
	chr16	265752	267214	
	chr16	271684	272766	
	chr16	343083	344285	
	chr16	353862	355393	
	chr16	371926	373475	
	chr16	391235	392389	
	chr16	458042	459289	
	ENr313	chr16	60988052	
ENr213	chr18	24011725	24012891	
ENm007	chr18	24064501	24065770	
	chr19	59062973	59064186	
	chr19	59065358	59067169	
	chr19	59077897	59080362	
	chr19	59093470	59095647	
	chr19	59104275	59106633	
	chr19	59138068	59139337	
	chr19	59155330	59157173	
	chr19	59174913	59178718	
	chr19	59186590	59187790	
	chr19	59206956	59208156	
	chr19	59225201	59226338	
	chr19	59340003	59341368	
	chr19	59357776	59359107	
	chr19	59385445	59387234	
	chr19	59588883	59590167	
	chr19	59633614	59634656	
	ENr333	chr20	33449379	
chr20		33520581	33521619	
chr20		33593121	33594722	
chr20		33604647	33606279	
chr20		33652973	33654116	
chr20		33667226	33668962	
chr20		33763526	33764778	
chr20		33764778	33764778	
ENm005	chr21	32736476	32738137	
	chr21	32825833	32827210	
	chr21	32924855	32926140	
	chr21	33143912	33145612	
	chr21	33314649	33316910	
	chr21	33321658	33323457	
	chr21	33371196	33372523	
	chr21	33445752	33446859	
	chr21	33697620	33699322	
	chr21	33773906	33775426	
	chr21	33847294	33849173	
	chr21	33869514	33870830	
	chr21	33937141	33938736	

continued

Table S1. Continued

ENCODE region	Chromosome	Beginning	End	Origin class
ENr133	chr21	39299170	39300343	
	chr21	39377239	39378501	
	chr21	39607655	39608789	Isolated
ENm004	chr22	30314293	30315055	
	chr22	30336429	30337386	
	chr22	30349689	30350809	
	chr22	30802515	30804264	
	chr22	30843605	30844819	
	chr22	30896070	30897269	
	chr22	31030409	31031414	Isolated
	chr22	31156785	31158209	
	chr22	31195329	31197371	
	chr22	31366116	31367680	
	chr22	31437500	31439348	
	chr22	31530713	31531675	
	chr22	31545005	31546201	
ENr324	chrX	122559999	122561372	
	chrX	122590866	122592123	
	chrX	122624067	122625387	
	chrX	122660854	122662000	
	chrX	122719627	122720715	
	chrX	122820974	122822324	Isolated
	chrX	122959521	122961890	
ENm006	chrX	152662042	152663313	
	chrX	152689689	152691828	
	chrX	152704242	152705388	
	chrX	152721163	152722370	
	chrX	152757248	152758442	
	chrX	152816486	152818401	
	chrX	152883775	152885328	
	chrX	153070355	153071501	
	chrX	153081356	153082612	
	chrX	153123710	153124807	
	chrX	153130060	153131533	
	chrX	153181711	153184284	
	chrX	153236591	153237728	
	chrX	153249941	153251354	
	chrX	153327230	153328684	
	chrX	153338595	153339968	OriG6PD
	chrX	153818861	153819900	Isolated

Table S2. Short ssDNA segments obtained with stringent criteria

Chromosome	Beginning	End
Chr1	148302434	148303315
Chr11	4730926	4731822
Chr11	63994433	63995679
Chr14	98908400	98909476
Chr2	234629711	234630595
Chr21	34015696	34017159
Chr5	56295243	56296211
Chr7	113942120	113943850
Chr8	119342198	119343366

Table S3. Short ssDNA segments obtained with less stringent criteria

Chromosome	Beginning	End
Chr1	147994346	147995114
Chr1	148045616	148046395
Chr1	148089197	148090135
Chr1	148135326	148136324
Chr1	148140049	148140794
Chr1	148142404	148143288
Chr1	148281719	148282862
Chr1	148289583	148291287
Chr1	148302434	148303315
Chr1	148330817	148332181
Chr1	148389782	148391035
Chr10	55227638	55228527
Chr10	55459529	55460848
Chr11	2093511	2094406
Chr11	2167384	2168284
Chr11	2191849	2192739
Chr11	2298623	2299539
Chr11	4730926	4732277
Chr11	5133531	5134393
Chr11	5313319	5314209
Chr11	5595938	5596841
Chr11	5680414	5681160
Chr11	5687697	5689222
Chr11	63990114	63990772
Chr11	63994206	63995679
Chr11	64021389	64022128
Chr11	64102109	64103140
Chr11	64189138	64190272
Chr11	64273426	64274383
Chr11	64391691	64392701
Chr11	116157738	116158738
Chr12	38626913	38627983
Chr12	38909853	38911145
Chr13	29568800	29569735
Chr13	112488810	112489461
Chr13	112606624	112607524
Chr13	112808792	112809833
Chr14	53151432	53152397
Chr14	53231677	53233208
Chr14	53349472	53351117
Chr14	98848330	98849434
Chr14	98906288	98907151
Chr14	98908400	98909476
Chr15	41546187	41547066
Chr15	41705071	41705723
Chr15	41711222	41711881
Chr15	41716599	41717383
Chr15	41727851	41728706
Chr15	41835861	41836733
Chr15	41947175	41948454
Chr16	37951	38703
Chr16	153019	153663
Chr16	222700	223746
Chr16	239942	240852
Chr16	350344	351446
Chr16	403327	404203
Chr16	422121	422871
Chr16	25980622	25981932
Chr16	26075229	26076594
Chr16	61111353	61112457
Chr18	24039017	24040002
Chr18	59497731	59498624

continued

Table S3. Continued

Chromosome	Beginning	End
Chr19	59083008	59083928
Chr19	59207244	59208156
Chr19	59276176	59277045
Chr19	59392364	59393406
Chr19	59439949	59440608
Chr19	59483497	59484141
Chr19	59503354	59504013
Chr19	59572518	59573340
Chr19	59613466	59614421
Chr19	59666309	59667473
Chr19	59759818	59761248
Chr19	59852668	59854113
Chr19	59879347	59880087
Chr2	51690973	51692246
Chr2	51906020	51907160
Chr2	51925613	51926744
Chr2	118106655	118107792
Chr2	118245142	118246802
Chr2	118388678	118389547
Chr2	220518360	220519278
Chr2	234384126	234385388
Chr2	234481585	234482489
Chr2	234575429	234576374
Chr2	234630670	234631327
Chr2	234667846	234669209
Chr2	234698012	234699267
Chr20	33406227	33407354
Chr20	33420723	33421839
Chr21	32768862	32770011
Chr21	32850099	32850751
Chr21	32886148	32887036
Chr21	32917818	32918949
Chr21	33059062	33060306
Chr21	33539418	33540507
Chr21	33560645	33561527
Chr21	33629320	33630285
Chr21	33704941	33705870
Chr21	33779372	33780125
Chr21	33782613	33783974
Chr21	33866109	33867218
Chr21	34015696	34017159
Chr21	34323487	34324414
Chr21	39664436	39665180
Chr21	39734404	39735264
Chr22	30163672	30164633
Chr22	30231331	30232276
Chr22	30289137	30289796
Chr22	30291389	30292112
Chr22	30360511	30361402
Chr22	30425554	30426213
Chr22	30486758	30487763
Chr22	30509856	30510798
Chr22	30634326	30635177
Chr22	30788719	30789626
Chr22	31047522	31048424
Chr22	31250037	31251055
Chr22	31293476	31294587
Chr22	31357188	31358002
Chr22	31609686	31610563
Chr4	118641431	118642304
Chr4	119048139	119049046
Chr5	55965960	55967436
Chr5	56161424	56162939

continued

Table S4. Continued

Gene name	Forward primer sequence validation origins	Reverse primer sequence	Amplicon position
background1	AGGGCTGAGCCATAATTCTTCT	CTGCAATGCACTCACAACAAC	chr2:51629857 + 51630010
background2	GGCAAGGGAAGGGGTAAC	ATTCCCACTCATTGCTCAGG	chr4:118920576 + 118920719
background3	CACTGTCAAAGTCTGTGAAGCTATT	AGGCCAAACGTTTTTCTTTG	chr6:132581326 + 132581470
background4	TAGGCGCAGCTTGTAGGACT	CAGGATTTGGGACAACCTGG	chr11:4898934 + 4899089
background5	CTTGACAATGCCTCACTCA	GAAAACACCAGCCACCAGAA	chr7:89922194 + 89922313
background6	GATGCTGAGGGGTTTCAAA	TGCAAGCCACTTTCGGTTAG	chr7:116505991 + 116506123
background7	TGACATACAACCAACCAGAAC	GGAATGAGATGGACAGCAAATC	chr7:26880046 + 26880177
background8	CTTTGAAGGAAATCCCAGTTGTC	CCGGCTTCCCCTGTAGTTATT	chr7:27109288 + 27109393
Highly dense origins			
T1	GGCTTGGAGTTTCAACAGGA	CGCAGGTACGAGTCAACAAC	chr7:115734636 + 115734791
T2	GGTGTAGATTGGGGTTTTGT	TGTCCCAGTTGTCCAGTC	chr2:234664712 + 234664826
T3	CAGTGAATGGGATGCTCTG	CTGTCTCCCTGTGGGTGAG	chr7:26927452 + 26927569
T4	CCTGCCCTTGGTAGGTCTC	CAGCTCATCTGGCAACACAG	chr2:220197823 + 220197951
T5	TGGTTGGCTGGCTACTCT	CGCGTCTACTCTGATGAC	chr14:98783632 + 98783755
T6	AGCACCTCTGTCTGCCTCAA	CCTTCCCCTGTGCTAGTTC	chr9:128864132 + 128864273
T7	CCTGGAATGTAGGCTGTGTA	CCCCTCTAGCACCCCAACT	chr11:64339669 + 64339802
T8	CCTTCTCCATCTGTGCTGCT	CTGTGAGCCCTTCAAATTGC	chr13:112428887 + 112429038
Isolated origins			
T9	TCGTGGCGTCATTCTGTGTA	CTTAAAGACGTGCTCCACTGC	chrX:153819393 + 153819546
T10	AGCGAGATTGGGTGTTTCTAG	CAACAGCCCCTGAGCTCTATT	chr7:89871198 + 89871332
T11	AGCGTGCAATGTGAGGAAAC	GGAGAGAAGCGAAGCAATGA	chr7:116557558 + 116557666
T12	GGGAAGGAAATGCAAGACAA	AATTTGGCTGCTTAGCATGG	chr7:126268071 + 126268212
ENCODE regions without origins			
T13	CCTCATTGTAGAAGTATGTCCT	GGTGTAATTAAGTCTGAGTGGACT	chr11:5280585—5280735
T14	AGAAAAAGGAGCTTCCAGTCG	TGTGGGTGCAATGAACTGT	chr4:118853686 + 118853826
T15	AGTGGAACAGCCAGGTGTGA	TGATCTTTTCTCGGGGTGTG	chr10:55415286 + 55415418
T16	CTGCCTTGTGAGTGTGAAG	CGCTGAGCCAGGATTTATGT	chr18:59660223 + 59660371
T17	ACCACCATCTGGGAAAAAAC	GTTGGGAACCTGCCCTTTC	chr11:116160417 + 116160536
T18	CACGCAGACTGTCAAAGAA	GGAGAGGGGAAATACTGGA	chr16:26054593 + 26054749
T19	CTGGACAGGAATCCAGGAAA	CCTTGATTGGTACCCACACC	chr16:25955735 + 25955892