

Supplementary Text

Comparisons of Shemesh's and CENTP annotations

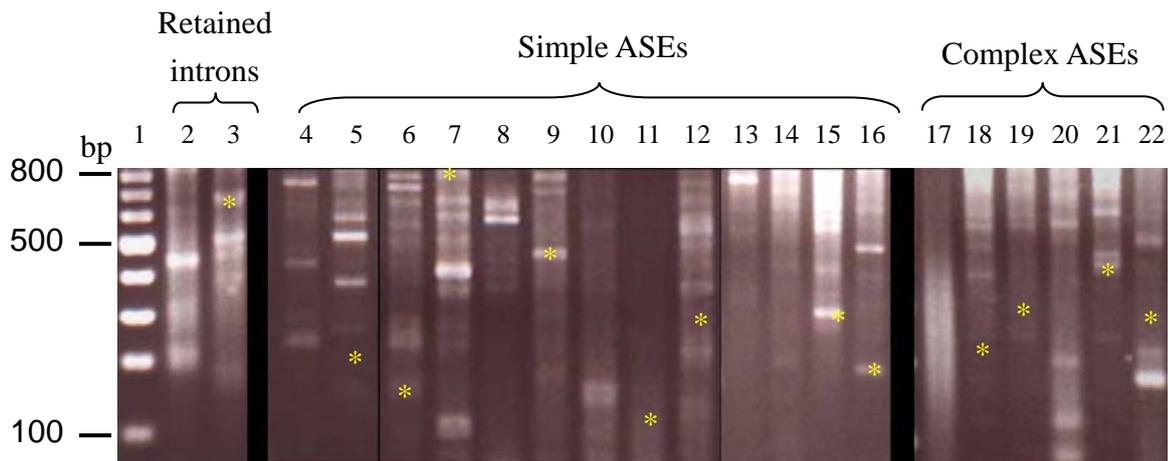
Compared with a recent study that applied human PPGs to ASV detection (Shemesh et al. 2006), CENTP identifies a much larger number of human PPG-based novel exons (i.e. the CENTP_{H_PPG} exons; 89 vs. 22 for cassette-on exons and 32 vs. 3 for retained introns). Possible reasons for the difference include: (1) Shemesh et al.'s (2006) study used only the PPGs identified by Zhang et al. (Zhang et al. 2003) but not the Ensembl-annotated PPGs, whereas both are included in our analysis; (2) Our analysis is based on a more updated version of the human genome as compared with the previous study (NCBI Build 36.1 vs. Build 34); (3) We align the PPGs to all of the well-annotated gene loci, while the previous study only compared PPG-parent gene pairs; (4) The two studies use different filtering criteria to reduce false positives, thus leading to different prediction results; and (5) CENTP identifies complex cassette-on exons, which are not considered in the previous study.

Experimental validation of the CENTP-identified exons/ASVs

Human cDNA was synthesized by RT-PCR with the SuperScriptTM First Strand Synthesis kit (Invitrogen. U.S.A) from the Human Total RNA Survey Panel (Ambion Inc. U.S.A), including the 20 following tissues: adipose, bladder, brain, cervix, colon, esophagus, heart, kidney, liver, lung, ovary, placenta, prostate, skeleton muscle, small intestine, spleen, testes, thymus, thyroid, and trachea. The cDNA was mixed and used as the template for PCR by the *Ex Taq* polymerase kit (TaKaRa Bio Inc. Japan). The primers were designed specifically to bind the flanking exons of the novel exons to be verified. The PCR products of expected sizes were then purified by the QIAquick Extraction kit (QIAGEN Inc. U.S.A) and auto-sequenced.

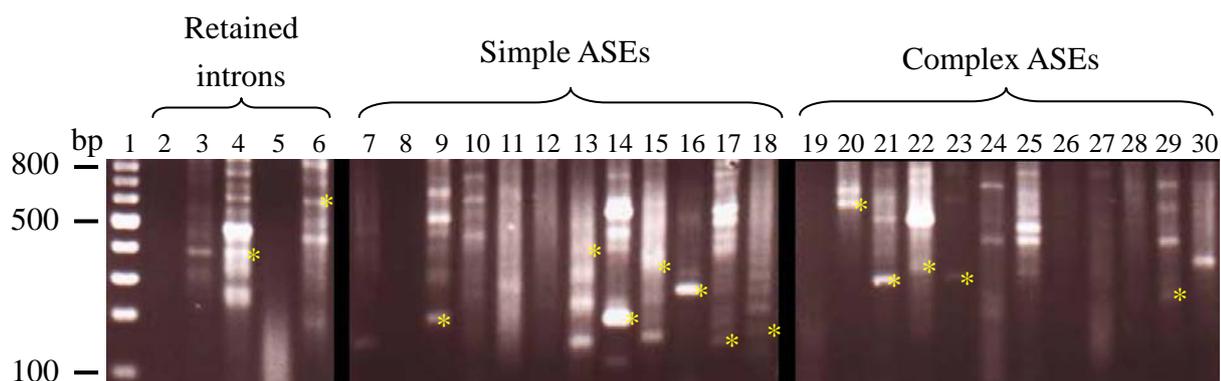
Shemesh, R., Novik, A., Edelheit, S., and Sorek, R. 2006. Genomic fossils as a snapshot of the human transcriptome. *Proc Natl Acad Sci U S A* **103**: 1364-1369.

Zhang, Z., Harrison, P.M., Liu, Y., and Gerstein, M. 2003. Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. *Genome Res* **13**: 2541-2558.



AS type	Lane	Original transcript accession ID of chimpanzee PPG	Human transcript accession ID	Experiment confirmed?
Retained introns	2	ENSTPTRP00000027558	U50529	No
	3	ENSTPTRP00000028865	AK095492	Yes
Simple ASEs	4	ENSTPTRP00000008507	AK095342	No
	5	ENSTPTRP00000008507	BC072018	Yes
	6	ENSTPTRP00000008507	NM_000422	Yes
	7	ENSTPTRP00000031668	NM_138459	Yes
	8	ENSTPTRP00000037527	NM_174961	No
	9	ENSTPTRP00000017168	AJ555149	Yes
	10	ENSTPTRP00000023017	NM_005342	No
	11	ENSTPTRP00000002870	NM_014864	Yes
	12	ENSTPTRP00000028611	NM_004174	Yes
	13	ENSTPTRP00000031135	NM_001014441	No
	14	ENSTPTRP00000034816	NM_000971	No
	15	ENSTPTRP00000008506	NM_005554	Yes
	16	ENSTPTRP00000038681	AK126842	Yes
	Complex ASEs	17	ENSTPTRP00000028924	NM_004660
18		ENSTPTRP00000037527	NM_005635 NM_174961	Yes
19		ENSTPTRP00000031135	NM_001014441	Yes
20		ENSTPTRP00000014279	AF285758 NM_005548	No
21		ENSTPTRP00000016258	NM_002266	Yes
22		ENSTPTRP00000038681	AK126842	Yes

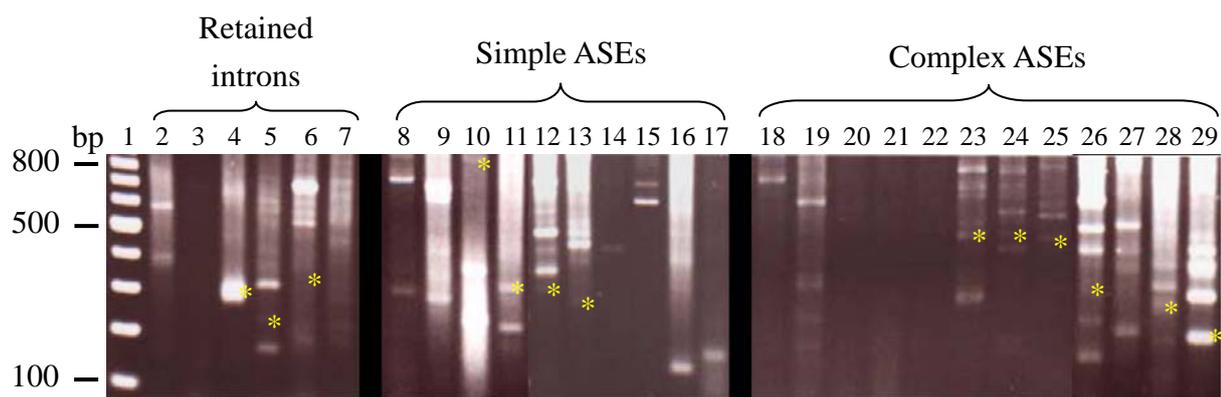
Supplementary Figure 1. RT-PCR verification results of CENTP exons supported by chimpanzee PPGs. Lane 1: 100bp-DNA marker; Lanes 2~3: retained introns; Lanes 4~16: simple cassette exons; Lanes 17~22: complex cassette exons. The stars indicate the location of the novel ASVs.



AS type	Lane	Human PPG ID	Human transcript ID	Experiment confirmed?	
Retained introns	2	Q9H8Q9.Human.chr18.mb55	AK126787	No	
	3	Q9H8Q9.Human.chr17.mb55	AK126787	No	
	4	Q9H8Q9.Human.chr15.mb51	AK126787	Yes	
	5	ENSP00000328595.Human.chr16.mb0	BC048328	No	
	6	Q96Q57.Human.chr15.mb80	AF316855	Yes	
	Simple ASEs	7	OFD1_HUMAN.Human.chr5.mb37	AF332238	No
8		OFD1_HUMAN.Human.chr5.mb37	AF332238	No	
9		ENSP00000349032.Human.chr19.mb62	BC019034	Yes	
10			NM_182752		
Complex ASEs		11	Q9NYV6.Human.chr2.mb143	BC104660	No
		12		NM_018427	
		13	ENSP00000351114.Human.chr7.mb149	AK127288	No
		14	ZNF85_HUMAN.Human.chr19.mb20	NM_024498	No
		15	ENSP00000335042.Human.chr7.mb64	NM_007153	Yes
		16	ENSP00000339073.Human.chr1.mb238	NM_000984	Yes
	17	ZN492_HUMAN.Human.chr12.mb56	AK126842	Yes	
	18	HMG4L_HUMAN.Human.chr3.mb26	NM_005342	Yes	
Complex ASEs	19	WEE1_HUMAN.Human.chr11.mb35	NM_173677	Yes	
	20	ENSP00000345525.Human.chr7.mb63	NM_007153	Yes	
	21	SSX5_HUMAN.Human.chr6.mb119	NM_005635	No	
	22	ENSP00000335042.Human.chr7.mb64	NM_024498	Yes	
	23	ENSP00000293618.Human.chr15.mb34	AF370416	Yes	
	24	Q6P4C7.Human.chr12.mb60	NM_005554	Yes	
	25	ENSP00000236251.Human.chr12.mb93	NM_004905	Yes	
Complex ASEs	26	ZN492_HUMAN.Human.chr12.mb56	NM_178549	No	
	27	ENSP00000347298.Human.chr17.mb18	NM_001007530	No	
	28				
	29				
	30				

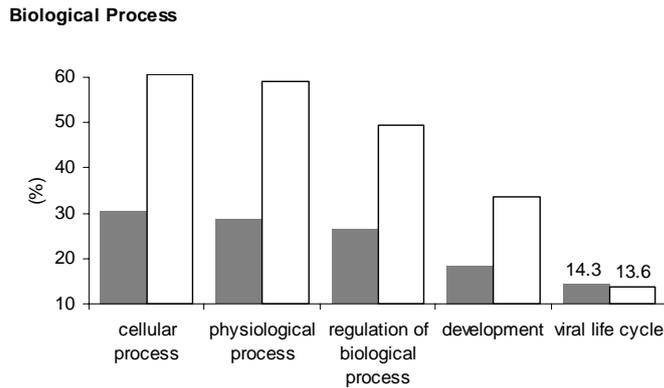
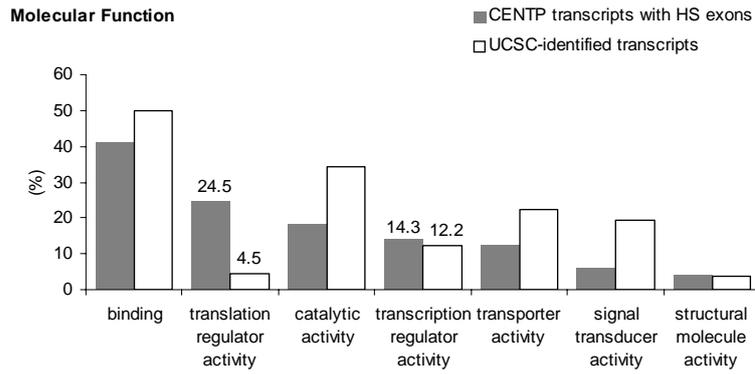
26	ENSP00000339575.Human.chr7.mb62	NM_001004321	No
27	ENSP00000266116.Human.chr19.mb58	NM_007153	No
28	Q6ZNB9.Human.chr3.mb32	NM_007153	No
29	ENSP00000352322.Human.chr13.mb50	NM_001039567	Yes
30	KI2S2_HUMAN.Human.chr19.mb59	ENST00000336077	No

Supplementary Figure 2. RT-PCR verification results of CENTP exons supported by human PPGs. Lane 1: 100bp-DNA marker; Lanes 2~6: retained introns; Lanes 7~18: simple cassette exons; Lanes 19~30: complex cassette exons. The stars indicate the location of the novel ASVs.



AS type	Lane	Chimpanzee gene ID	Human transcript ID	Experiment confirmed?
Retained introns	2	ENSPTRT00000027598	BC035973	No
	3	ENSPTRT00000038090	NM_016018	No
	4	ENSPTRT00000031258	AK092379	Yes
	5	ENSPTRT00000003028	NM_001460	Yes
	6	ENSPTRT00000059742	NM_032129	Yes
	7	ENSPTRT00000046316	B075864	No
	Simple ASEs	8	ENSPTRT00000046547	NM_012341
9		ENSPTRT00000058464	NM_001018095	No
10		ENSPTRT00000040955	NM_018977	Yes
11		ENSPTRT00000044763	AF263742	Yes
12		ENSPTRT00000004006	AK125187	Yes
13		ENSPTRT00000026230	CR456364	Yes
14		ENSPTRT00000012576	NM_001038640	No
15		ENSPTRT00000055857	NM_020420	No
16		ENSPTRT00000046791	AF332229	No
17		ENSPTRT00000017564	NM_007168	No
Complex ASEs	18	ENSPTRT00000046791	AF332229	No
	19	ENSPTRT00000049503	NM_174961	No
	20	ENSPTRT00000015425	AK054645	No
	21	ENSPTRT00000033416	NM_015245	No
	22	ENSPTRT00000043329	NM_004406	No
	23	ENSPTRT00000043329	NM_004406	Yes
	24	ENSPTRT00000046026	NM_017579	Yes
	25	ENSPTRT00000023338	NM_021007	Yes
	26	ENSPTRT00000055399	BC043599	Yes
	27	ENSPTRT00000033416	NM_015245	No
	28	ENSPTRT00000044186	AF146698	Yes

Supplementary Figure 3. RT-PCR verification results of CENTP exons supported by Ensembl chimpanzee genes. Lane 1: 100bp-DNA marker; Lanes 2~7: retained introns; Lanes 8~17: simple cassette exons; Lanes 18~29: complex cassette exons. The stars indicate the location of the novel ASVs.



GO categories ^a	Observed	Expected	Observed/Expected	<i>P</i> ^b	Corrected <i>P</i> ^c
Molecular Function					
binding	20	21.80	0.92	0.389	NS
catalytic activity	9	12.13	0.74	0.211	NS
signal transducer activity	3	7.12	0.42	0.090	NS
structural molecule activity	2	1.31	1.53	0.517	NS
transcription regulator activity	7	4.69	1.49	0.253	NS
translation regulator activity	12	1.46	8.24	0.001	<0.05
transporter activity	6	12.66	0.47	0.047	NS
Biological Process					
cellular process	15	23.51	0.64	0.015	NS
development	9	13.92	0.65	0.144	NS
physiological process	14	23.15	0.60	0.007	NS
regulation of biological process	13	19.77	0.66	0.059	NS
viral life cycle	7	5.13	1.36	0.374	NS

^aOne transcript might have multiple GO assignments.

^b*P* values were determined by the two-tailed Fisher's exact test.

^cBonferroni corrected for 12 tests. "NS" designated "not significant".

Supplementary Figure 4. Gene ontology analysis of the human transcripts that

contain the potential absent-in-chimpanzee exons identified by CENTP. Here we compare the Gene Ontology (GO) classifications (Gene Ontology Consortium 2001) of the transcripts that contain the identified absent-in-chimpanzee exons with the genome-wide average. Three important trends are observed. Firstly, the overall distribution profile of these transcripts follows that of UCSC-annotated transcripts. This observation implies that most of these changes may have occurred by chance. Nevertheless, such “accidental” changes may still be evolutionarily important given that they have the potential of altering the structures and/or functions of the affected proteins, leading to functional divergence between human and chimpanzee. Secondly, in the majority of GO categories, these transcripts are underrepresented as compared with the genome-wide average. In other words, a significant portion of these transcripts (~ 40%) has no GO annotations, indicating that their functions remain unexplored. This observation appears to echo the hypothesis that the key of speciation often lies in non-essential genes (Wu and Ting 2004). Thirdly, the absent-in-chimpanzee exons are overrepresented in three functional categories – translation regulatory activity, transcription regulatory activity, and viral life cycle. Particularly noticeable is the strongly significant difference in “translation regulatory activity” (P -value < 0.01 by the two-tailed Fisher’s exact test, see Supplementary Fig. 4). After Bonferroni corrections for multiple hypothesis testing, this difference remains significant (P -value < 0.05). There are two possible reasons for the enrichment (which are not mutually exclusive). Firstly, since the relative abundance of PPGs is directly related to the expression levels of their parent genes, it is likely that such enrichment actually results from the high expression levels of the parent genes in these functional categories. Secondly, the enrichment may indicate remarkable disparity in ASV compositions in the related functional categories, and possibly functional divergence between human and chimpanzee. In other words, the related exons/ASVs may have contributed to *Homo-Pan* divergences at both RNA and protein levels, and been associated with the differences in immune response against viral infections between these two species. The latter implication is of particular interest considering that human and chimpanzee differ greatly in susceptibility to such virus-infected disease as AIDS (Dash et al. 2001; Shimizu et al. 1998) and hepatitis B/C (Muchmore et al. 1988). Exploring functional impacts of these newly identified exons/ASVs may be fruitful for understanding hominoid evolution and developing therapies for human viral diseases.

Dash, S., Kalkeri, G., McClure, H.M., Garry, R.F., Clejan, S., Thung, S.N., and Murthy,

- K.K. 2001. Transmission of HCV to a chimpanzee using virus particles produced in an RNA-transfected HepG2 cell culture. *J Med Virol* **65**: 276-281.
- Gene Ontology Consortium. 2001. Creating the gene ontology resource: design and implementation. *Genome Res* **11**: 1425-1433.
- Muchmore, E., Popper, H., Peterson, D.A., Miller, M.F., and Lieberman, H.M. 1988. Non-A, non-B hepatitis-related hepatocellular carcinoma in a chimpanzee. *J Med Primatol* **17**: 235-246.
- Shimizu, Y.K., Igarashi, H., Kiyohara, T., Shapiro, M., Wong, D.C., Purcell, R.H., and Yoshikura, H. 1998. Infection of a chimpanzee with hepatitis C virus grown in cell culture. *J Gen Virol* **79** (Pt 6): 1383-1386.
- Wu, C.I. and Ting, C.T. 2004. Genes and speciation. *Nat Rev Genet* **5**: 114-122.

Supplementary Table 1. RT-PCR-sequencing validation results (number of experimentally confirmed exons/number of tested exons) of the potentially novel exons identified by CENTP.

	CENTP _{C_PPG}	CENTP _{H_PPG}	CENTP _{C_gene}	Summary
<u>Cassette-on</u>				
Simple exon	8/13	7/12	4/10	19/35
Complex exon	4/6	5/12	6/12	15/30
Subtotal	12/19	12/24	10/22	34/65
<u>Retained intron</u>				
Retained intron	1/2	2/5	3/6	6/13
Total Verified	13/21	14/29	13/28	40/78

Supplementary Table 2. Substitution rates between the human/chimpanzee PPGs and their parent genes.

	Chr No.	Exon start	Exon stop	Substitution Rate (%)	Is the CENTP exon found/identified in the macaque/ mouse orthologous Gene?
<u>Chimpanzee PPG ID (Set ①)</u>					
ENSPTRP00000044280.Chimp.chr4.mb174	11	4854558	4854528	11.67	Yes (chimpanzee-specific loss)
ENSPTRP00000027558.Chimp.chr14.mb31	13	31949651	31950984	2.66	Yes (chimpanzee-specific loss)
ENSPTRP00000031135.Chimp.chr2.mb76*	13	96334463	96334441	7.23	Yes (chimpanzee-specific loss)
ENSPTRP00000031135.Chimp.chr2.mb76*	13	96334947	96334924	7.23	Yes (chimpanzee-specific loss)
ENSPTRP00000019599.Chimp.chr2.mb54	19	45231721	45231692	16.91	Yes (chimpanzee-specific loss)
ENSPTRP00000037527.Chimp.chr5.mb121	X	52677731	52677753	22.58	Yes (chimpanzee-specific loss)
ENSPTRP00000037540.Chimp.chr11.mb31	X	135788948	135788842	7.91	Yes (chimpanzee-specific loss)
ENSPTRP00000028924.Chimp.chr3.mb121	Y	13536543	13536579	8.66	Yes (chimpanzee-specific loss)
<u>Chimpanzee PPG ID (Set ②)</u>					
ENSPTRP00000046373.Chimp.chr21.mb33	2	32475547	32475529	22.71	No (chimpanzee loss)
ENSPTRP00000006263.Chimp.chr9.mb49	11	48411211	48411137	17.05	No (chimpanzee loss)
ENSPTRP00000008506.Chimp.chr5.mb167	12	51169452	51169444	5.86	Yes (chimpanzee-specific loss)
ENSPTRP00000008506.Chimp.chr10.mb68	12	51169743	51169737	6.88	No (chimpanzee loss)
ENSPTRP00000008507.Chimp.chr5.mb87	17	37032889	37032863	7.54	Yes (chimpanzee-specific loss)
ENSPTRP00000037527.Chimp.chr5.mb121	X	52677731	52677753	22.58	Yes (chimpanzee-specific loss)
ENSPTRP00000017168.Chimp.chr9.mb84	X	70666108	70666227	9.32	Yes (chimpanzee-specific loss)
ENSPTRP00000023017.Chimp.chr2.mb27	X	149905018	149905029	8.73	Yes (chimpanzee-specific loss)
<u>Human PPG ID (Set ③)</u>					
ENSP00000288960.Human.chr20.mb31	2	32475547	32475529	13.68	No (chimpanzee loss)
Q6IFL4.Human.chr11.mb48	11	48411211	48411137	18.24	No (chimpanzee loss)
Q6P4C7.Human.chr6.mb162	12	51169452	51169444	9.39	Yes (chimpanzee-specific loss)
Q6P4C7.Human.chr12.mb60	12	51169743	51169737	11.21	No (chimpanzee loss)
ENSP00000228652.Human.chr6.mb86	17	37032889	37032863	4.32	Yes (chimpanzee-specific loss)
SSX5_HUMAN.Human.chr6.mb119	X	52677731	52677753	19.88	Yes (chimpanzee-specific loss)
Q8WUM3.Human.chr11.mb83	X	70666108	70666227	7.23	Yes (chimpanzee-specific loss)
HMG4L_HUMAN.Human.chr3.mb26	X	149905018	149905029	10.95	Yes (chimpanzee-specific loss)

* The transcript contains two CENTP exons.

Supplementary Table 3. Substitution rates between the human PPGs and their parent genes. Note that some of the identified transcripts contain more than one CENTP exons. The substitution rates $\leq 2.6\%$ are shaded.

Human PPG ID in Sets ③	Chr. No.	Exon start	Exon stop	Human Original Transcript ID	Substitution rate (%)	Is the CENTP exon
						found/identified in the macaque/mouse orthologous Gene?
KI2S2_HUMAN.Human.chr19.mb59	19	59973205	59973276	ENST00000344867	1.7	Yes (Chimp-specific loss)
Q9H8Q9.Human.chr15.mb51	15	51993444	51993866	ENST00000294993	12.7	Yes (Chimp-specific loss)
ENSP00000336566.Human.chr3.mb33	3	10328912	10328883	ENST00000337354	13.4	Yes (Chimp-specific loss)
ZNF91_HUMAN.Human.chr19.mb58	7	64075639	64075277	ENST00000300619	14.3	Yes (Chimp-specific loss)
Q96T57.Human.chr15.mb80	15	80552490	80551879	ENST00000313380	14.3	Yes (Chimp-specific loss)
Q96T57.Human.chr15.mb80	15	80771498	80772109	ENST00000313380	14.3	Yes (Chimp-specific loss)
Q96T57.Human.chr15.mb80	15	80939074	80938463	ENST00000313380	19.8	Yes (Chimp-specific loss)
ENSP00000334575.Human.chr15.mb72	9	130059957	130060307	ENST00000339078	0.4	No
ENSP00000334575.Human.chr15.mb72	10	30694142	30694379	ENST00000339078	0.4	No
ENSP00000334575.Human.chr15.mb73	9	130060307	130059957	ENST00000339078	0.4	No
ENSP00000334575.Human.chr15.mb73	10	30694509	30694142	ENST00000339078	0.4	No
ENSP00000334575.Human.chr15.mb73	15	60323244	60323002	ENST00000339078	0.4	No
ENSP00000339388.Human.chr15.mb82	10	30694427	30694131	ENST00000338706	0.5	No
ENSP00000339388.Human.chr15.mb82	15	83553818	83553480	ENST00000338706	0.5	No
ENSP00000334575.Human.chr15.mb75	9	130059958	130060307	ENST00000339078	0.5	No
ENSP00000334575.Human.chr15.mb75	10	30694152	30694509	ENST00000339078	0.5	No
ENSP00000312753.Human.chr12.mb110	12	110284889	110285509	ENST00000321753	0.7	No
ENSP00000328595.Human.chr16.mb0	9	6954	7499	ENST00000331147	1.5	No
KI2S2_HUMAN.Human.chr19.mb59	19	59973427	59973471	ENST00000344867	1.7	No
KI2S2_HUMAN.Human.chr19.mb59	19	59973500	59973517	ENST00000344867	1.7	No
KI2S2_HUMAN.Human.chr19.mb59	19	59983880	59983903	ENST00000344867	1.7	No
Q8N7T0.Human.chr7.mb34	7	102707520	102707835	ENST00000320297	2.6	No
WDR5_HUMAN.Human.chr4.mb39	4	39158500	39159478	ENST00000358625	5.9	No
OFD1_HUMAN.Human.chr5.mb37	Y	19203992	19203921	ENST00000326325	5.9	No
OFD1_HUMAN.Human.chr5.mb37	Y	19358959	19359030	ENST00000326325	5.9	No
ENSP00000293618.Human.chr15.mb34	12	49153950	49153974	ENST00000293618	6.8	No
KR106_HUMAN.Human.chr21.mb44	21	44844965	44846016	ENST00000334670	7.8	No
ENSP00000339575.Human.chr7.mb62	7	62858278	62859018	ENST00000340897	8.0	No
ENSP00000339575.Human.chr7.mb62	7	62859957	62861178	ENST00000340897	8.0	No
Q9H8Q9.Human.chr17.mb55	15	51993837	51993584	ENST00000294993	8.8	No
Q9H8Q9.Human.chr18.mb55	15	51993577	51993749	ENST00000294993	10.5	No

ENSP00000349032.Human.chr19.mb62	1	3534398	3534415	ENST00000356620	13.3	No
ENSP00000266116.Human.chr19.mb58	7	64075823	64075784	ENST00000377618	13.6	No
ENSP00000335042.Human.chr7.mb64	7	64075606	64075457	ENST00000340708	13.7	No
ENSP00000335042.Human.chr7.mb64	7	64075807	64075692	ENST00000340708	13.7	No
ENSP00000351114.Human.chr13.mb44	9	68470493	68470383	ENST00000324819	13.8	No
ENSP00000339073.Human.chr1.mb238	17	24074072	24074083	ENST00000355731	14.1	No
ENSP00000351114.Human.chr13.mb63	9	70128613	70128660	ENST00000324819	16.1	No
ENSP00000351114.Human.chr13.mb63	9	70129274	70129368	ENST00000324819	17.0	No
ENSP00000351114.Human.chr13.mb63	9	70129450	70129526	ENST00000324819	17.0	No
ENSP00000351114.Human.chr7.mb149	9	68470258	68470209	ENST00000324819	17.0	No
ENSP00000351114.Human.chr7.mb149	9	68470622	68470602	ENST00000324819	17.2	No
ENSP00000351114.Human.chr7.mb149	9	70128582	70128602	ENST00000324819	17.2	No
ENSP00000351114.Human.chr7.mb149	9	70129260	70129361	ENST00000324819	17.2	No
ENSP00000351114.Human.chr7.mb149	9	70129428	70129532	ENST00000324819	17.2	No
Q9NYV6.Human.chr2.mb143	16	15085651	15085640	ENST00000198767	17.2	No
ZN492_HUMAN.Human.chr12.mb56	1	225901771	225901810	ENST00000355095	18.3	No
ZN492_HUMAN.Human.chr12.mb56	19	23720110	23720087	ENST00000355095	18.4	No
ENSP00000351114.Human.chr9.mb127	9	68470061	68469998	ENST00000324819	18.4	No
ENSP00000351114.Human.chr9.mb127	9	70128711	70128809	ENST00000324819	19.6	No
ENSP00000351114.Human.chr9.mb127	9	70129434	70129526	ENST00000324819	19.6	No
ZNF85_HUMAN.Human.chr19.mb20	7	64075387	64075361	ENST00000328178	19.6	No
ENSP00000351689.Human.chr19.mb23	19	23627570	23629094	ENST00000357491	22.9	No

Supplementary Table 4. Gene ontology analysis of the human transcripts that contain the CENTP exons identified by chimpanzee transcripts. The overrepresented subcategories, i.e., translation regulatory activity and viral life cycle, are shaded (both P -values < 0.01).

GO categories ^a	Observed	Expected	Observed/ Expected	P^b	Corrected P^c
Molecular Function					
antioxidant activity	7	2.02	3.47	0.089	NS
signal transducer activity	70	57.47	1.22	0.123	NS
structural molecule activity	7	10.54	0.66	0.238	NS
transcription regulator activity	42	37.85	1.11	0.318	NS
translation regulator activity	31	11.75	2.64	0.001	<0.05
transporter activity	69	102.16	0.68	0.002	<0.05
binding	159	175.83	0.90	0.125	NS
catalytic activity	79	97.87	0.81	0.062	NS
enzyme regulator activity	9	13.76	0.65	0.199	NS
motor activity	2	2.12	0.94	1.000	NS
Biological Process					
behavior	13	19.05	0.68	0.143	NS
cellular process	185	189.68	0.98	0.388	NS
development	128	112.33	1.14	0.123	NS
physiological process	177	186.72	0.95	0.260	NS
regulation of biological process	159	159.46	1.00	0.500	NS
viral life cycle	66	41.39	1.59	0.006	NS

^aOne transcript might have multiple GO assignments.

^b P values were determined by the two-tailed Fisher's exact test.

^cBonferroni corrected for 16 tests. "NS" designated "not significant".