Supporting Information

Gibbons et al. 10.1073/pnas.1416878112

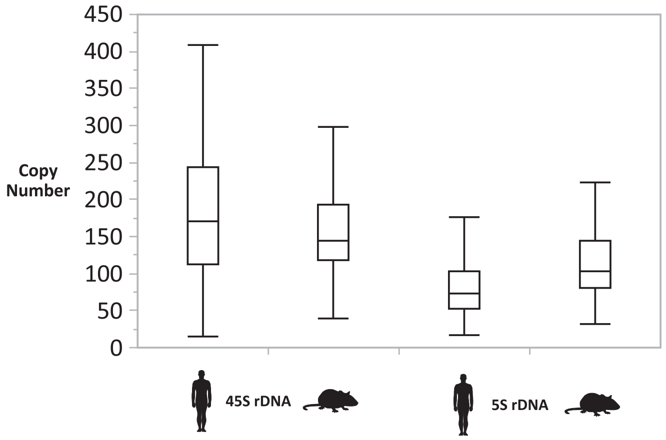

Fig. S1. 5S and 45S rDNA CN display extensive variability. Box plots of 5S and 45S rDNA CN in humans and mice. Box borders represent first and third quartiles, the horizontal line inside the box is the median, whiskers extend to one and one half times the interquartile range, and points represent outliers. The *y* axis reports haploid CN.

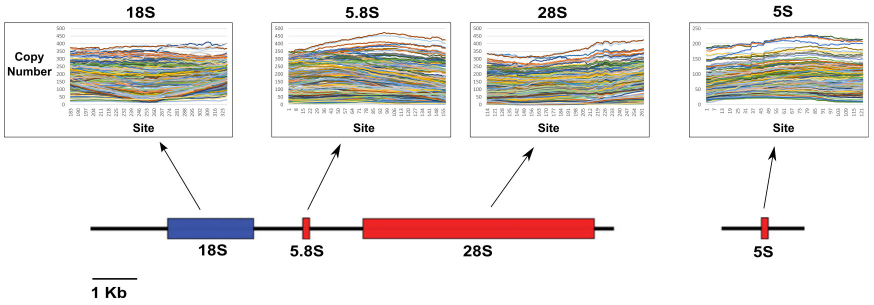

Fig. S2. rDNA CN estimates display low variance along selected segments in each rDNA component. CN estimates (*y* axis) per site (*x* axis) for segments of each 45S rDNA component (*A*–*C*) for the 168 human individuals. Segments exhibit low variance in CN estimates across sites and no nucleotide variation between human and mouse genomes (*Materials and Methods*). The nucleotide sites in each segment are reported below the line graphs.

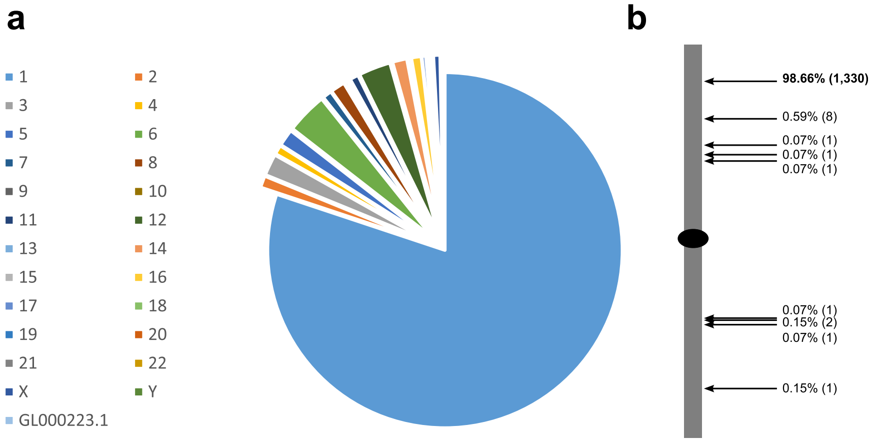

Fig. S3. 5S rDNA reads map to the expect location at 1q42. Display data for 5S rDNA read mapping breakdown from a single randomly selected individual (NA11193). (*A*) Pie chart shows the distribution of 5S rDNA read mapping across the entire reference human genome and supercontigs. (*B*) Display of the location of the 5S reads that map to chromosome 1. Arrows show chromosomal region, numbers represent the percentage of chromosome 1 mapped reads, and raw counts are in given in parentheses On average, fewer than 10% of the reads that map to the 5S rDNA in each individual also map to other locations outside of 1q42.

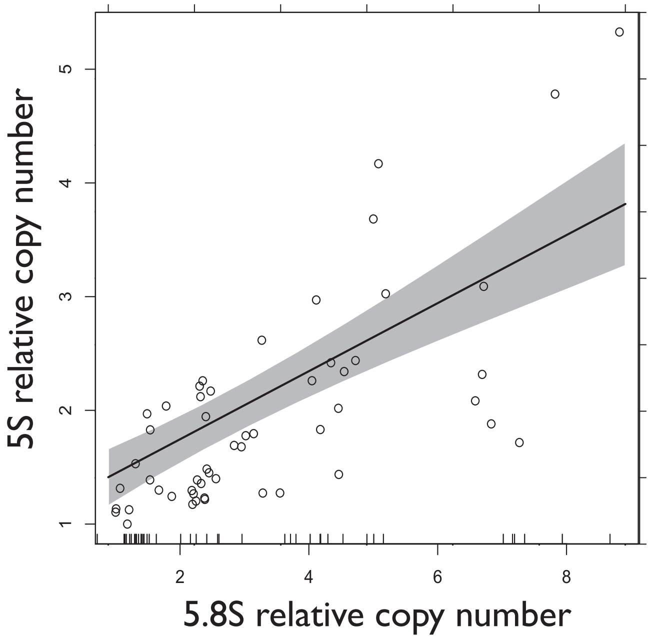

Fig. S4. Experimental determination of cCNV across individuals in human populations. Linear association between CN of 5S rDNA locus (*y* axis) and 5.8S rDNA locus (used as proxy for 45S rDNA locus; *x* axis). Estimates of rDNA CN were obtained with qPCR (*Materials and Methods*; *n* = 54 individuals). The gray shaded area represents the 95% CIs. Estimates of CN were divided by the value of the individual with the smallest CN (relative CN). Spearman correlation ρ is 0.71 ($P = 4.58 \times 10^{-9}$) across all individuals.

Table S1. Spearman correlations and associated *P* values between 5S rDNA CN and 45S rDNA CN

Copy	Combined (*n* = 168)		CEU (*n* = 90)		YRI (*n* = 78)		Male (*n* = 77)		Female (*n* = 91)	
Human	ρ	*P* value	ρ	*P* value	ρ	*P* value	ρ	*P* value	ρ	*P* value
18S	0.61	2.00E-18	0.54	8.80E-08	0.68	4.00E-12	0.55	3.00E-07	0.62	3.00E-11
5.8S	0.80	3.00E-38	0.73	1.00E-15	0.83	8.00E-22	0.77	3.00E-16	0.79	3.00E-21
28S	0.73	4.00E-29	0.67	1.00E-12	0.79	3.00E-18	0.70	2.00E-12	0.73	2.00E-16
45S rDNA	0.74	1.00E-30	0.68	7.00E-13	0.80	6.00E-19	0.70	2.00E-12	0.74	2.00E-17
	Combined (*n* = 27)		Lab (*n* = 17)		Wild (*n* = 10)					
Mouse	ρ	*P* value	ρ	*P* value	ρ	*P* value				
18S	0.64	0.0004	0.50	0.042	0.52	0.128				
5.8S	0.78	1.30E-06	0.60	0.011	0.83	0.003				
28S	0.73	1.50E-05	0.57	0.017	0.44	0.200				
45S rDNA	0.77	2.40E-06	0.60	0.012	0.65	0.043				

Data are shown for the correlation between estimates of CN for the 5S rDNA locus vs. individual components of the 45S locus (18S, 5.8S, and 28S), and the entire 45S rDNA locus (average of 18S, 5.8S and 28S; *Methods*).

Table S2. Population samples used for the experimental evaluation of cCNV

Individuals from CEU origin (*n* = 15)	Individuals from Autism collection (*n* = 12)	Families (*n* = 3)	Ethnically diverse individuals (*n* = 29)
NA12872, NA12812, NA12814, NA12234, NA12873, NA12874, NA12875, NA12827, NA12828, NA12829, NA12830, NA12763, NA12716, NA12751, NA12043.	NA10021, NA10022, NA10023, NA10033, NA10038, NA10039, NA10139, NA10140, NA10143, NA10144, NA10145, NA10149.	Family 1447, Family 1456, Family 1459	12ZF6, SBM3X, YCCZ6, WLB37, ZIUU4, 2KGZW, 15WN4, NQSUI, KXZ6O, XDMNS, 7PYWG, 8TDIN, MWMLL, 1AL7A, 6MC7I, K2NRJ, 61KQS, XLFF1, Q3AE3, S8SNY, X79QS, K5L23, XH8QD, 2N42J, BNPO5, UMYA7, QYQPK, YAW38

Further information on population samples and experimental evaluation of cCNV is provided in *Methods*.

Table S3. Sequence of primers used for qPCR analysis

Gene	Forward primer	Reverse primer
TP53	5′TGTCCTTCCTGGAGCGATCT3′	5′CAAACCCCTGGTTTAGCACTTC3′
5S rDNA	5′TCGTCTGATCTCGGAAGCTAA3′	5′AAGCCTACAGCACCCGGTAT3′
5.8S rDNA	5′CGACTCTTAGCGGTGGATCA3′	5′GATCAATGTGTCCTGCAATTC3′
18S rDNA	5′GACTCAACACGGGAAACCTC3′	5′AGACAAATCGCTCCACCAAC3′
28S rDNA	5′GCGGGTGGTAAACTCCATCT3′	5′CACGCCCTCTTGAACTCTCT3′

Further information on qPCR is provided in *Methods*.