

Figure 1

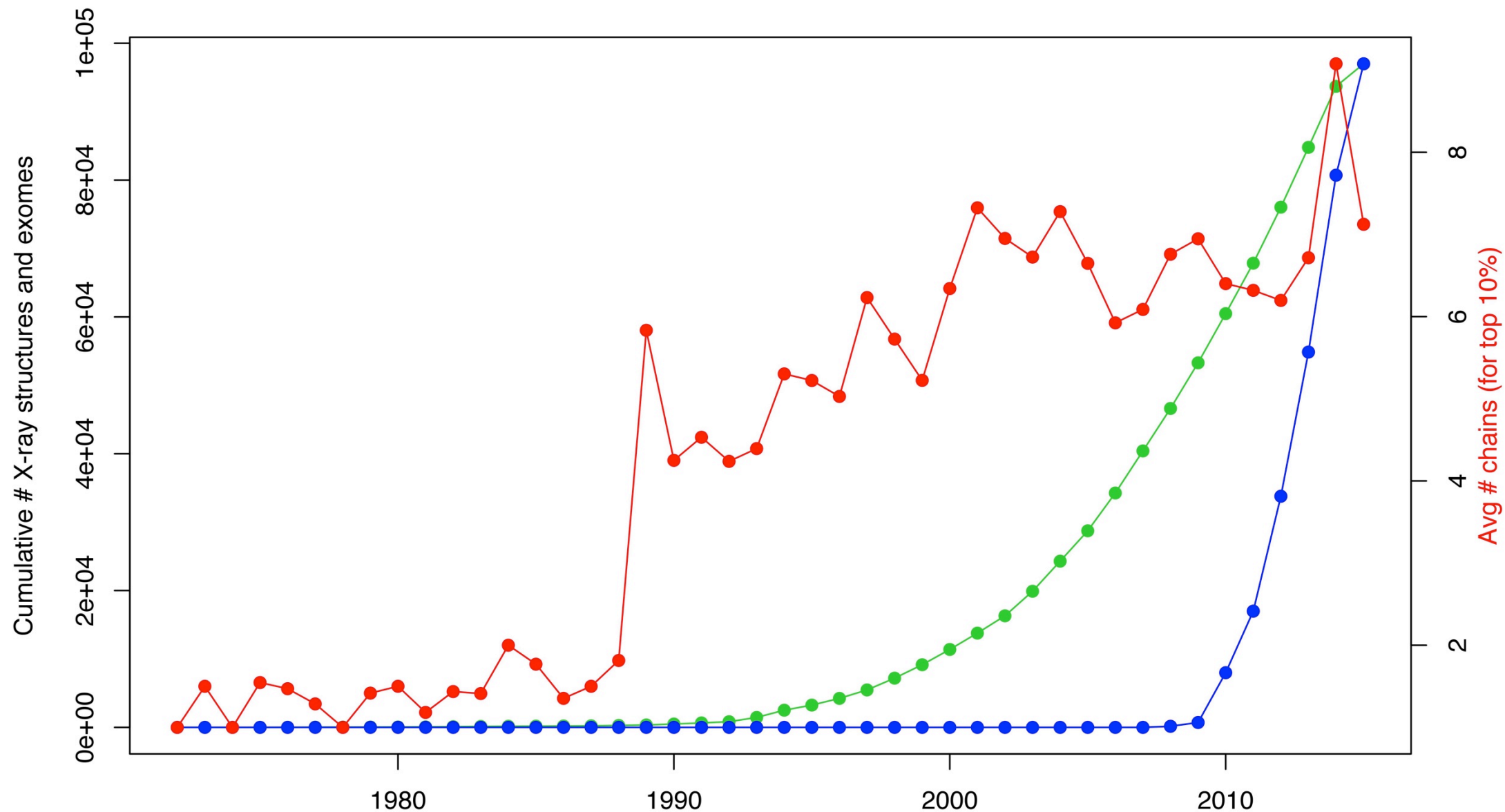
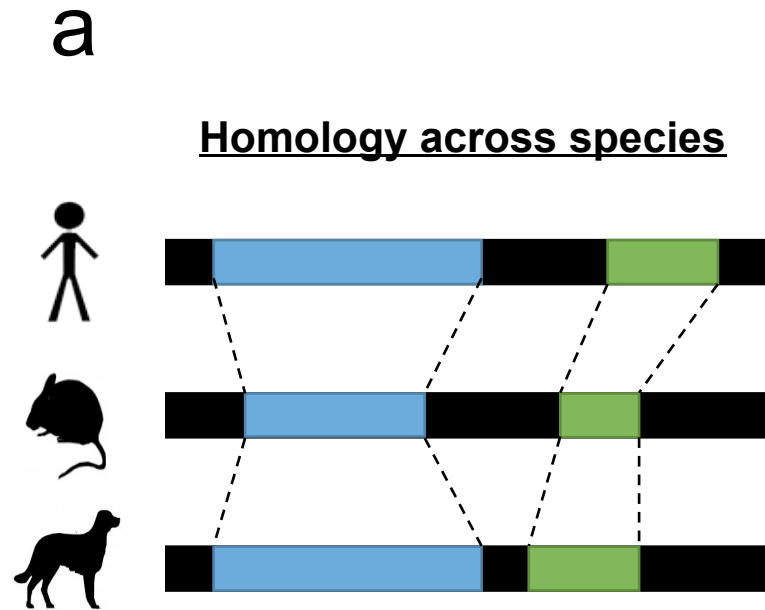


Figure 2

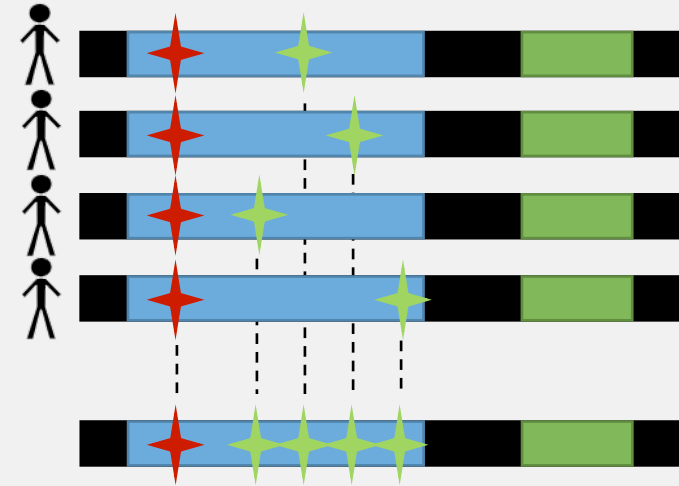


**Conservation**

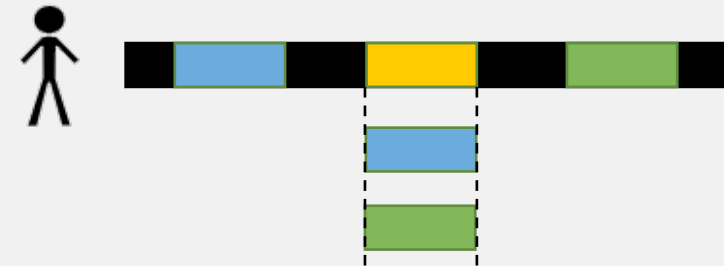
- Typically defined by comparison across species

**b**

(1) Homology within human population



(2) Homology within a single human genome



**Human- or domain-specific 'conservation'**

- Comparison between homologous sequences within or across humans

	SYNONYMOUS				NON-SYNONYMOUS				ALL SNVs			
	DAF <0.5%	DAF 0.5- 5%	DAF >5%	Total	DAF <0.5%	DAF 0.5- 5%	DAF >5%	Total	DAF <0.5%	DAF 0.5-5%	DAF >5%	Total
<b>AVERAGE</b>	<b>277</b>	<b>961</b>	<b>13,017</b>	<b>14,255</b>	<b>406</b>	<b>1022</b>	<b>10,932</b>	<b>12,360</b>	<b>57,915</b>	<b>253,982</b>	<b>3,803,457</b>	<b>4,115,354</b>
ASW	506	2,005	12,650	15,161	645	1,954	10,496	13,095	116,216	558,996	3,728,104	4,403,316
LWK	642	2,359	12,309	15,310	806	2,278	10,195	13,279	152,267	649,303	3,619,039	4,420,609
YRI	547	2,468	12,190	15,205	697	2,377	10,056	13,130	122,978	683,239	3,592,643	4,398,860
CEU	175	594	13,237	14,006	298	709	11,173	12,180	30,034	143,987	3,854,212	4,028,233
FIN	166	594	13,243	14,003	279	717	11,170	12,166	32,918	146,737	3,874,781	4,054,436
GBR	167	591	13,216	13,974	280	711	11,183	12,174	31,944	146,802	3,863,507	4,042,253
IBS	139	586	13,196	13,921	224	715	11,130	12,069	36,129	155,270	3,855,295	4,046,694
TSI	203	592	13,232	14,027	331	703	11,198	12,232	33,425	146,066	3,856,263	4,035,754
CHB	218	497	13,077	13,792	355	563	11,026	11,944	36,081	119,668	3,811,719	3,967,468
CHS	226	478	13,058	13,762	365	540	11,008	11,913	40,974	118,137	3,803,199	3,962,310
JPT	240	500	13,067	13,807	387	571	11,012	11,970	45,107	120,115	3,796,180	3,961,402
CLM	215	740	13,284	14,239	340	833	11,171	12,344	45,184	189,257	3,869,561	4,104,002
MXL	194	641	13,246	14,081	320	740	11,104	12,164	38,654	163,737	3,856,439	4,058,830
PUR	237	810	13,232	14,279	357	892	11,124	12,373	48,892	214,430	3,867,461	4,130,783

*Table 1. Per individual whole genome SNV load in the 1000 Genomes Project Phase 1 data<sup>2</sup>. The number of synonymous, non-synonymous and all SNVs (including non-coding SNVs) are categorized by the derived allele frequency (DAF; defined as the allele alternative to the ancestral allele). DAF<0.5% are considered 'rare' and DAF>5% are considered 'common'. Ancestry legend: ASW: Americans of African ancestry in Southwestern USA; LWK: Luhya in Webuye, Kenya; YRI: Yoruba in Ibadan, Nigeria; CEU: Utah Residents (CEPH) with Northwestern European ancestry; FIN: Finnish in Finland; GBR: British in England and Scotland; IBS: Iberian population in Spain; TSI: Toscani in Italy; CHB: Han Chinese in Beijing; CHS: Southern Han Chinese; JPT: Japanese in Tokyo, Japan; CLM: Columbians from Medellin, Columbia; MXL: Mexican Ancestry from Los Angeles, USA; PUR: Puerto Ricans from Puerto Rico.*

Figure 3

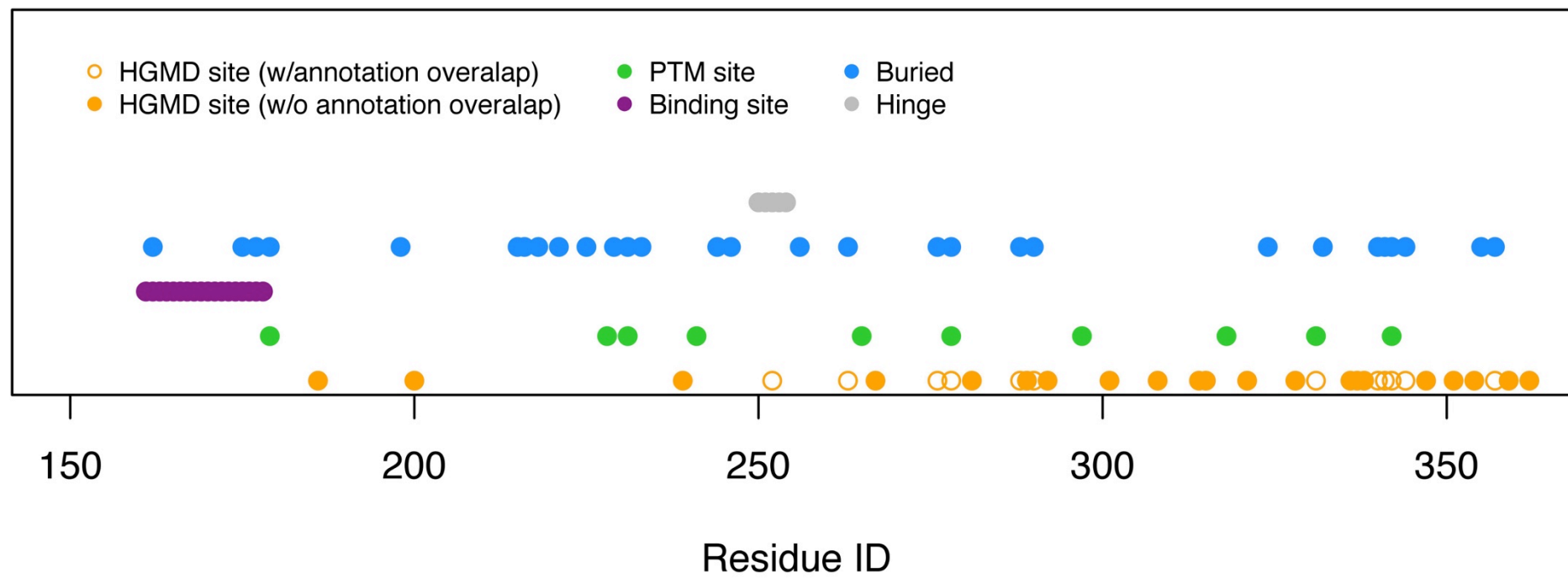
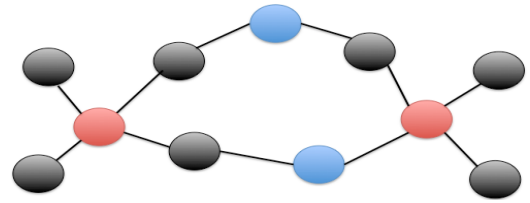
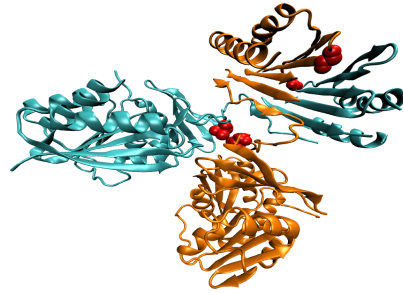


Figure 4

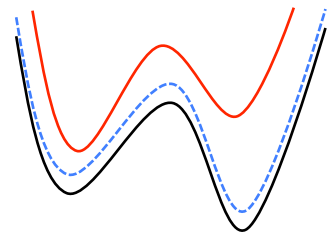
```
ATGCCGTAATTCGTTAATCGTCG
ATGCCGAAAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
ATGCCGTAATTCGTTAATCGTCG
```



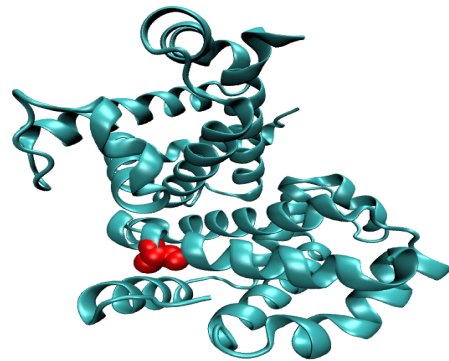
Impact on PPI topology



Impact on binding interactions



Ensemble level changes



Global changes



Local transitions