

file	contents	length	σ
<i>SOURCES</i>	C/Java source code.	210,866,607	230
<i>PITCHES</i>	MIDI Pitch values.	55,832,855	133
<i>PROTEINS</i>	Protein sequences.	1,184,051,855	24
<i>DNA</i>	DNA sequences.	403,927,746	16
<i>ENGLISH</i>	English texts.	1,073,741,824	226
<i>XML</i>	XML files.	296,135,874	97