Table I. Statistical comparison of the best performing 12 groups and the NAÏVE_consensus method in QA1.1 mode (global quality estimates assessed on a per-target basis)

		312	**359**	**371**	**000**	**407**	**2**	**386**	**426**	**369**	**319**	**397**	**78**	**490**
MUFOLD-WQA	**312**	X	116	116	116	116	116	113	116	116	116	116	116	113
MUFOLD-QA	**359**	0.98	X	116	116	116	116	113	116	116	116	116	116	113
QMEANclust	**371**	0.86	0.70	X	116	116	116	113	116	116	116	116	116	113
NAÏVE-consensus	**000**	0.36	0.09	<0.01	X	116	116	113	116	116	116	116	116	113
United3D	**407**	0.50	0.33	0.13	0.74	X	116	113	116	116	116	116	116	113
Multicom-cluster	**2**	0.30	0.06	<0.01	0.44	0.61	X	113	116	116	116	116	116	113
Mufold	**386**	0.24	<0.01	<0.01	0.19	0.29	0.33	X	113	113	113	113	113	113
MetaMQAPclust	**426**	0.18	0.07	0.07	0.37	0.37	0.42	0.46	X	116	116	116	116	113
MQAPmulti	**369**	0.06	0.06	0.04	0.30	0.26	0.33	0.41	0.97	X	116	116	116	113
Pcons	**319**	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.22	0.18	X	116	116	113
ModFOLDclust2	**397**	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.22	X	116	113
IntFOLD-QA	**78**	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.15	<0.01	X	113
MULTICOM	**490**	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.14	0.42	0.53	X

Results of the two-tailed paired t-tests on Pearson's correlation coefficients. The upper right part of the table contains the numbers of common targets predicted. The lower part displays the probability that the observed differences are due to chance. Shaded cells highlight pairs of statistically indistinguishable groups at the 10^{-2} significance level.