



HAL
open science

Imputation of low frequency variants is using the HapMap3 benefits from large, diverse reference sets

Luke Jostins, Katherine Morley, Jeffrey C Barrett

► To cite this version:

Luke Jostins, Katherine Morley, Jeffrey C Barrett. Imputation of low frequency variants is using the HapMap3 benefits from large, diverse reference sets. *European Journal of Human Genetics*, 2011, 10.1038/ejhg.2011.10 . hal-00618484

HAL Id: hal-00618484

<https://hal.science/hal-00618484>

Submitted on 2 Sep 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Imputation of low frequency variants using the HapMap3 benefits from large, diverse reference sets

Short Title: Large, diverse imputation reference sets

Luke Jostins,¹ Katherine I. Morley,^{1,2} Jeffrey C. Barrett^{1*}

1/6/2011

¹ Statistical and Computational Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1HH, UK

² Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health, The University of Melbourne, Melbourne, VIC 4010, Australia

* Corresponding Author: barrett@sanger.ac.uk, Tel: +44 (0)1223 492351, Fax: +44 (0)1223 496826

Keywords: Imputation, Reference Sets, Rare Variants

1 **Abstract**

2
3 Imputation allows the inference of unobserved genotypes in low-density
4 datasets, and is often used to test for disease association at variants that are
5 poorly captured by standard genotyping chips (such as low frequency variants).
6 While much effort has gone into developing the best imputation algorithms,
7 less is known about the effects of reference set choice on imputation accuracy.
8 We assess the improvements afforded by increases in reference size and
9 diversity, specifically comparing the HapMap2 dataset that has been used to
10 date for imputation, and the new HapMap3 dataset, which contains more
11 samples from a more diverse range of populations. We find that, for imputation
12 into Western European samples, the HapMap3 reference provides more
13 accurate imputation with better calibrated quality scores than HapMap2, and
14 that increasing the number of HapMap3 populations included in the reference
15 set grants further improvements. Improvements are most pronounced for low
16 frequency variants (frequency < 5%), with the largest and most diverse
17 reference sets bringing the accuracy of imputation of low frequency variants
18 close to that of common ones. For low frequency variants, reference set
19 diversity can improve the accuracy of imputation independent of reference
20 sample size. HapMap3 reference sets provide significant increases in
21 imputation accuracy relative to HapMap2, and are of particular use if highly
22 accurate imputation of low frequency variants is required. Our results suggest
23 that although the sample sizes from the 1000 Genomes Pilot Project will not
24 allow reliable imputation of low frequency variants, the larger sample sizes of
25 the main project will.

27 **1 Introduction**

28 Genome-wide association studies (GWAS) comparing thousands of disease cases and healthy controls at
29 hundreds of thousands of single nucleotide polymorphisms (SNPs) have led to the recent discovery of
30 hundreds of *bona fide* associations between common SNPs and risk for complex human diseases¹². To
31 add further value, a wide variety of statistical refinements have been applied to maximize the power of
32 these studies. Genotype imputation is one such approach which predicts untyped markers in target (i.e.
33 GWAS) samples using a densely typed reference set (e.g. the HapMap³). Imputation allows meta-
34 analysis of studies genotyped on different commercial SNP chips, and allows association testing of
35 variants which are not in high LD with any single genotyped SNPs, and are thus not well captured by the
36 chips (such as rare mutations⁴).

37 Many recent papers have investigated various factors that influence imputation performance; these
38 include method used^{5,6,7,8}, SNP density in target sample^{5,9}, quality of reference haplotype phasing^{7,8}
39 and settings of method-specific parameters^{7,9}. Many studies have measured how imputation
40 performance increases with reference sample size^{9,8,10}. Other studies have investigated the specific
41 composition of the reference set: Huang et al¹⁰ showed that specific mixtures of HapMap 2 populations
42 gave better performance than any single population when performing imputation in 29 target populations
43 from around the world. These results were reviewed by Li et al¹¹, who recommended a combination of
44 all HapMap2 samples for imputing into samples from certain populations. Similarly, Marchini and
45 Howie⁷ showed that combining all HapMap 2 samples from all populations increased imputation
46 performance for low frequency SNPs. More recently, the HapMap3 dataset was used¹² to show that a
47 mixture of samples from two European populations (CEU and TSI) could give improvements in
48 imputation performance for target samples from Western Europe.

49 Most imputation work to date has used the HapMap2 reference panel³, which comprises 60 unrelated
50 individuals each of European and African origin, and 90 of East Asian origin, genotyped at over 2
51 million sites. While this reference set has been shown to provide highly accurate imputation for nearly all
52 common variation in samples of European origin, an open question remains about how the size (in terms

53 of number of samples and number of SNPs) and quality of new and planned reference datasets will affect
54 imputation. Specifically, the HapMap3¹² reference set contains more samples (over 1000 individuals
55 from 11 sample collections with diverse ancestry) genotyped at a restricted set of approximately 1.5
56 million variants. Conversely, the pilot phase of the 1000 genomes project plans to release genotypes at
57 many millions of novel sites in the relatively small HapMap2 sample set. The full project will sequence
58 nearly all of the HapMap3 samples, as well as a number of samples from other populations, to give a
59 high-density reference set greater in size than the HapMap.

60 To date, no in-depth analysis has been performed to investigate the effect of reference set size and
61 diversity in mixed-population reference sets. The release of the large, diverse HapMap3 dataset allows
62 such an investigation. We perform imputation into European target samples using HapMap 2 and
63 HapMap 3 reference sets of various sizes and population diversities, and measure the difference in
64 imputation accuracy, quality score performance, and computational resources required. We also perform
65 experiments to tease out the effect of reference set size, diversity and closeness of genetic match to the
66 target population. Our comparative analysis focuses on three areas: (1) what effect does the higher
67 quality of genotyping from HapMap3 compared to HapMap2 have on imputation? (2) what
68 improvements can the large increase in sample size and diversity of mixed reference sets have on
69 imputation accuracy and predicted quality scores, especially for low frequency SNPs? and (3) what can
70 we infer about the relationship between imputation performance and closeness of match between the
71 ancestry of reference and target samples?

72

73 **2 Materials and Methods**

74 **2.1 Performing and Scoring Imputation**

75 For the target set, we used 1 374 individuals from the 1958 British Birth Cohort¹³, genotyped on both
76 the Illumina HumanHap550 BeadChip and Affymetrix GeneChip® Human Mapping 500k chips as our
77 target set. We used the Illumina data to perform imputation, and checked the answers using the
78 Affymetrix data (Illumina chips having been previously shown to be more powerful for imputation¹⁴).
79 For the target reference sets, we used the approximately 2.5M polymorphic SNPs of the HapMap2 CEU
80 samples, and various mixtures of HapMap3 samples, with approximately 1.4M polymorphic SNPs
81 (Table 1).

82 To perform the imputation we used the imputation program Beagle⁹⁸. We also tested a subset of our
83 results using IMPUTE v1¹⁵ and IMPUTE v2⁶, and compared the computation requirements of all three
84 programs (Supplemental Table 2). For some of our analyses, we removed poorly-imputed SNP by
85 applying a filter that removed SNPs with a predicted dosage r^2 of less than 0.9. For several analyses we
86 compare common ($MAF > 5\%$) and low frequency ($MAF \leq 5\%$) SNPs.

87 To score the imputation results, we measured both the accuracy of imputation and the usefulness of
88 the predicted quality scores that the imputation method provides. Accuracy was measured using dosage
89 r^2 , which measures the correlation between the actual gene dosages and those predicted by imputation.
90 The dosage r^2 is useful as it is not confounded by minor allele frequency, and thus can be used to
91 compare rare and common SNPs, as well as having a simple relationship to power in a GWAS¹⁴. For
92 predicted quality scores, both Beagle and IMPUTE give a predicted dosage r^2 for each SNP (a prediction
93 of what the dosage r^2 would be for that SNP), which was evaluated using four criteria: (1) the
94 calibration, or mean difference between predicted and actual dosage r^2 (2) the quality r^2 , or the
95 correlation between predicted and actual dosage r^2 , (3) the number of overconfident calls, i.e. the number
96 of SNPs that are poorly imputed despite having high predicted dosage r^2 , and, vice versa, (4) the number

97 of underconfident calls. We are particularly interested in the number of overconfident SNPs, as when
98 genotypes are incorrectly imputed with high confidence, any differential effect of these errors between
99 cases and controls can yield false positive associations. Following up these errors in replication studies
100 can be a costly waste of time.

101 **2.2 Reference Set Quality**

102 While the majority of SNPs in both HapMap2 and HapMap3 are of high quality, HapMap2 data were
103 generated using a variety of genotyping technologies in the period from 2003-2007, some of which were
104 not as robust as the GWAS chips used to generate the HapMap3 data in 2008. To investigate whether
105 this increase in reference set quality had a significant effect on imputation, we performed genome-wide
106 imputation on the target set using two ‘reduced’ HapMap reference sets, and measured differences in
107 dosage r^2 . These reduced sets contained only the 56 CEU samples and 1M SNPs that HapMap2 and
108 HapMap3 have in common.

109 **2.3 Reference Set Size**

110 To assess the effect of larger HapMap sample sizes, we performed genome-wide imputation on the target
111 set, using five reference sets of increasing size and diversity. We used the HapMap2 and HapMap3 CEU
112 samples (HM2CEU and HM3CEU), which should be the best match to the UK target set, as well as a
113 mixed reference set of HapMap3 European samples (CEU+TSI). To give a large, but still partially
114 matched reference set, we used the HapMap3 European samples mixed with the Indian and Mexican
115 samples (CEU+TSI+GIH+MEX), as these populations cluster together on the first two principal
116 components (see Supplemental Figure 2 from ¹²). Finally, we examined all HapMap3 individuals
117 (WORLD), in order to assess a very large and very diverse reference set. Sample sizes are shown in
118 Table 2.

119 **2.4 Reference Set Diversity**

120 We investigated the importance of population matching, independent of sample size, in two ways.
121 Firstly, we compared genome-wide imputation using the HapMap3 CEU+TSI reference set to a
122 CEU+JPT+CHB reference set of the same size and non-CEU proportion. This allows us to investigate
123 the effect of adding poorly matched samples on imputation. Second, we created a number of equally-
124 sized reference sets for chromosome 17 by combining a range of mixture proportions of either CEU and
125 TSI, or CEU and CHB+JPT. We measured the accuracy of imputation using these reference sets for low
126 frequency variants. We denote these constant-sized mixed reference sets as CEU/TSI and
127 CEU/CHB+JPT, in order to distinguish between reference sets in which sample size is not held constant
128 (e.g. CEU+TSI).

129

130 **3 Results**

131 **3.1 Reference Set Quality**

132 We found a small but significant difference due to genotyping quality (unfiltered mean dosage r^2 0.841
133 vs 0.84, Supplemental Figure 1), but not enough to explain a meaningful difference in imputation quality
134 between HapMap2 and HapMap3.

135 **3.2 Reference Set Size**

136 We found that HapMap3 yields a substantial increase in imputation accuracy compared to HapMap2,
137 with the number of SNPs in the highest score category (>95%) increasing, and the number in all lower-
138 scoring categories decreasing (Figure 1). A further increase in imputation accuracy is seen when adding
139 the HapMap3 TSI samples. The number of SNPs that pass the filter (have a predicted r^2 greater than 0.9)
140 rises as imputation accuracy increases, although this falls as samples from many populations are added
141 due to a decrease in the imputation software's predicted confidence (see below). The dosage r^2 of filtered
142 SNPs shows a trend of improved imputation with increasing sample sizes. This increase is statistically
143 significant ($p < 10^{-16}$) for all increases in sample size, with the exception of the WORLD set (Table 2). A
144 corresponding increase is seen in computational time, especially for the WORLD set; however, the
145 CEU+TSI+GIH+MEX reference set only takes 55% longer to process than just CEU, despite being
146 nearly 3 times larger.

147 The improvement for low frequency SNPs is the most striking. The HM2CEU mean dosage r^2 score
148 for unfiltered low frequency SNPs is low, especially compared to common SNPs (0.89 vs 0.96). If all
149 samples from all HapMap3 populations are included, this gap nearly disappears (0.96 vs 0.98). In
150 general, fewer low frequency SNPs pass the imputation quality filter (63% at most), but the accuracy of
151 these imputed low frequency SNPs can become very high. The improvement in dosage r^2 is inversely

152 proportion to the frequency of the SNP, with the greatest improvement observed for the very rarest SNPs
153 (Figure 2).

154 For small reference sets, the calibration of predicted quality scores tends towards overconfidence. As
155 the reference set increases in size, the calibration improves, though very diverse reference sets lead the
156 confidence scores towards underconfidence (Supplemental Table 1). The correlation between predicted
157 and actual dosage r^2 improves, though with a slight decrease for the most diverse sets. These trends are
158 stronger in low frequency variants than in common ones; low frequency variants tend to have less well
159 calibrated and correlated predicted quality scores. Larger reference sets decrease the number of
160 overconfident mistakes and the number of underconfident mistakes (with the exception of the WORLD
161 set, which causes a slight inflation in underconfident calls, Figure 3).

162 **3.3 Reference Set Diversity**

163 We found that, while the mismatched CEU+JPT+CHB reference set gives a lower imputation accuracy
164 than CEU+TSI, it still yielded a substantial improvement over the CEU reference set alone. Half of the
165 improvement in imputation accuracy from CEU to CEU+TSI was also gained with the CEU+JPT+CHB
166 reference. This implies that while matching the reference set to the target set is important, even the
167 addition of unrelated samples yields increases in imputation accuracy.

168 Increased diversity initially correlates with increased imputation accuracy for both CEU/TSI and
169 CEU/CHB+JPT (Figure 4), though the former is far less marked than the latter. Beyond a certain
170 proportion of non-CEU samples accuracy starts to fall off as the effect of diversity is outweighed by the
171 effect of mismatching. The optimum population mix is 22% for CEU/TSI, and 17% for CEU/CHB+JPT.
172 It is only above 43% TSI do we see a decrease in imputation accuracy for adding TSI over pure CEU; for
173 CHB+JPT this figure is 33%. This relationship is specific to low frequency variants.

174

175 **4 Discussion**

176 Higher quality reference data and larger sample sizes yield improved imputation accuracy. Using
177 HapMap3 as a reference set compared to using HapMap2 demonstrates this improvement, especially at
178 sites with a low minor allele frequency. While this result was expected we did not anticipate the
179 substantial improvement achieved with large and genetically diverse reference sets. Including samples
180 from such diverse populations as MEX and GIH can provide significant improvement in imputation into
181 UK samples of alleles with a minor allele frequency of less than 5%. Larger reference sets also improve
182 predicted quality scores, with a decrease in overconfident mistakes without inflating underconfident
183 calls.

184 Overall, an imputation reference set consisting of CEU, TSI, MEX and GIH improves the quality of
185 imputation in all frequency ranges, and greater improvement for very rare SNPs was achieved with very
186 large and highly mixed reference sets. The latter came at the cost of computational power, as well as
187 overly conservative predicted quality scores. Imputation is robust to the precise mix of samples of
188 closely related ancestry (such as CEU/TSI), and small amounts of divergent ancestry can actually
189 improve accuracy (such as CEU/CHB+JPT). However, crude population matching is important, as
190 demonstrated by the reduced accuracy of the CEU+JPT reference compared to CEU+TSI.

191 These results imply a set of relatively simple rules for picking imputation reference sets: for the best
192 trade-off between accuracy and computation time, the most diverse mixture of populations that still
193 approximately cluster with the target samples of interest on a world-wide PCA plot should be used.
194 However, if imputing genotypes for low frequency variants with high accuracy is required, all samples
195 available should be used, with the understanding that this will increase computational time, and cause
196 quality scores to be somewhat conservative.

197 Of the programs we tested, Beagle takes greatest advantage of the highly divergent sample mixes,
198 possibly because IMPUTE v2's only uses haplotypes with small Hamming distance from the target
199 sample during phasing, and thus is less likely to take full advantage of the more divergent haplotypes.
200 However, this is a function of the parameter values chosen: increasing the value of k in IMPUTE v2 will
201 increase the number of haplotypes considered, thus increasing accuracy at the expense of resource use.

202 As IMPUTE v1 always uses all reference haplotypes, it seems likely that it would also be able to take
203 advantage of divergent populations, but its prohibitive resource usage makes it a poor choice for large
204 reference sets.

205 That badly matched reference sets lead to increasingly conservative quality scores is an interesting
206 observation. This effect is observed in Beagle and IMPUTE V1, but not in IMPUTE v2 (Table S2) is
207 more puzzling. This lowering of predicted quality is likely to be due to the poor match of haplotype
208 frequencies in the reference and target sets. As the true haplotypes in the target are likely to be rarer in
209 the reference, this will effectively lower the prior on correctly guessed haplotypes, leading a deflation of
210 the posterior. IMPUTE v2, by only examining haplotypes close to the target sample, will not suffer from
211 this problem.

212 It should be noted that these results were obtained by imputation into European individuals, and
213 further studies will be needed to assess how these conclusions generalize to other populations, notably
214 African populations.

215 Accurate imputation of low frequency SNPs using HapMap3 samples could allow new associations
216 to be mined from existing GWAS datasets. HapMap3 contains nearly 150 000 SNPs with a frequency of
217 less than 5%, a large fraction of which can be accurately imputed. This approach will be even more
218 powerful when applied to the millions of new low frequency variants catalogued by the 1000 Genomes
219 Project. The promise of such analyses must be tempered, however, by the observation that high quality
220 genotypes in hundreds of samples will be required to provide accurate imputation. The HapMap2-like
221 sample sizes of the 1000 Genomes pilot, coupled with less accurate genotypes derived from low
222 coverage sequence may well not be sufficient to allow powerful imputation. However, the diverse and
223 extensive set of samples being sequenced for the final project (including TSI, UK and Finnish samples),
224 coupled with improvement on genotype calls from sequence offer the exciting prospect of imputing
225 millions of low frequency variants into existing GWAS datasets.

226 **Acknowledgements**

227 The authors would like to thank Carl Anderson, Richard Durbin and Eleftheria Zeggini for helpful
228 comments on this manuscript. JCB is funded by Wellcome Trust grant WT089120/Z/09/Z.

229 **Conflict of Interest**

230 The authors declare no conflict of interest.

231 **Supplemental Data**

232 Supplementary information is available at the European Journal of Human Genetics' website.

References

- [1] Hindorff, L., Sethupathy, P., Junkins, H., *et al* (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U.S.A.* *106*, 9362–9367.
- [2] Zhernakova, A., van Diemen, C., and Wijmenga, C. (2009). Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat. Rev. Genet.* *10*, 43–55.
- [3] Frazer, K., Ballinger, D., Cox, D., *et al.* (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature* *449*, 851–861.
- [4] Barrett, J. and Cardon, L. (2006). Evaluating coverage of genome-wide association studies. *Nat. Genet.* *38*, 659–662.
- [5] Nothnagel, M., Ellinghaus, D., Schreiber, S., Krawczak, M., and Franke, A. (2009). A comprehensive evaluation of SNP genotype imputation. *Hum. Genet.* *125*, 163–171.
- [6] Howie, B., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* *5*, e1000529.
- [7] Marchini, J. and Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* pp. 499–511.
- [8] Browning, B. and Browning, S. (2009). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* *84*, 210–223.
- [9] Browning, S. and Browning, B. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* *81*, 1084–1097.
- [10] Huang, L., Li, Y., Singleton, A., Hardy, J., Abecasis, G., Rosenberg, N., and Scheet, P. (2009). Genotype-imputation accuracy across worldwide human populations. *Am. J. Hum. Genet.* *84*, 235–250.
- [11] Li, Y., Willer, C., Sanna, S., and Abecasis, G. (2009). Genotype imputation. *Annu Rev Genomics Hum Genet* *10*, 387–406.

- [12] The International HapMap3 Consortium. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* 467, 52–58.
- [13] Power, C. and Elliott, J. (2006). Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol* 35, 34–41.
- [14] Anderson, C., Pettersson, F., Barrett, J., *et al* (2008). Evaluating the effects of imputation on the power, coverage, and cost efficiency of genome-wide SNP platforms. *Am. J. Hum. Genet.* 83, 112–119.
- [15] Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* 39, 906–913.

Figure Titles and Legends

Figure 1: Effects Of Reference Set On Imputation Accuracy

A histogram of dosage r^2 scores across unfiltered SNPs genome-wide for samples imputed with HapMap2 and HapMap3 CEU, as well as HapMap3 CEU+TSI, and a reference set consisting of HapMap3 CEU+JPT+CHB of the same size as the CEU+TSI set.

Figure 2: Imputation Improvement Is Most Striking At Low Allele Frequency

The genome-wide increase in dosage r^2 for unfiltered imputed SNPs relative to HapMap2 CEU, plotted against minor allele frequency, for the four HapMap3 sample mixtures.

Figure 3: Overconfident And Underconfident Imputation

The rates of overconfident and underconfident mistakes in imputation, using various reference sets. An overconfident mistake is any SNP that is imputed with a predicted dosage $r^2 > 0.9$, but an actual dosage $r^2 \leq 0.8$, and an underconfident mistake has a predicted dosage $r^2 \leq 0.8$ and an actual dosage $r^2 > 0.9$.

Figure 4: Ancestry Mixtures Can Improve Rare Imputation

The relationship between the mean dosage r^2 across unfiltered SNPs and the proportion of non-CEU samples in a 100-sample reference set. The trend lines are quadratic least squared regression curves, and both explain the data significantly better than a linear relationship ($N = 207$, $p < 10^{-4}$ and $N = 159$, $p < 10^{-16}$ for TSI and CHB+JPT respectively). The insert shows an expansion of the trend lines between 0 and 50%.

Tables

Table 1: HapMap Samples

A summary of the HapMap sample sets and their sizes in the HapMap2 and HapMap3 datasets. We used release 21 of the phased HapMap2 data, and release 2 of the phased HapMap3 data.

Population	Code	HapMap2	HapMap3
African Americans	ASW	0	63
North Europeans	CEU	60	117
Chinese Americans	CHD	0	85
Gujarati	GIH	0	88
Japanese and Chinese	JPT+CHB	90	170
Luhya	LWK	0	90
Mexicans	MEX	0	52
Maasai	MKK	0	143
Toscani	TSI	0	88
Yoruba	YRI	60	155

Table 2: Effect Of Reference Set On Imputation

Information on Genome-Wide imputation using various reference sets. The CPU columns shows the number of CPU hours used in the imputation, which increases with the size and SNP density of the reference set. The proportion of SNPs that passed the ffil (predicted dosage $r^2 \geq 0.9$), and the mean dosage r^2 of those that passed, are shown for common (MAF > 0.05) and rare (MAF \leq 0.05) SNPs.

Reference Set	Size	CPU	Passed Filter		Filtered Dosage r^2	
			Common	Rare	Common	Rare
HM2CEU	60	514h ^a	83.7% ^b	52.5% ^b	0.957	0.889
CEU	117	296h	85.1%	59.7%	0.968	0.921
CEU+TSI	205	350h	86.1%	63.1%	0.974	0.934
CEU+TSI+GIH+MEX	345	458h	85.3%	60.3%	0.978	0.957
WORLD	1010	1207h	83.8%	55.5%	0.979	0.968

^a HM2 has a large SNP set, hence the longer imputation time

^b HM2 has a larger number of SNPs in total







