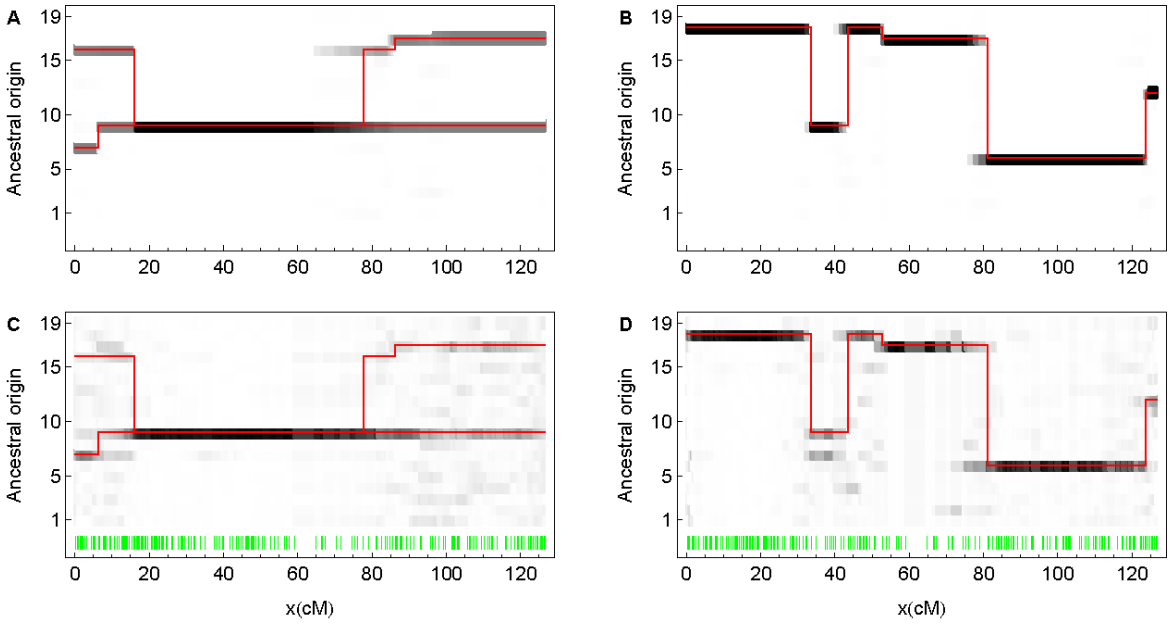

Figure S1 The posterior probabilities of the 19 ancestral origins obtained from the simulated datasets MAGIC-F5 (left panels) and MAGIC-F11 (right panels). The top and bottom panels denote the results obtained from RABBIT (jointModel) and HAPPY, respectively. The mode of HAPPY is diploid for the MAGIC-F5 and haploid for the MAGIC-F11. The probabilities are represented by gray levels, with white =0 and black =1. The red lines denote the true ancestral origins. The green vertical bars in the bottom panels denote the marker locations.

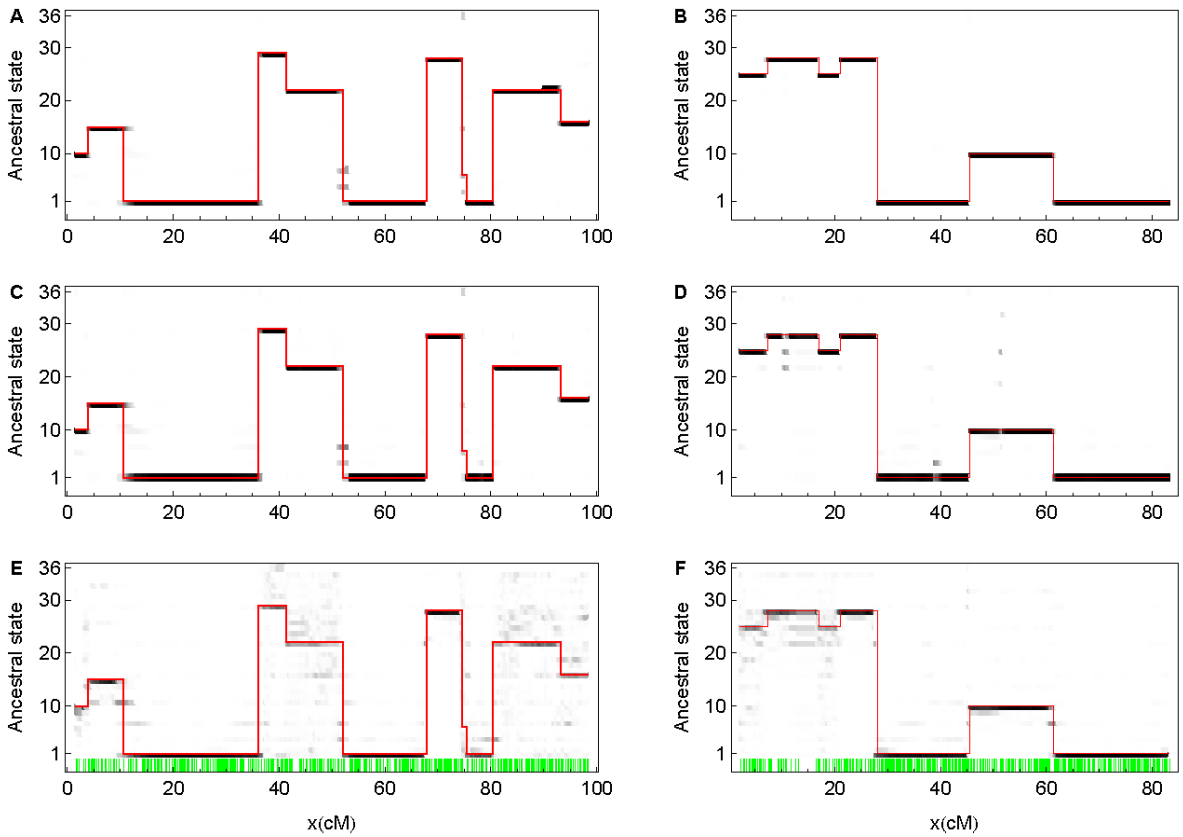

Figure S2 The posterior probabilities of the 36 ancestral origin states obtained from the simulated datasets CC-F11-AA (left panels) and CC-F11-XX (right panels). The top, middle, and bottom panels denote the results obtained from RABBIT (jointModel), GAIN, and HAPPY (diploid), respectively. The probabilities are represented by gray levels, with white =0 and black =1. The red lines denote the true ancestral origin states. The green vertical bars in the bottom panels denote the marker locations.

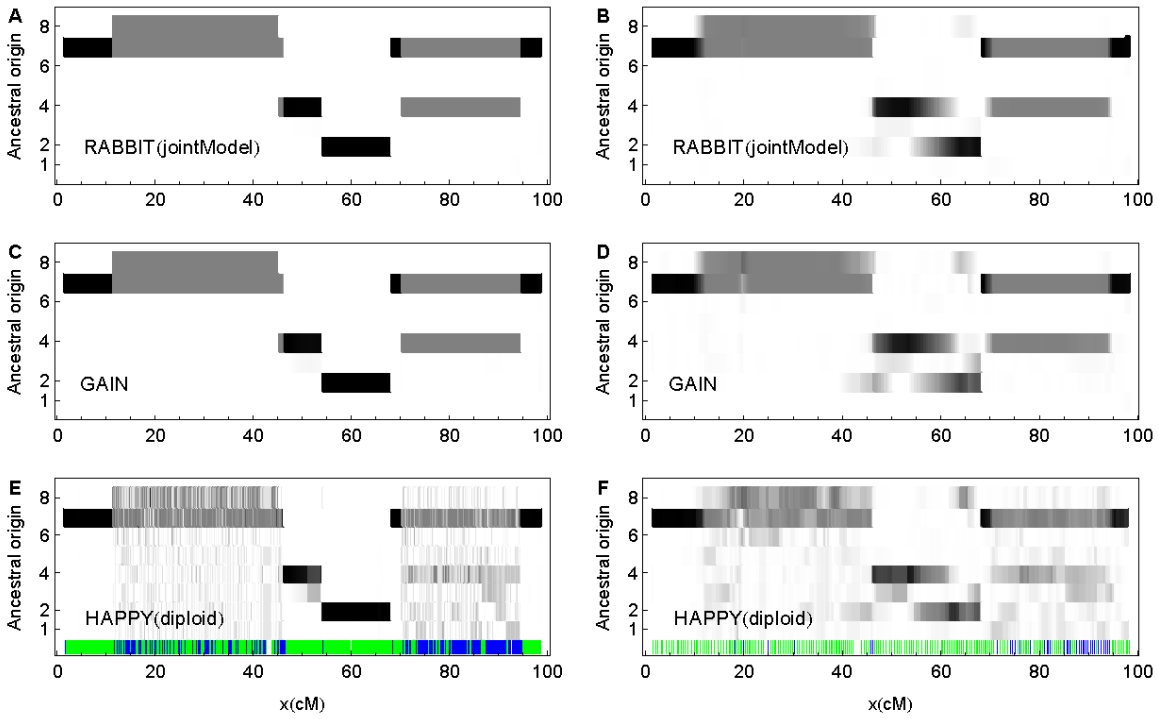

Figure S3 Similar to Figure 7 but for the posterior probabilities of the eight ancestral origins along the first pair of autosomes of the example pre-CC line (IL-18).

Table S1 Probability $P(Y|\boldsymbol{Z},\ \epsilon)$ of the observed genotype Y given the true phased genotype $\boldsymbol{Z}$, and the allelic typing error probability ϵ. Dashes denote missing alleles in a sampled individual. In practice, genotypes $1/-$ and $2/-$ are rarely called from probe intensity data.

Observed genotype Y	True phased genotype $\boldsymbol{Z}$			
	(1, 1)	(1, 2)	(2, 1)	(2, 2)
$-/-$	1	1	1	1
$1/-$	$1-\epsilon$	1/2	1/2	ϵ
$2/-$	ϵ	1/2	1/2	$1-\epsilon$
$1/1$	$(1-\epsilon)^2$	$\epsilon(1-\epsilon)$	$\epsilon(1-\epsilon)$	ϵ^2
$1/2$	$2\epsilon(1-\epsilon)$	$\epsilon^2+(1-\epsilon)^2$	$\epsilon^2+(1-\epsilon)^2$	$2\epsilon(1-\epsilon)$
$2/2$	ϵ^2	$\epsilon(1-\epsilon)$	$\epsilon(1-\epsilon)$	$(1-\epsilon)^2$

Table S2 Probability $P(\boldsymbol{D}|\boldsymbol{Z},\boldsymbol{O},\epsilon_F)$ of the derived genotype $\boldsymbol{D}$ given the true phased genotype $\boldsymbol{Z}$, the latent ancestral origin state $\boldsymbol{O}$, and the allelic typing error probability ϵ_F. The δ is an indicator of the latent IBD, and the question marks denote missing alleles derived from founders.

Derived genotype $\boldsymbol{D}$	True phased genotype $\boldsymbol{Z}$			
	(1, 1)	(1, 2)	(2, 1)	(2, 2)
(?, ?)	1	$1-\delta$	$1-\delta$	1
(?, 1)	$(1-\delta)(1-\epsilon_F)$	$(1-\delta)\epsilon_F$	$(1-\delta)(1-\epsilon_F)$	$(1-\delta)\epsilon_F$
(1, ?)	$(1-\delta)(1-\epsilon_F)$	$(1-\delta)(1-\epsilon_F)$	$(1-\delta)\epsilon_F$	$(1-\delta)\epsilon_F$
(2, ?)	$(1-\delta)\epsilon_F$	$(1-\delta)(1-\epsilon_F)$	$(1-\delta)\epsilon_F$	$(1-\delta)(1-\epsilon_F)$
(?, 2)	$(1-\delta)\epsilon_F$	$(1-\delta)\epsilon_F$	$(1-\delta)(1-\epsilon_F)$	$(1-\delta)(1-\epsilon_F)$
(1, 1)	$\delta(1-\epsilon_F)+(1-\delta)(1-\epsilon_F)^2$	$(1-\delta)\epsilon_F(1-\epsilon_F)$	$(1-\delta)\epsilon_F(1-\epsilon_F)$	$\delta\epsilon_F+(1-\delta)\epsilon_F^2$
(1, 2)	$(1-\delta)\epsilon_F(1-\epsilon_F)$	$(1-\delta)(1-\epsilon_F)^2$	$(1-\delta)\epsilon_F^2$	$(1-\delta)\epsilon_F(1-\epsilon_F)$
(2, 1)	$(1-\delta)\epsilon_F(1-\epsilon_F)$	$(1-\delta)\epsilon_F^2$	$(1-\delta)(1-\epsilon_F)^2$	$(1-\delta)\epsilon_F(1-\epsilon_F)$
(1, 2)	$\delta\epsilon_F+(1-\delta)\epsilon_F^2$	$(1-\delta)\epsilon_F(1-\epsilon_F)$	$(1-\delta)\epsilon_F(1-\epsilon_F)$	$\delta(1-\epsilon_F)+(1-\delta)(1-\epsilon_F)^2$

Table S3 Probability $P(Y|Z,\ \epsilon)$ of the observed allele Y given the true allele Z, and the allelic typing error probability ϵ. Dashes denote missing alleles in a sampled individual.

Observed allele Y	True allele Z	
	1	2
–	1	1
1	$1-\epsilon$	ϵ
2	ϵ	$1-\epsilon$

Table S4 Probability $P(D|Z, O, \epsilon_F)$ of the derived allele *D* given the true allele *Z*, the latent ancestral origin *O*, and the allelic typing error probability ϵ_F. The question marks denote missing alleles derived from founders.

Derived allele *D*	True allele *Z*	
	1	2
?	1	1
1	$1 - \epsilon_F$	ϵ_F
2	ϵ_F	$1 - \epsilon_F$