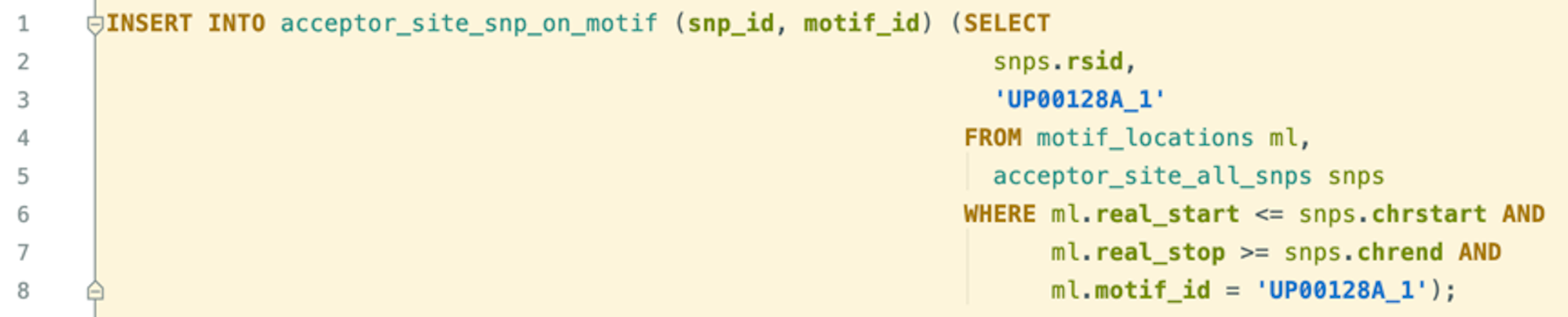
## MOTIF SCORING WITH SPLICE VARIANT SNP DATA

***Calculating SNPs’ Co-located With Motifs***



Sample SQL script for calculating SNPs’ co-location on significant motifs. “acceptor\_site\_snp\_on\_motif” table stores the SNP-motif pairs for co-locations. The query snippet shown in this figure is run in the same manner for all motifs.

***SNPs’ Splicing Effect Prediction With SPANR***

After downloading SPIDEX data, we stored that into our local database and computed the SPANR scores of our SNPs by running the following SQL queries.

1. **SPANR scores table creation**



Table creation for storing SPANR scores of splice region variants. This table is populated with SPIDEX data(<http://download.openbioinformatics.org/spidex_download_form.php>).

1. **Arranging alleles**



Queries for arranging alleles of acceptor region SNPs into separate columns.

1. **Reference and most common allele detection**



Queries for detecting the reference and most common alleles of acceptor region SNPs. The allele with greatest allele frequency indicates the reference allele which refers to the nucleotide base on the reference assembly at the SNP's position. Then, the second greatest indicates the most common allele change.

1. **Allele mapping to positive strand**



Queries for mapping alleles of acceptor region SNPs to positive strand.

1. **Setting SNP scores**



Query for setting SPANR scores of acceptor region SNPs.