**SUPPLEMENTAL MATERIAL**

**Phenotypic Assessment**

For HbF levels, peripheral blood drawn from SIT Trial study subjects was mixed 1:1 with Alsever's solution (Sigma, Saint Louis, MO), stored at 4°C, and processed within 1 week of being drawn. Hemolysates were prepared for each subject by lysing cell pellets in 8-12 volumes of distilled water, centrifuging at 12,000 rpm for 15 minutes, and storing the clear supernatant at -80°C. Lysates were analyzed by HPLC for HbF levels. To avoid confounding due to systematic differences in local measurements of HbF levels at participating sites, we only included the 456 (42.14%) SIT Trial samples whose HbF levels were measured in a central laboratory at Johns Hopkins University. Similarly, for the CSSCD cohort, analysis was restricted to those samples whose HbF data was centrally generated at Centers for Disease Control and Prevention (n=764). None of the patient from the SIT Trial and CSSCD cohorts was using hydroxyurea when the HbF measurements were performed.

**Hemoglobinopathy Status**

To ascertain the genetic basis for the hemoglobinopathy in SIT Trial samples, DNA was isolated from patient-derived EBV-transformed lymphoblast cells and the β-globin gene sequence was determined bidirectionally on a 3730 DNA Analyzer with BigDye® Terminator v1.1 and compared to reference sequence (NC_000011) obtained from the National Center for Biotechnology Information (NCBI) using the SeqScape v2.5 software (Applied Biosystems, Foster City, CA). In both SIT Trial and CSSCD cohorts, analysis was restricted to Hb-SS patients.

**Statistical Analysis**

Cryptic relatedness in the study cohorts was determined by examining pair-wise identity-by-state (IBS) and the first degree related samples (including full siblings, parent-offspring etc) were excluded from the analysis. Between each detected pair of related individuals, we adopted a criterion of retaining the first recruited sample for the analysis and the other related sample was dropped. The differences in the demographic and hematological parameters between the SIT Trial and CSSCD participants were compared using student's t-test. All association analysis was performed using R statistical computing environment (http://www.r-project.org/) (version 2.14.1). We found the distribution of the pain episodes in both the cohorts was over-dispersed and the variance-to-mean relationship suggests that a Poisson model, with correction for over-dispersion, fits best to our data. The distribution of HbF levels was slightly skewed; therefore, a $\log_{10}$-transformation was applied. In our study, we did not perform analyses using any model selection criteria, but instead relied on the models used in a seminal study conducted by Platt *et al*, which analyzed the CSSCD cohort. In this study, Platt *et al* comprehensively investigated various clinical and hematological parameters that influence pain rate (episodes per year), using the Poisson regression model, and showed that among all the covariates used, only age, hematocrit and HbF levels are associated with high pain rates. Similarly, in our study, pain episodes were modeled using a multivariable Poisson regression, adjusting for age, sex, and hematocrit levels. In the multivariable analysis, the significance for the associations is reported as a two-tailed p-value. Vaso-occlusive pain episode variance explained by these covariates was estimated using change in the deviance from the null (with only an intercept) to the covariate model after adjustments for the estimated dispersion parameter. Further, to estimate the overall effect of HbF levels on severe vaso-occlusive pain, association results were combined in a fixed effect meta-analysis using MetABEL package (http://www.genabel.org/packages/MetABEL).