

1. 課題名

ゲノムデータのブリッジ・スタディのための遺伝統計学の挑戦

2. 代表者名

江口 真透（統計数理研究所）

3. 研究成果の概要

遺伝子発現データ，高密度 SNP ハプロタイプデータ，プロテオーム関数データなどを含むゲノムデータの解析に関して「 $p \gg n$ 」問題にチャレンジした．これはデータの次元 p がサンプル数 n よりもきわめて大きくなるケースを指す．異なる施設のデータを併合することによって「 $p \gg n$ 」問題の解決を目指した．具体的には，遺伝子発現のデータをプールして，その中から共通な遺伝子に絞って，異なる施設のからの被験者を参照することである．特に，遺伝研の DDBJ (DNA Data Bank of Japan) に蓄積されつつある遺伝子発現データベース CIBEX に基づき大規模な解析実験を計画した．この方法は，SNP ハプロタイプのデータに対しても適用可能になることも検討した．