別紙２

総括研究報告書

１．研究開発課題名：個別化医療を実現するプライバシ保護ゲノム情報解析

２．研究開発代表者：　清水 佳奈（産業技術総合研究所）

３．相手国研究代表者：Antti Honkela（ヘルシンキ大学（フィンランド））

４．研究開発の成果

当該年度は H26 年度に引き続き，プライバシ保護ゲノム配列検索の開発に取り組んだ．主として以下の三つの課題に取り組んだ．

1) 遺伝子型秘匿検索システムの開発

H26 年度に実装した遺伝子型検索の暗号プロトコルをコアとして，データベースと連動して動作する秘匿検索システムを完成させた．さらに，完成したシステムを運用中の個人ゲノムデータバンクに導入して運用試験を行った．実運用に耐えうるシステムを構築するために，H26 年度に設計したプロトコルの改良を行い，通信量のオーバーヘッドを O(N)から O(√N)に改善した．また，暗号プロトコルの安全性のみならず，ユーザー認証からクライアントソフトウェアの配布に至るまで，システム全体として高い安全性を達成するため，通信プロトコルのパラメータなど詳細な部分に渡って検討した．また，利用者の利便性を考慮してシステムの操作を行うための GUI も開発した．

2) 高い機能性を有する秘匿ゲノム検索プロトコルの開発

クエリを暗号化したまま，データベースから類似するゲノム配列を検索することのできる方法論を考案した．本手法では，Positional Burrows Wheeler Transform と呼ばれる離散データ構造と，ベクトルから指定した要素のみを安全に取得する紛失通信という暗号技術を組み合わせることによって検索を行う．従来手法では，クエリの長さに対して指数関数的な計算量が必要であったが，本手法では線形の計算量で部分/最長文字列の一致を検索することができる． 2184 本のゲノム配列を含む 1000 ゲノムプロジェクトのデータを持いて実験を実施したところ，25 塩基のパターンを数秒で検索することができた．また，考案した手法のプロトタイプ実装を web 上で公開した．本手法の応用範囲は広く，ゲノム配列のみならずアミノ酸配列やその他の文字列に対しても用いることが可能．

3) ゲノム情報の特殊性を考慮したプロトコルの考案

ゲノム情報は子孫に遺伝する性質を持つため，通常の個人情報よりも長期間に渡って保護をする必要がある．ところが一般的な公開鍵暗号方式ではそのような長期間に渡って平文を保護することを想定していないため，ゲノムに適した保護の方策が必要である．本研究では，サーバーとクライアントのみが共有するかく乱要素を暗号文（もしくは平文）に作用させることにより，第三者の盗聴からゲノム情報を長期間に渡って保護することが可能な紛失通信プロトコルを考案した．

年度終盤には，H28 年度に予定しているプライバシ保護モデル学習の技術開発に向けた準備を行った．

上記の研究開発に加えて，本研究課題に関連する学術コミュニティの活性化を目的として，バイオインフォマティクスの国際会議(GIW/InCoB 2015) に併設した国際ワークショップ PRIVAGEN2015 を開催した．本ワークショップは，日本側の研究者とフィンランド側の研究者の双方により共同開催し，10 か国以上から 80 名の参加があった．さらに，国内最大のバイオインフォマティクスの会議（IIBMP2015）においては，H27 年度に改正法が施行された個人情報保護法に関して，ゲノム情報解析にどのような影響が及ぶのかを議論するセッションを開催した．

今年度は，5 件の論文発表（査読付き国際誌 4 件，査読なし国内誌 1 件）を達成した他，学会より 2 件受賞（生命医薬情報学連合大会 2015 年大会において「研究奨励賞」及び「最優秀口頭発表賞」を受賞），所属組織より 1 件の受賞（2015 年度 産総研 理事長賞（研究）受賞）を達成した．