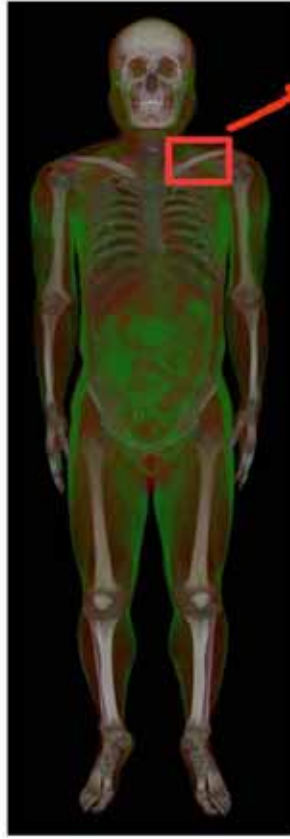
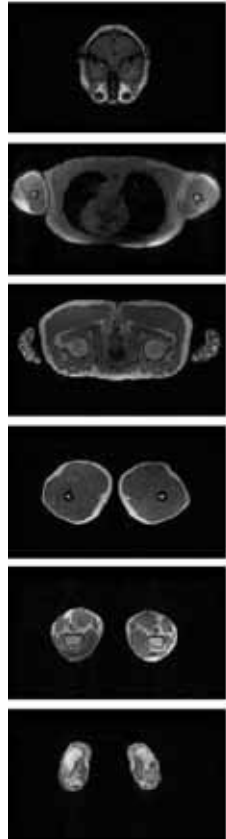


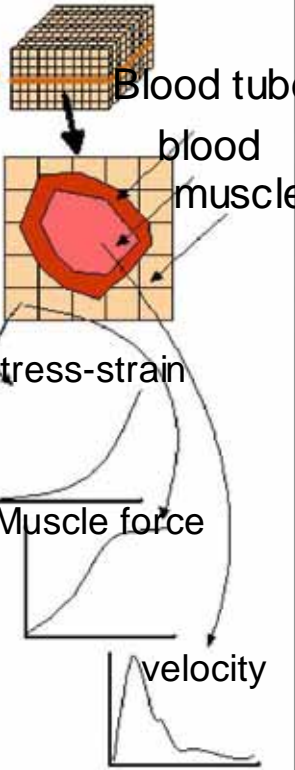
臓器全身スケールに関する開発

Medical images

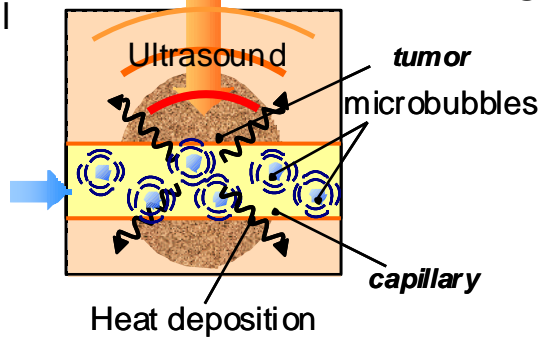


Whole body voxel model

Whole body



Multi physics



Surface model

Local precise model



Cerebral artery model



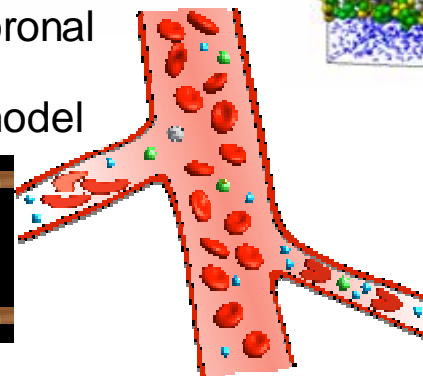
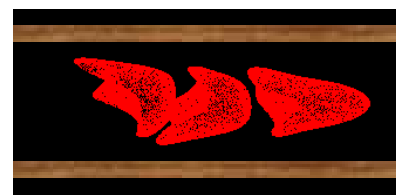
Eyeball model



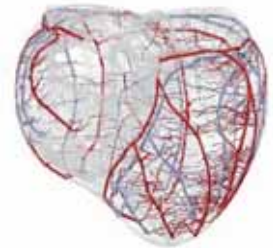
Leg model

coronal

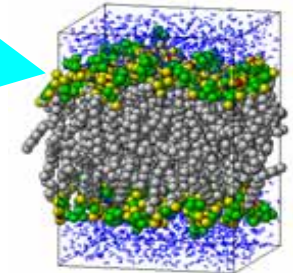
Blood flow model



Heart model



Membrane model



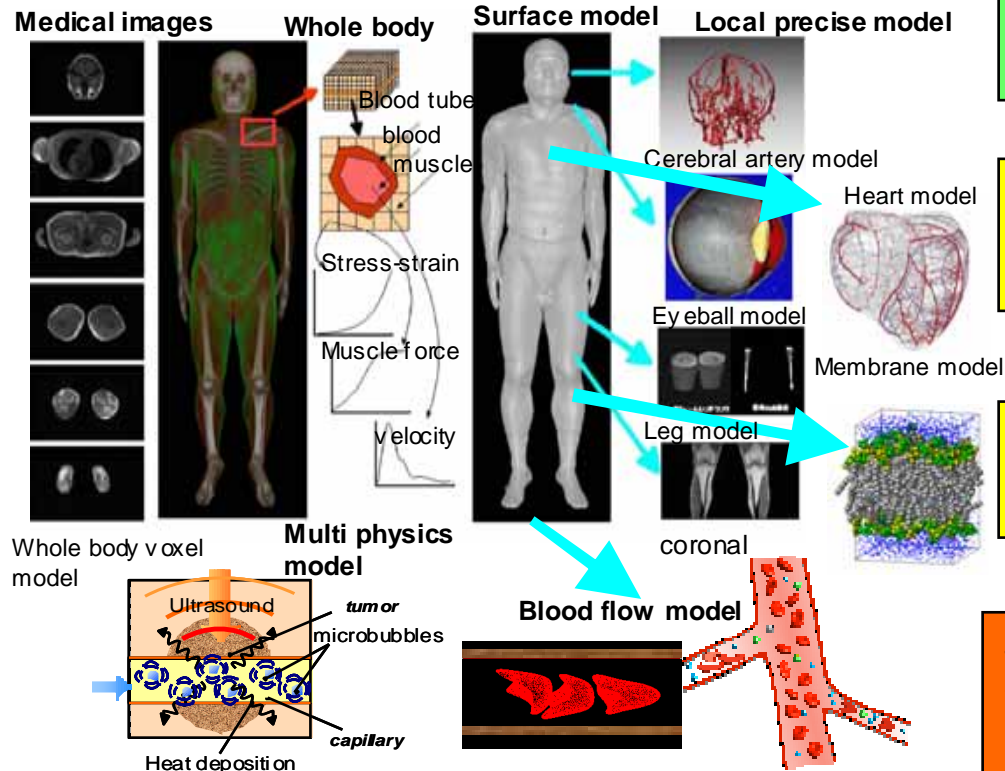
臓器全身スケールシミュレーション

筋骨格、種々の臓器、循環器系、呼吸器系、神経系など、全て備えた人体モデル化技術開発、及び癌治療と循環器系疾患に関連するモデル開発をベースに、主に構造流体連成解析、循環器系シミュレーション等により姿勢を変化させたときの癌の体内位置予測や治療法の開発を目指す

筋骨格、種々の臓器、循環器系、呼吸器系、神経系など、全て備えた人体モデル化技術、及び癌治療と循環器系疾患に関連するモデル化技術の確立

個別に開発されたシミュレーションモデルを、理研で開発中のV-CAD人体モデルを元に、各種シミュレーションが統一的に取り扱える技術の開発

血管と他の臓器、筋肉、骨なども連成した統一的力学シミュレーションにより、血管狭窄、血栓形成など血流に関連した疾患の治療や診断、手術方法の検討が可能



このような研究のためには

筋骨格、種々の臓器、循環器系、呼吸器系、神経系など、全て備えた人体モデル化技術、及び癌治療と循環器系疾患に関連するモデル化技術の確立が必要であり、統一したモデルで構造流体連成等の統一的力学シミュレーション技術の開発が必須

しかしながら現時点では、従来の手法を用いても

従来、臓器のシミュレーションモデルは個別に扱われており、更に、計算能力の問題で、統一したモデルでのシミュレーションは実行不可能

ボトルネックは

従来、精密なシミュレーションモデルは循環器、呼吸器、筋骨格モデル、神経系モデル等、個別に扱われており、統一したモデルが未確立

本グループにおける挑戦的問題解決

個別に開発されたシミュレーションモデルを、V-CAD人体モデルを元に、各種シミュレーションが統一的に取り扱える技術を確立する

さらには、新規方法論を用いて

次世代スーパーコンピュータをフルに活用する事で、力学的マルチスケール臓器モデルを構築し、血管と他の臓器、筋肉、骨なども連成した統一的力学シミュレーションが可能になる

これらにより、従来は不可能であった

血流に関連した疾患の治療や診断、手術方法の検討に応用でき、重イオン粒子を使った癌の放射線治療時に問題となる姿勢変化時の臓器の移動を予測が可能。また、人体の運動に関してはリハビリや補助具の開発にも貢献

(4) データ解析融合に関する研究

研究開発の内容: データ解析のための統計モデルやシミュレーションモデルなど、生命体に関する様々なモデルが研究されてきているが、現状では、問題の大規模化・複雑化、及び生産されるデータに十分に対応できていない。本研究では、ペタスケール計算の能力を用いて、こうしたモデルに、データの情報を動的かつ適切に取り込むためのデータ同化技術を開発することにより、モデル及びデータ、それぞれ単独では得られない有用な情報を抽出することが可能なデータ駆動型の大規模データ解析及びシミュレーション技術を開発することを目標とする。これにより、ペタスケールデータ解析とシミュレーションの融合に基づく、個人差を考慮した投薬量・最適投与プロセスなどの開発、タンパク質構造や大規模遺伝子ネットワーク推定による創薬ターゲット・毒性関与パスウェイ探索の技術が成果として期待される。シミュレーションモデルと現実データとの乖離を埋める技術として、地球物理の分野で発展してきたデータ同化とよばれる枠組みを用いた研究を進める。また、個々のサブテーマに出現する並列化問題に対してはアーキテクチャとアルゴリズムの観点から、効率的並列化の手法を研究する。

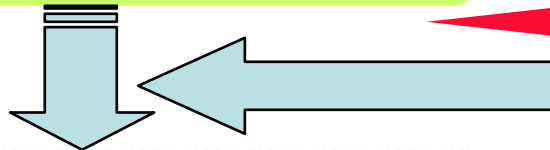
研究開発の優位性: 生物に関しては方程式のはっきりと分かった領域と未知の領域が存在する。分子スケールと全身スケールはそれぞれ支配方程式が分かっているが、細胞スケールでは未知のものが多い。データ解析融合チームはこの方程式が未知であるが、膨大なデータからなんらかの方法で生物で起こっている現象に迫ろうというものである。具体的には、飛躍的に増大しているゲノムや遺伝子発現データを解析する、高度なデータマイニング技術に基づいた応用技術と、データ同化によるデータとシミュレーションモデルの融合を図るモデル構築技術にある。これらは、ペタスケールの計算能力があって初めて実現できるものであり、今後の生命科学研究において極めて重要な技術となる。その優位性は、データ同化と呼ばれる枠組みにあり、モデル及びデータ、それぞれ単独では得られない、有用な情報を解析することができる。

データ解析融合の概観

超高次元大規模ヘテロデータ解析技術と生体生命シミュレーションの融合

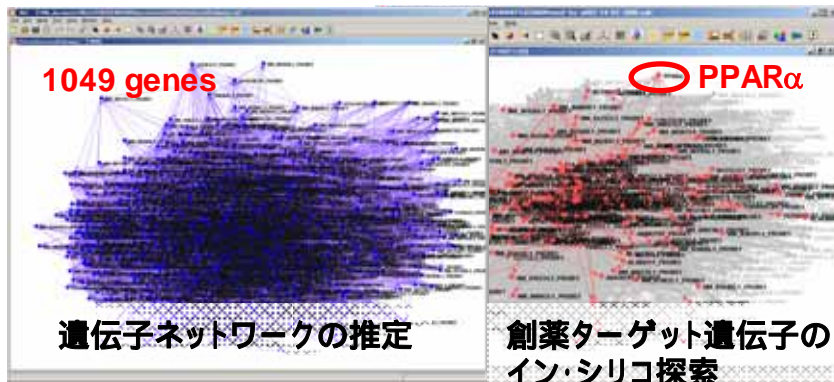
実験データの飛躍的増大

- ・超高次元化: 遺伝子からエクソンレベルへ
- ・極めてヘテロ: 配列, 構造, ダイナミクス, ...
- ・不観測性: 情報の欠損・欠如



期待される成果

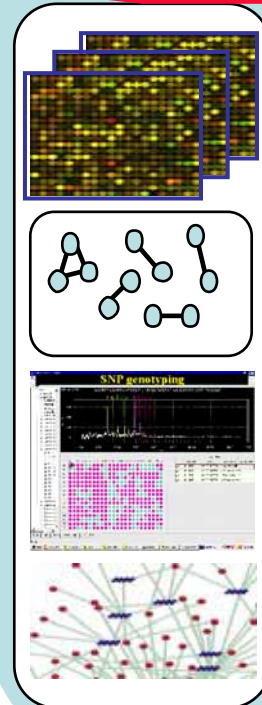
- ・ペタスケールデータ解析による個人差を考慮した投薬量・最適投与プロセスなどの開発
- ・ 10^4 スケール遺伝子ネットワーク推定による創薬ターゲット・毒性関与パスウェイ探索法の開発



ペタスケール
データ解析技術

機能のモデル化
技術

京速コンピューティング



データ駆動型大規模シミュレーション技術の創出



- ・データ同化技術による, シミュレーションモデルと現実データとの乖離を埋める技術
- ・大規模モデルリングの自動化技術
- ・高度なデータマイニング技術に基づいたデータ解析技術とシミュレーション科学の融合を可能にする情報基盤技術を創出

データ解析融合

高次元化・膨大化する遺伝子及び関連する情報に対して、高度なデータマイニング技術に基づいたデータ解析とシミュレーション科学を融合するデータ同化技術を確立し、生命体システムに対する予測と発見の基盤情報技術を構築する

データ同化技術による、シミュレーションモデルと現実データとの乖離を埋める技術の確立

モデル・データ単独では得られない有用な情報を抽出可能なデータ駆動型大規模データ解析・シミュレーション技術の開発

遺伝情報の網羅的解析に基づく個人差を考慮した投薬量・最適投与プロセスなどの開発や創薬ターゲット・毒性関与パスウェイ探索法の開発に貢献

このような研究のためには

超高次元・超ヘテロかつ膨大なデータの解析を可能とする理論と計算能力が不可欠である。その下でシミュレーションモデルと現実データとの乖離を埋める解決法が必要。

しかしながら現時点では、従来の手法を用いても

データ解析はデータの多様化と増大化に急速に引き離され、シミュレーションは生命体個別の現実データを反映できず予測能力に限界がある。

ボトルネックは

圧倒的計算能力の不足。例：遺伝子ネットワーク推定による創薬標的遺伝子探索法（計算時間が遺伝子数の3乗で増加）では、ヒト全遺伝子の1/30しか解析できない。また、シミュレーションモデルに現実データを現実的・合理的に融合する技術が未確立。

本グループにおける挑戦的問題解決

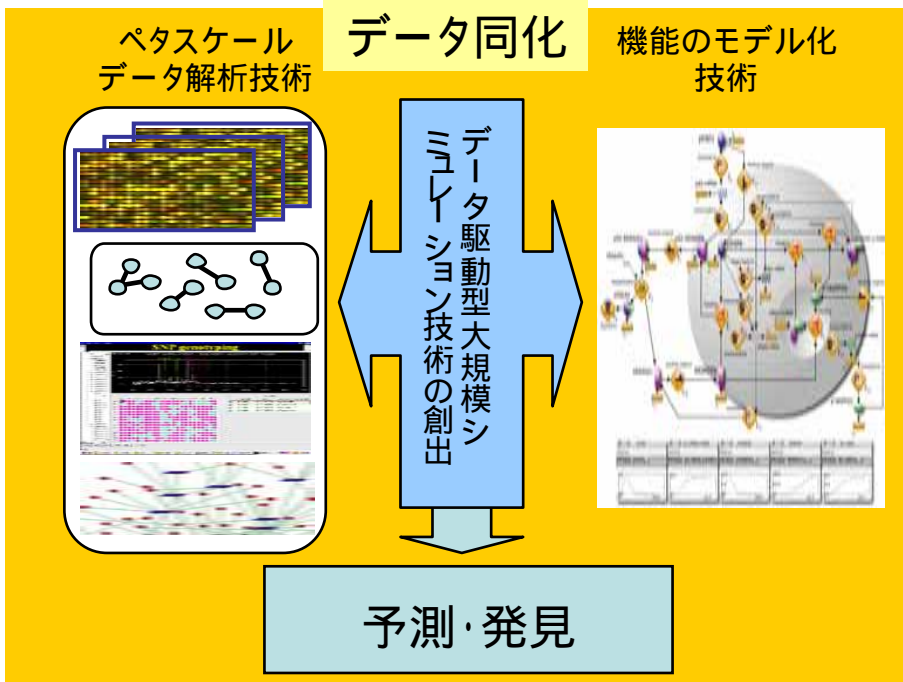
ベイズ統計によるデータ同化技術や異種情報統合法の開発、遺伝子ネットワーク推定や遺伝情報解析の並列アルゴリズムの採用により、次世代スーパーコンピュータの性能をフルに活用。

さらには、新規方法論を用いて

大規模生体モデルの自動構築により、ヒト全遺伝子を対象に現実データを利用したネットワーク推定による創薬標的探索が5PFLOPSで実現。更に複数モデルの評価・比較・選択・統合。

これらにより、従来は不可能であった

データ解析とシミュレーションを融合した生体システムに対する予測と発見のためのデータ駆動型のペタフロップスの計算方法論を構築。現実的問題に対してデータ同化が初めて実現。



生命体統合シミュレーションソフトの現状と 次世代スーパーコンピュータでの予測

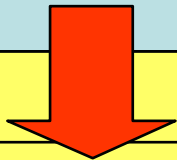
現状

分子スケール: ProteinDF

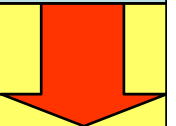
- ターゲット分子: インスリン6量体(306残基)
原子数4716、電子数:18552
- 用いた計算システム: SGI Altix3700
Itanium2 64CPU, MMU:512GB,
Peak性能:333GFLOPS
- 計算時間:10日

全身スケール:

- ターゲット:人工心臓(左右両心室、1拍)
- 用いた計算機システム:PCクラスタ
Pentium4x128CPU(3.2GHz), MMU/CPU:2GB
- 計算時間
 - マクロスケール
 - ・力学現象:260万自由度:7時間、
 - ・電気現象:2800万自由度:1時間
 - マルチスケール
 - ・26万自由度の細胞モデルを100要素:約1時間



次世代スーパーコンピュータ



- ターゲット分子:
原子数、電子数を3倍に増加
(演算量は約10倍に)
励起状態の計算も可能
- 必要な計算資源
Memory容量:0.5PB
Disc容量:1PB
- 推定計算時間:数時間

- ターゲット:人工心臓(全心臓モデル、1拍)
- 必要な計算資源
Memory容量:2.5PB
- 推定計算時間
 - マルチスケール
 - ・16.6万自由度の細胞モデルを88万マクロ要素
モデルへ組み込んだ連成計算:約10時間

(年次計画)

テーマおよびサブ テーマ	18年度	19年度	20年度	21年度	22年度	23年度	24年度
1分子スケールに関する研究	既存ソフトの改良によって、反応と分子運動との共役、薬剤設計。基質結合による構造変化、分子間相互作用、フォールディング等	既存ソフトの改良による薬剤設計・酵素反応、基質結合による構造変化・分子間相互作用、新モデリング法の開発等	既存ソフトの改良による薬剤設計・酵素反応、基質結合による構造変化、分子間相互作用、ロジモデリング法の継続開発等	新規ソフトウェアの開発による高速な方法論の開発、超巨大系・ウイルス丸ごとシミュレーション等。	新規ソフトウェアの開発によるより高速な方法論の開発、超分子複合体・ウイルス丸ごとシミュレーション等	シミュレーションプログラムの改良とテスト	開発された最先端・高性能汎用スーパーコンピュータ上での実行テスト・改良
(2)細胞スケールに関する研究	場（細胞分画の代謝情報、細胞膜を介した水・電解質移動等）を考慮した統合シミュレーションに向けたフォーマットの検討	場を考慮したシミュレーションに向けたフォーマットの開発シミュレータの拡張	細胞CADに関する研究 分化細胞（赤血球細胞等）シミュレータの開発	オルガネラ追加モデルの構築、複数シミュレーション連携、疾患時のシミュレーション開発	複数シミュレーション連携の研究 多細胞モデル、オルガネラ追、肝細胞1細胞シミュレーション等	複数シミュレーション連携に関する研究 改良と高速化、	統合プラットフォームの開発、開発された最先端・高性能汎用スーパーコンピュータ上での実行テスト・改良
(3)臓器全身スケールに関する研究	全身ボクセルデータ作成の高度化。生体内超音波伝播、超音波-マイクロバブル、心臓完全体シミュレータ、心臓全体完全シミュレータ等の研究に着手。	神経・筋・骨格系動態、神経・筋・骨格系動態シミュレーション、姿勢変更時の臓器位置検出シミュレータ等の改良	生体内超音波伝播シミュレータ、超音波-マイクロバブルシミュレータ、超音波診断・治療シミュレータ開発着手	リアルタイム診断・治療システム支援・メソスケール血流循環、全呼吸器系シミュレータ等開発着手	リアルタイム診断・治療システム支援シミュレータ全循環器系の開始、系の接合を開始	リアルタイム診断・治療システム支援シミュレータ等の高度化開発、各系のシミュレータ連結	開発された最先端・高性能汎用スーパーコンピュータ上での実行テスト・改良
(4)データ解析融合に関する研究	既存ソフトの改良により、疾患や薬物反応性などの個人表現型と配列情報を関連、データ同化技術で数十次元モデルの推定技術開発	ペタ級対応のため継続した改良開発、数百次元モデルに対応できるように改良	データ同化技術により、数千次元モデルの自動構築技術開発、ペタスケールデータ解析とモデル構築技術等開発	技術改良により、数万次元モデルへの適用、創薬・毒性関与パスウェイの探索応用技術開発	モデルの自動構築技術継続改良	シミュレーションプログラムの改良とテスト	開発された最先端・高性能汎用スーパーコンピュータ上での実行テスト・改良