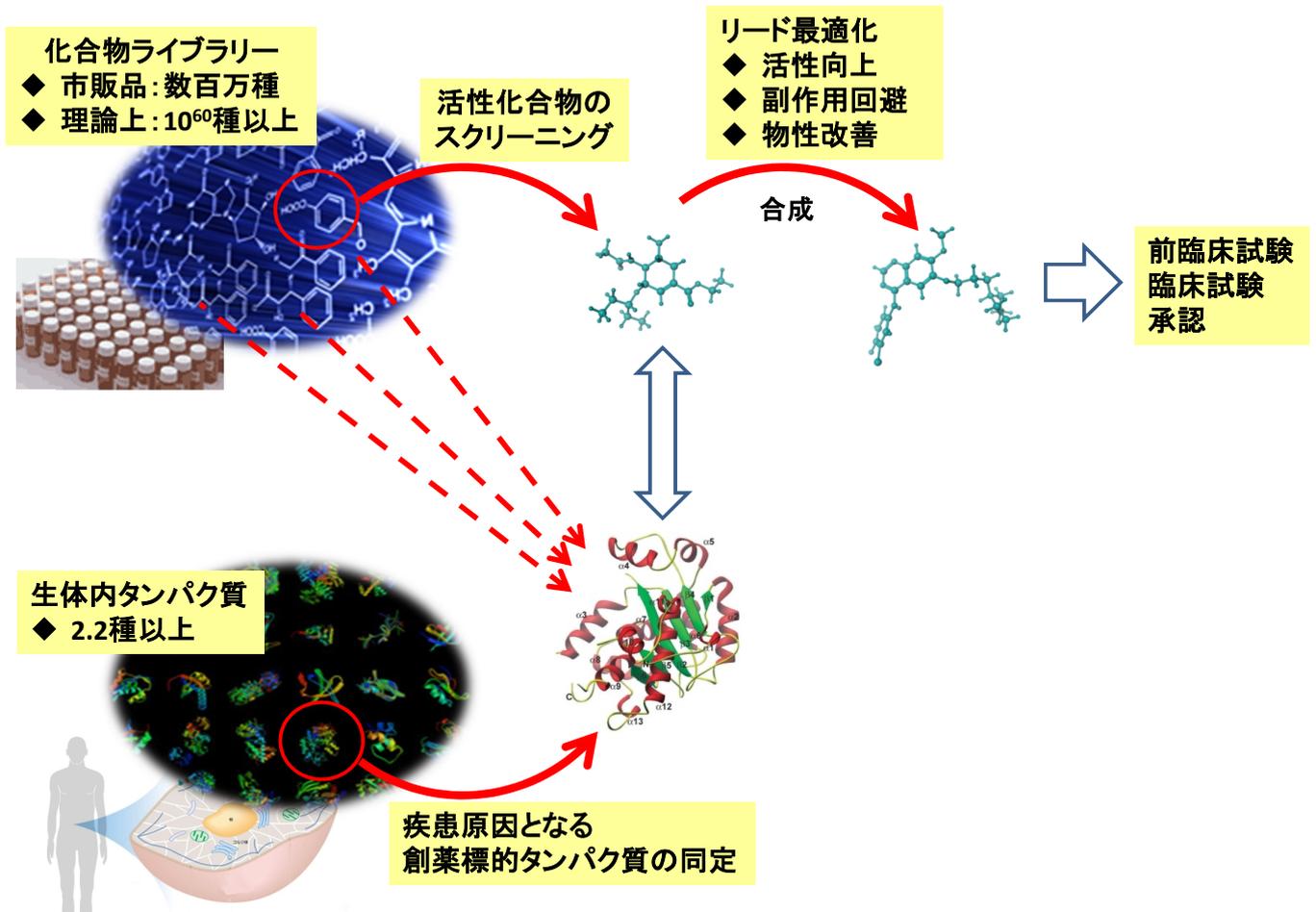


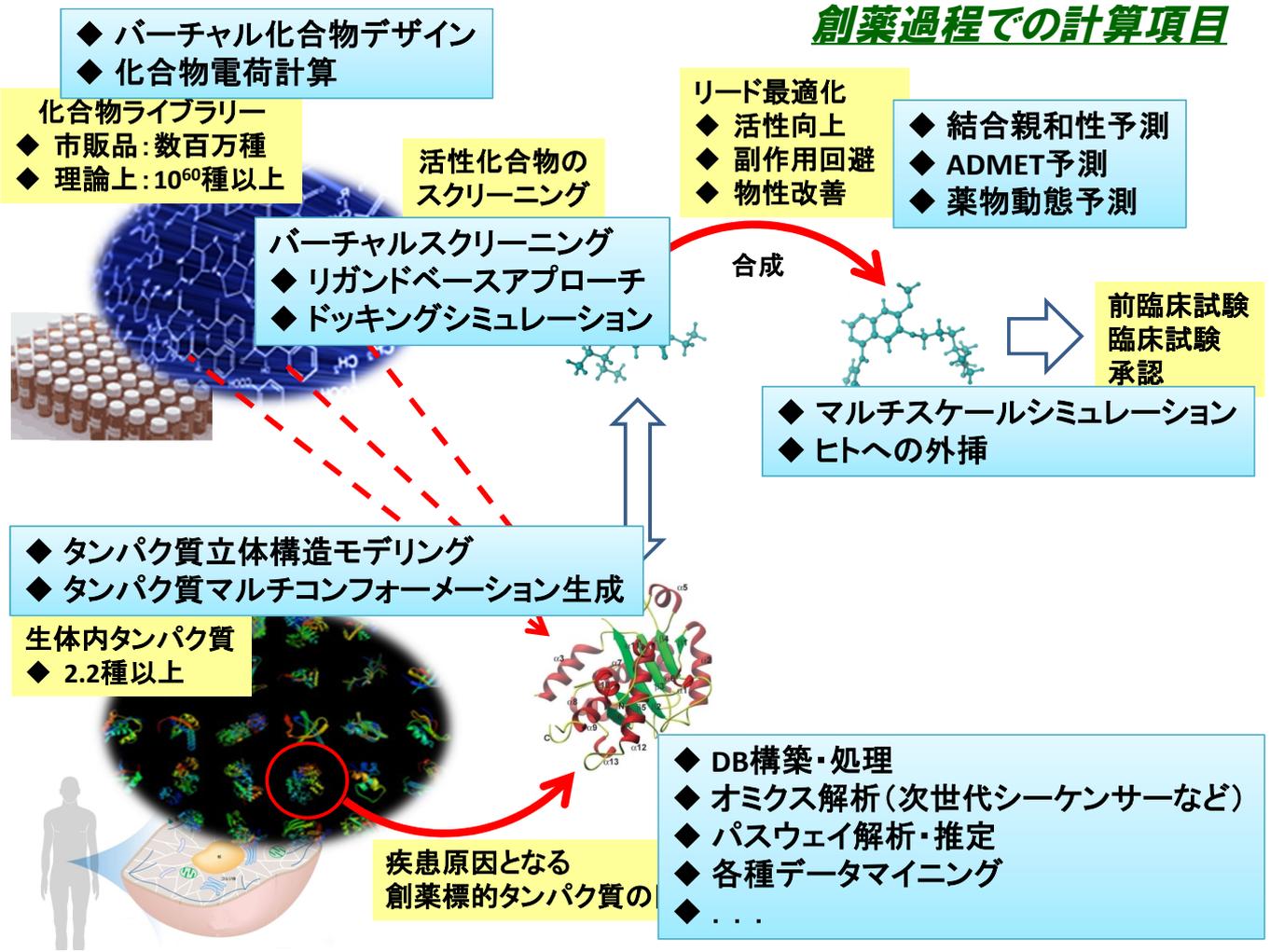
創薬分野における 計算アプリケーションの現状

京都大学 奥野恭史

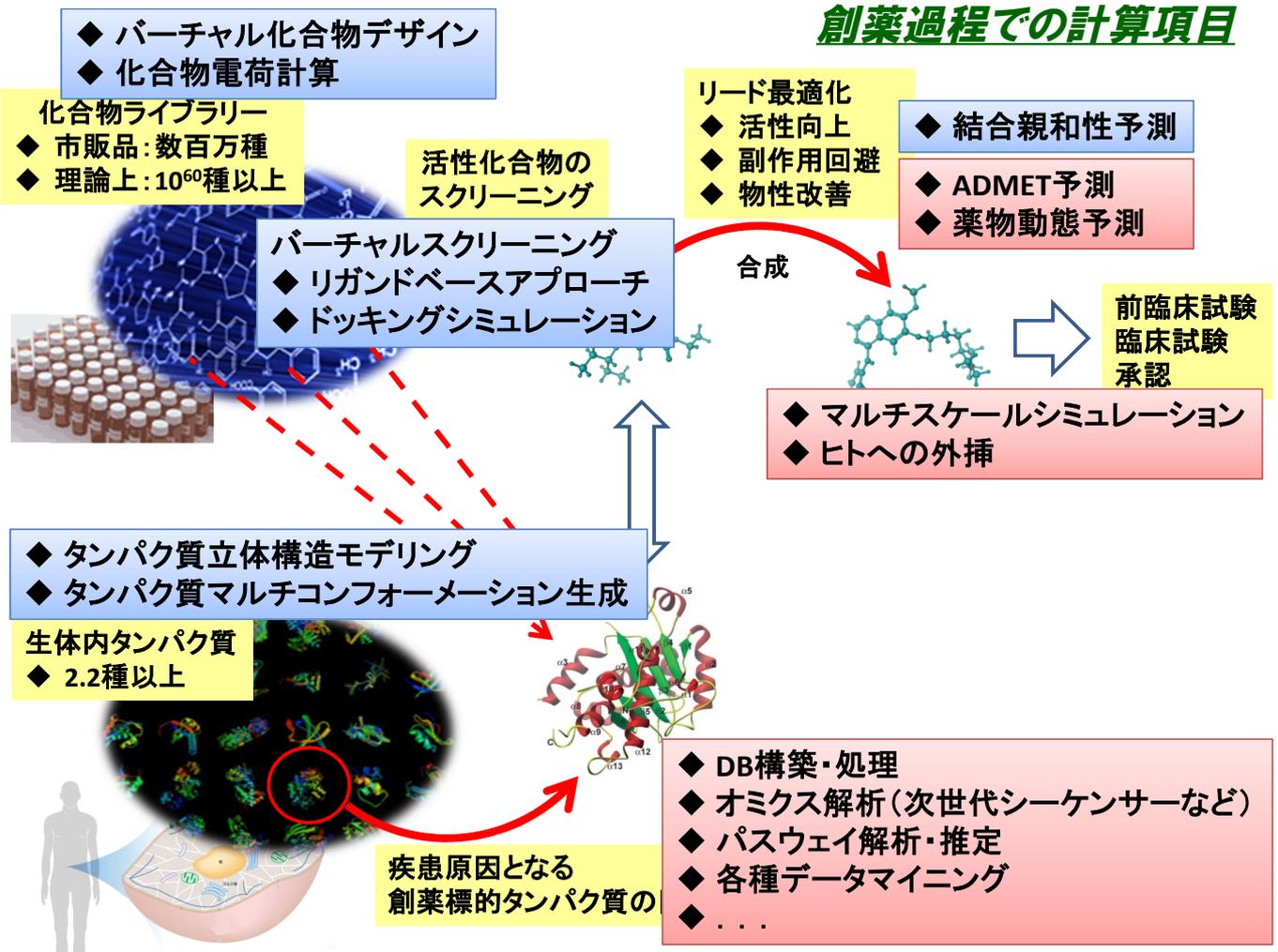
創薬(分子標的創薬)の流れ



創薬過程での計算項目



創薬過程での計算項目



創薬計算で利用されているソフトの現状 (1)

	ソフト	計算規模 (計算時間は汎用計算サーバー、数10~数100コアを目安)	実行の形態	実行頻度	ユーザーの要望	その他
DB構築・処理	SQL(オープンソース)	最大10の7乗程度(数万遺伝子×数千サンプル)のエントリー数に対し、数秒から数分で検索	インタラクティブジョブ	数~数十テーマ×各社/年	現状では実験の手間やコストによるサンプル数の制約が大きい。近い将来はサンプル数が急増する可能性大。よって、将来的に大規模データの高速処理のニーズが高くなる。	
オミクス解析(次世代シーケンサー(NGS)など)	各種パイオインフォマティクスツール(オープンソース、インハウス、市販)	最大10の7乗程度、または、NGSの場合は10の11乗程度(30億塩基×数百サンプル)のエントリー数に対し、数分から数時間で計算	シングルジョブ or バルクジョブ (10負荷のため比較的インメモリ容量大)	同上	同上	理研・東大(宮野G)らにより「京」アプリを開発
パスウェイ解析・推定	パイオインフォマティクスツール(オープンソース、インハウス、市販)	最大10の4乗程度(数百遺伝子×数百サンプル)のエントリー数に対し、数時間から数日で計算	シングルジョブ or バルクジョブ	同上	同上	同上
各種データマイニング	各種パイオインフォマティクスツール(オープンソース、インハウス、市販)	最大10の7乗程度のエントリー数に対し、数分から数時間で計算	シングルジョブ or バルクジョブ	同上	同上	同上

創薬計算で利用されているソフトの現状 (2)

	ソフト	計算規模 (計算時間は汎用計算サーバー、数10~数100コアを目安)	実行の形態	実行頻度	ユーザーの要望	その他
パーチャル化合物デザイン	市販、インハウス	10の数10乗の化学構造発生が可能だが、メモリも問題や、発生後の計算(ドッキング計算など)の制約から、最大でも数万の発生する程度	ドッキングプログラムなどのプラグインツールとして実行することが多い	数~数十テーマ×各社/年	・発生後の計算(ドッキング計算など)の処理速度が高速になれば、発生する化合物数も増加できる。 ・発生した化学構造の実際に合成可能かどうかの推定も重要(現有の反応可能性予測は、予測精度が良いとは言えない。)	
化合物電荷計算	市販: Gaussian, MOPAC フリー: GAMESS, MOPAC	分子量1000以下の1化合物のHF/RESP電荷計算には数分から数時間程度、AM1電荷計算では数秒程度。	シングルジョブ or バルクジョブ	同上	パーチャルスクリーニング目的などの大規模な化合物数の電荷計算には、簡易なGasteiger電荷やAM1電荷を利用しているが、計算速度がアップすれば、より正確なHF/RESP電荷計算の適用も可能になる。	
タンパク質立体構造モデリング	ホモロジーモデリングツール フリー: MODELLER, SWISS-MODEL等 市販: Mo, Pr等	1つのタンパク質について数秒から数分程度。GUIを使ってモデリングすることも多い	インタラクティブ	同上	タンパク質構造を目視しながらの作業が必要なので、GUIが具備されていることが必要である。また、その後の計算(ドッキングやMD)も同じGUIでシームレスに操作できることが望ましい。	ホモロジーモデリングと必要なので、そのものの結晶構造は無い場合に、類縁タンパクの構造を用いてモデリングする。
タンパク質マルチコンフォメーション生成	分子動力学計算 フリー: GROMACS 市販: AMBER等 国プロ: myPresto/cosgene	1つのタンパク質について数時間程度	バルクジョブ(数10ps~数nsのMD)	同上	・長時間MDへの期待 ・Platypus-REIN(レプリカ交換法)などにも期待	この後、複数のコンフォメーションモデルに対するドッキング計算を必要とするため、計算負荷の問題から現状では当該計算を行わない事も多い。
パーチャルスクリーニング: リガンドベースアプローチ	類似化合物検索(市販: Pi, Mo, Dr, Op, Ch等) ファーマコフォア検索(市販: Co, Li, Mo等) CGBVS(市販、フリー)	100万化合物のスクリーニングに、数時間程度 100万化合物のスクリーニングに、数時間程度	シングルジョブ or バルクジョブ	同上	現状では、最大でも数1000万化合物程度を対象とした計算規模であるが、計算速度がアップすれば、対象化合物数の増加や、複数の標的タンパク質に対する計算も可能になる。	
パーチャルスクリーニング: ドッキングシミュレーション	市販: Go, Gl, Mo等 フリー: AutoDock 国プロ: myPresto/sievgene	100万化合物のスクリーニングに、数時間から数日程度	シングルジョブ or バルクジョブ	同上	同上	京大(奥野G)らにより「京」アプリを開発

創薬計算で利用されているソフトの現状 (3)

	ソフト	計算規模 (計算時間は汎用計算サーバー、数10~数100コアを目安)	実行の形態	実行頻度	ユーザーの要望	その他
結合親和性予測	QM/MM (国プロ: Platypus、市販: Q等)	10数化合物の計算に数時間程度	シングルジョブ or バルクジョブ	数~数十テーマ×各社/年	<ul style="list-style-type: none"> 対象化合物数を増やすために、計算速度のアップが望まれる。 現場利用には、ドッキング計算の結果を開始構造とした計算が必須であるが、十分な検証がなされていない。 	
	FMO (フリー: GAMESS)	10数化合物の計算に数時間程度	シングルジョブ or バルクジョブ	同上	<ul style="list-style-type: none"> 対象化合物数を増やすために、計算速度のアップが望まれる。 現場利用には、ドッキング計算の結果を開始構造とした計算が必須であるが、十分な検証がなされていない。 	神大(北浦G)らにより「京」アプリを開発
	MM-PB/SA (市販: AMBER等)	10数化合物の計算に数時間から数日程度	シングルジョブ or バルクジョブ	同上	<ul style="list-style-type: none"> 対象化合物数を増やすために、計算速度のアップが望まれる。 現場利用には、ドッキング計算の結果を開始構造とした計算が必須であるが、十分な検証がなされていない。 	
	MP-CAFEE (フリー: GROMACS)	10数化合物の計算に数週間程度 (京をフル利用した場合は、5 or 日程度)	シングルジョブ or バルクジョブ	同上	<ul style="list-style-type: none"> 現場利用には計算速度の劇的アップが必須。 現場利用には、ドッキング計算の結果を開始構造とした計算が必須であるが、十分な検証がなされていない。 	東大(藤谷G)、京大(奥野G)らにより「京」アプリを開発
ADMET予測	市販: Ad等 インハウス: 回帰モデル、判別モデル	数10化合物の計算に数分程度	シングルジョブ	同上	<ul style="list-style-type: none"> 予測精度に難があり、現状では汎用的な利用されていない。 	
薬物動態予測	市販: Ga, Di等	数11化合物の計算に数分程度	シングルジョブ	同上	<ul style="list-style-type: none"> 予備実験無しで、新規な化合物の予測が望まれる。 	現在の方法論は、動物などを用いて初期パラメータを取得する実験が予め必要であり、実験無しで新規の化合物の予測はできない。
マルテスケールシミュレーション	—	—	—	—	<ul style="list-style-type: none"> 東大(杉浦G)らによる心臓シミュレーション(京アプリ)の薬物作用応用に期待。 心臓以外の臓器への適用にも期待。 	
ヒトへの外挿	—	—	—	—	同上	