

複数領域とタンパクコード領域の塩基配列データにおける塩基置換モデルの選択

東北大・院・生命科学 田辺昌史

Introduction

系統推定の現状

モデルベースの系統推定では分子進化モデルの選択が重要

↓
Modeltest (Posada & Crandall, 1998) の普及

現状の問題点

モデル選択では「検討対象外のモデル」は考慮されていない

↓
Modeltestの検討モデルで十分か?

タンパクコード領域の場合

- ・コドン位置ごとの置換速度の不均質性は十分考慮していない
- ・コドン位置ごとの置換速度の不均質性も検討すべきでは?
- ・コドン位置ごとに異なるモデルを当てはめることも検討すべきでは?

複数領域データの場合

- ・領域ごとに異なるモデルを適用したい場合、Modeltestでは手間がかかる
 - ・単一モデルを当てはめる場合、置換速度の不均質性の検討が不十分では?
- Kakusan (Tanabe, 2007) · Kakusan2の開発へ

Features of Kakusan2

- ・FASTA、PHYLIP、NEXUS、GenBank形式のファイルに対応
- ・Windows / MacOS X / その他UNIXのいずれでも動作
- ・basemlもしくはPAUP*を用いた並列化尤度計算
- ・多数の領域の塩基配列データに対してモデル選択を一度に実行
- ・AIC / AICc / BICを用いたモデル選択
- ・MrBayes3 / PAUP*用簡単設定ファイルを出力
- ・タンパクコード領域に適したいくつかのモデルに対応
- ・Kakusan2+basemlなら、タダ

対応モデル

塩基置換確率行列…Modeltest · Kakusan1 · Kakusan2で共通
置換速度の不均質性

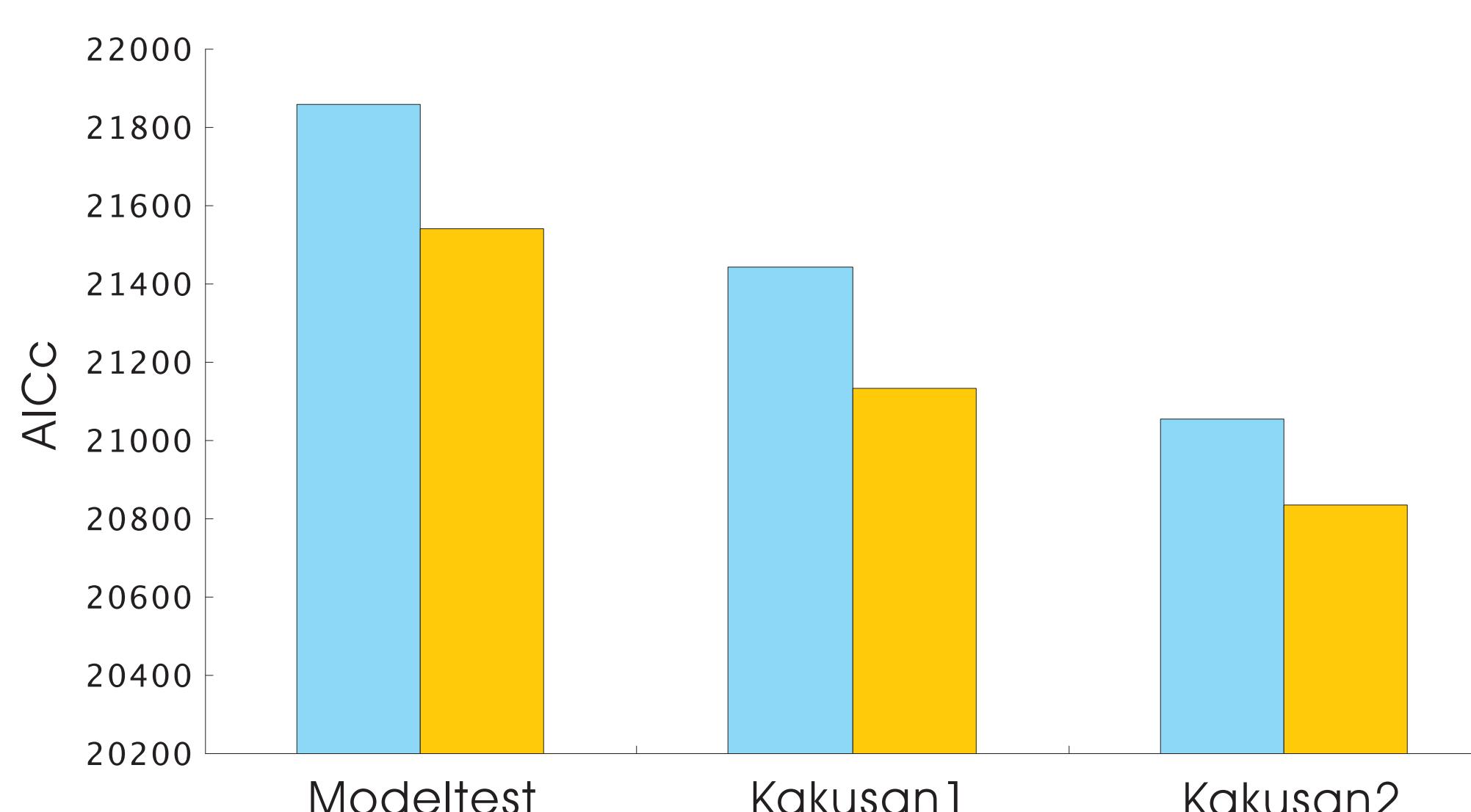
Homogeneous Gamma	Modeltest	Proportion of Invariant Gamma+Invariant
Codon Position/Partition Specific Rate		Kakusan1 (= ^{Kakusan2} _{+PAUP*})
Autocorrelated Gamma		
Codon Position/Partition Specific Rate+1Gamma		
Codon Position/Partition Specific Rate+NGamma		
Codon Position/Partition Specific Rate+1AGamma		
Codon Position/Partition Specific Rate+NAGamma		

Kakusan2+baseml

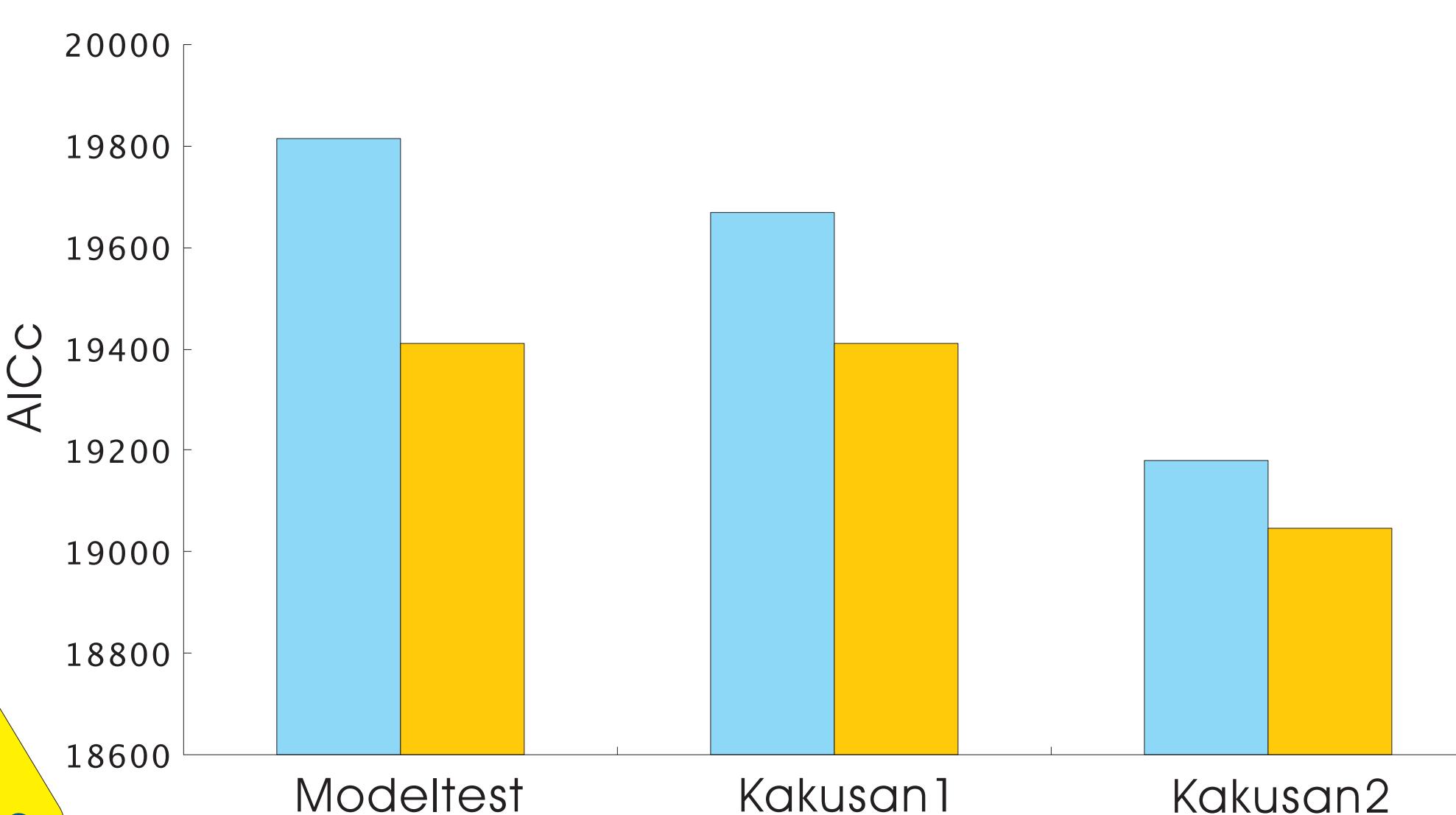
Comparison of AICc with previous softwares

- Materials Case1. ゴマシジミのCOI · COII · EF-1 α 領域の塩基配列 (Als *et al.*, 2004)
Case2. オナガザルのCOI · COII · α 1,3-GT · TSPY領域の塩基配列 (Davenport *et al.*, 2006)
Case3. 力ササギヒタキのND2 · Myo intron2領域の塩基配列 (Filardi & Moyle, 2005)
Case4. トカゲ · ヘビのC-mos · HOXA13 · R35 · RAG1 · RAG2領域塩基配列 (Fry *et al.*, 2006)
- Methods · Modeltest / Kakusan1 (=Kakusan2+PAUP*) / Kakusan2+baseml
を用いて最適なsingle modelとproportional modelをそれぞれ選択する
· モデル選択基準にはAICcを用いる
· 選択されたモデルを適用してMrBayes3で解析
· サンプリングされた全ての系統樹においてAICcを計算し、最小値を比較する

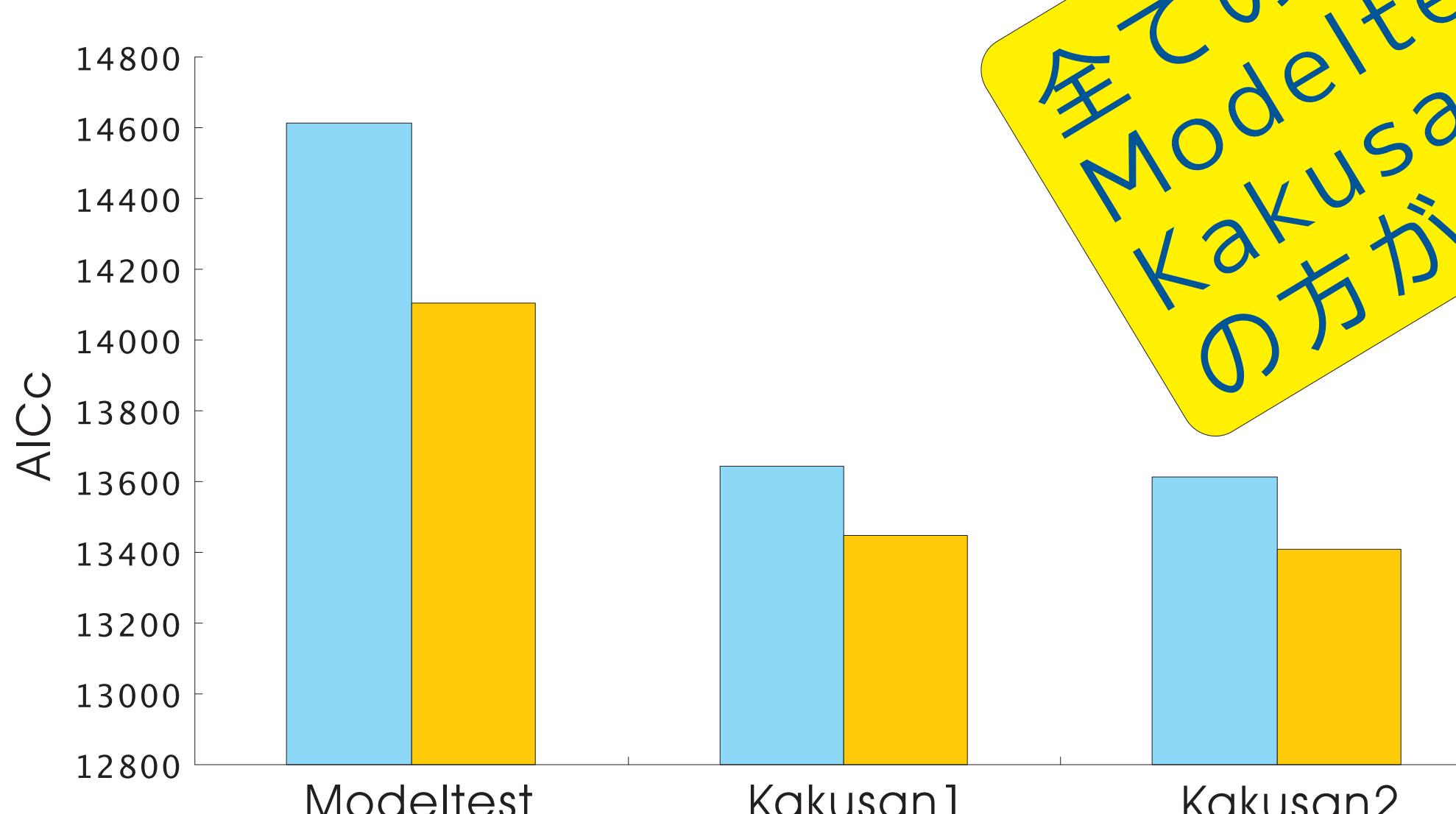
Case1: Als *et al.* (2004)



Case3: Filardi & Moyle (2005)

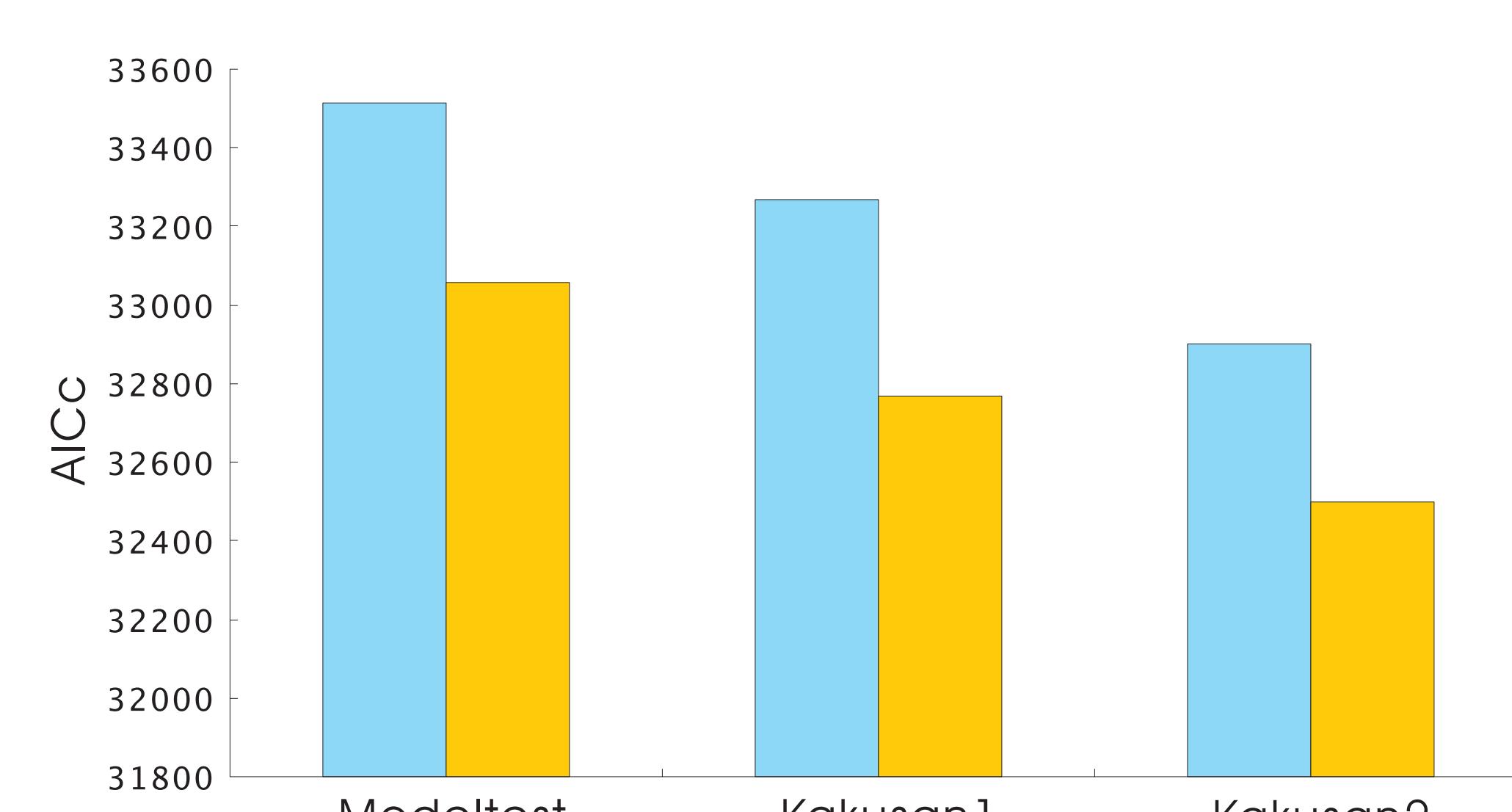


Case2: Davenport *et al.* (2006)



全てのデータセットで
ModeltestよりKakusan1、
Kakusan1よりKakusan2
の方が良好な結果を得られた。

Case4: Fry *et al.* (2006)



Future work

- ・候補モデルの事前選別機能の強化
- ・Treefinder (Jobb, 2004) 用設定ファイル出力への対応
- ・ステップダウン法によるモデル選択への対応
- ・rRNA/tRNA領域用モデルやコドン置換モデルへの対応

ただいま
検討中・・・