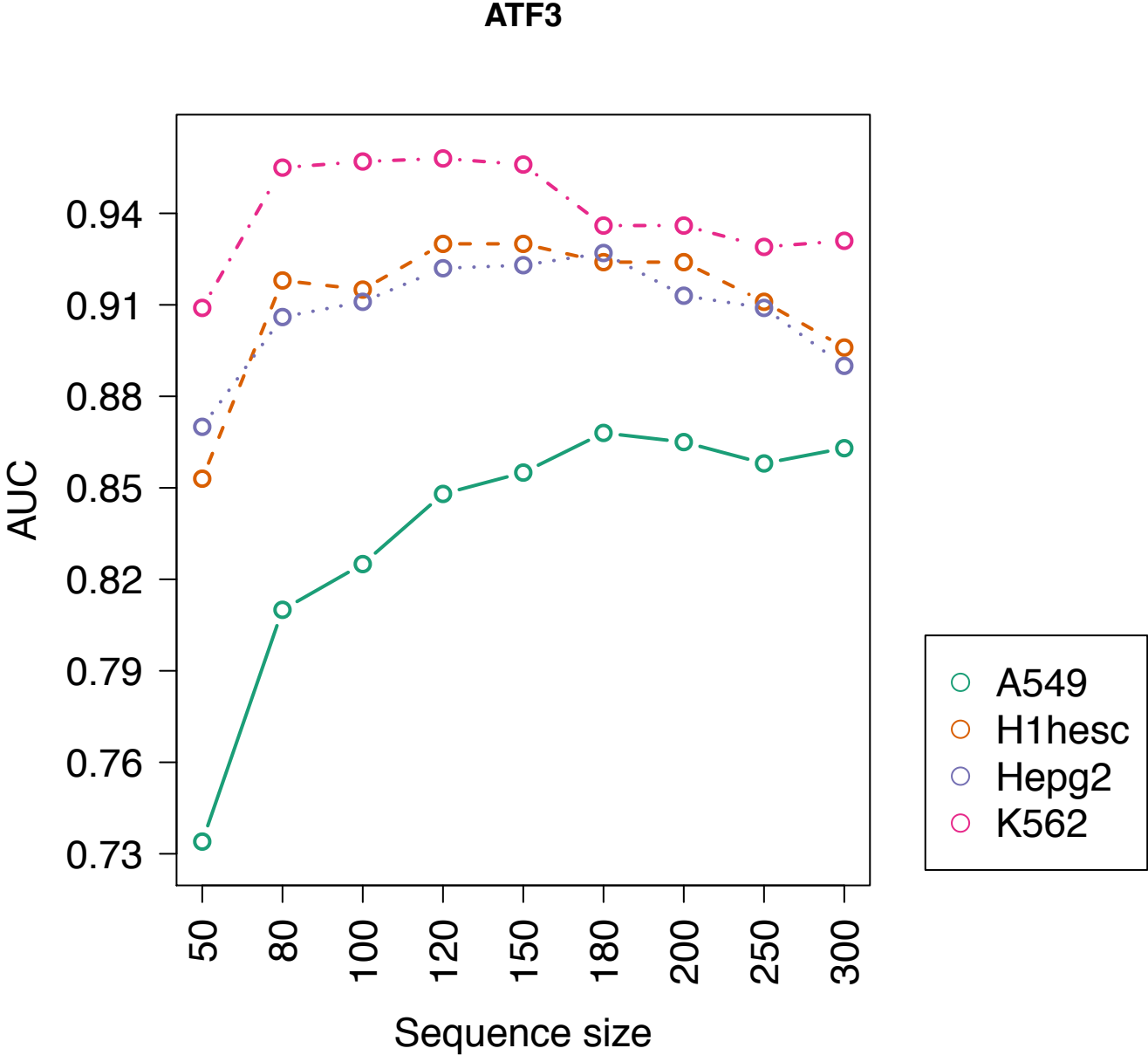
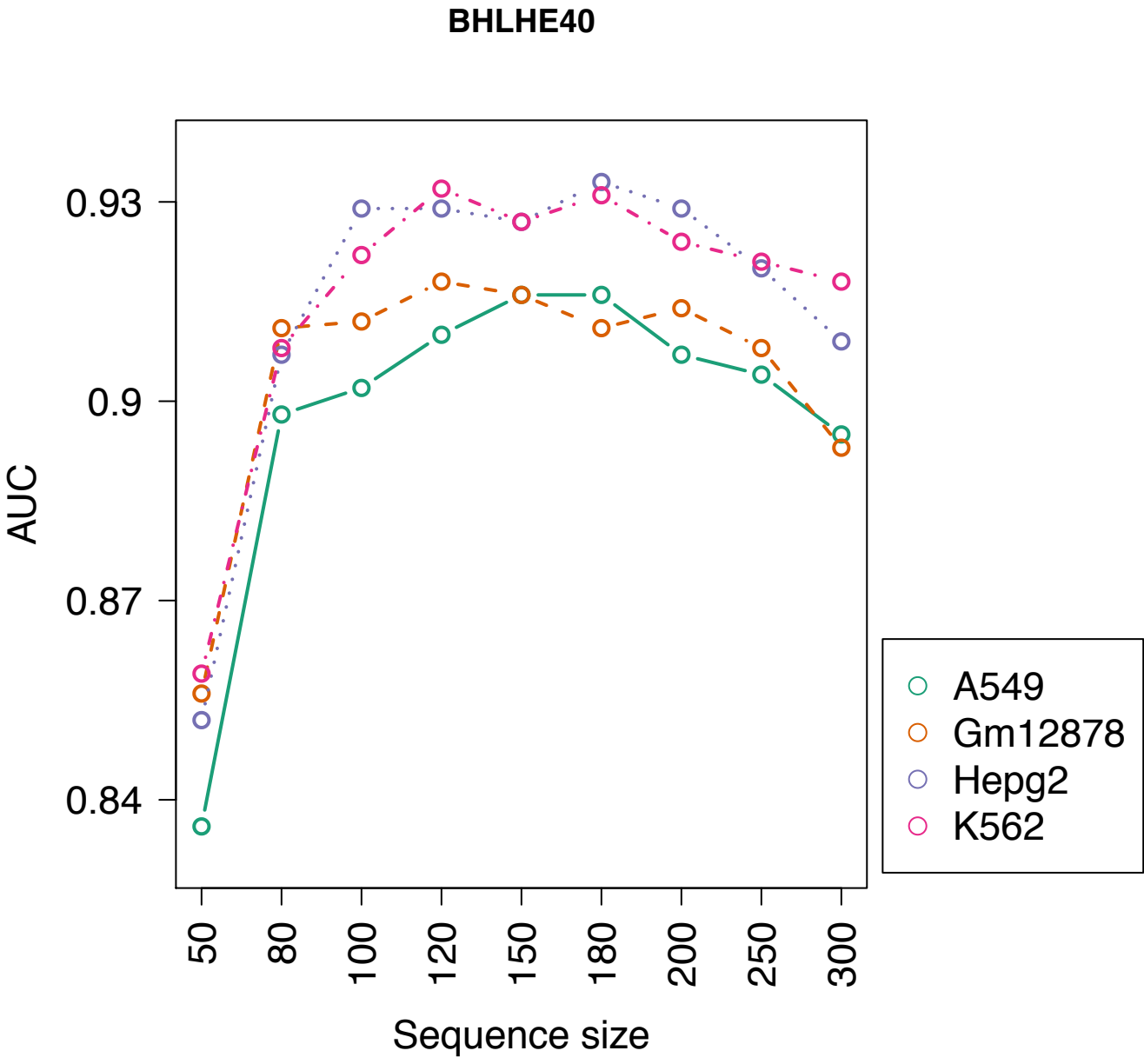
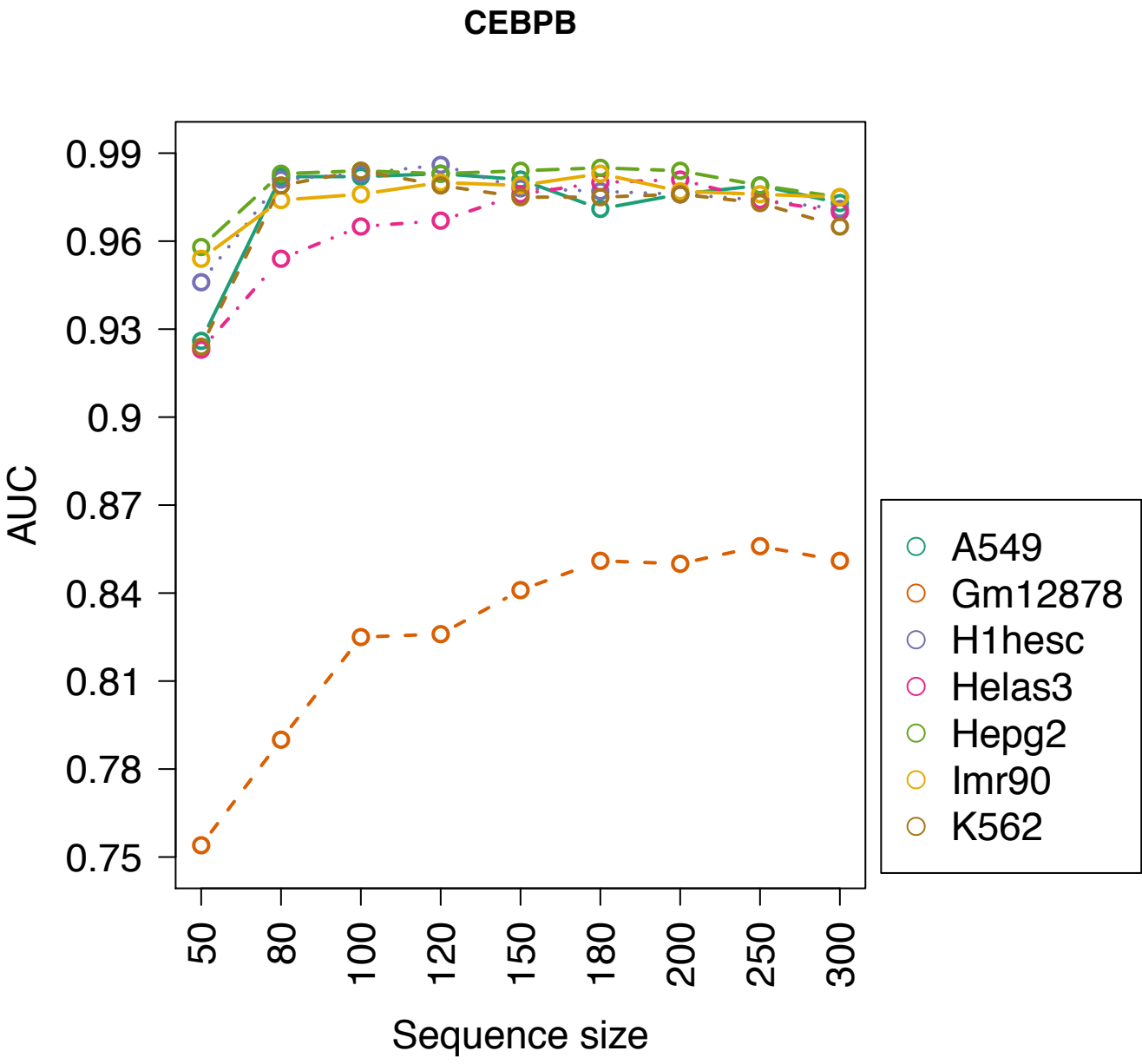
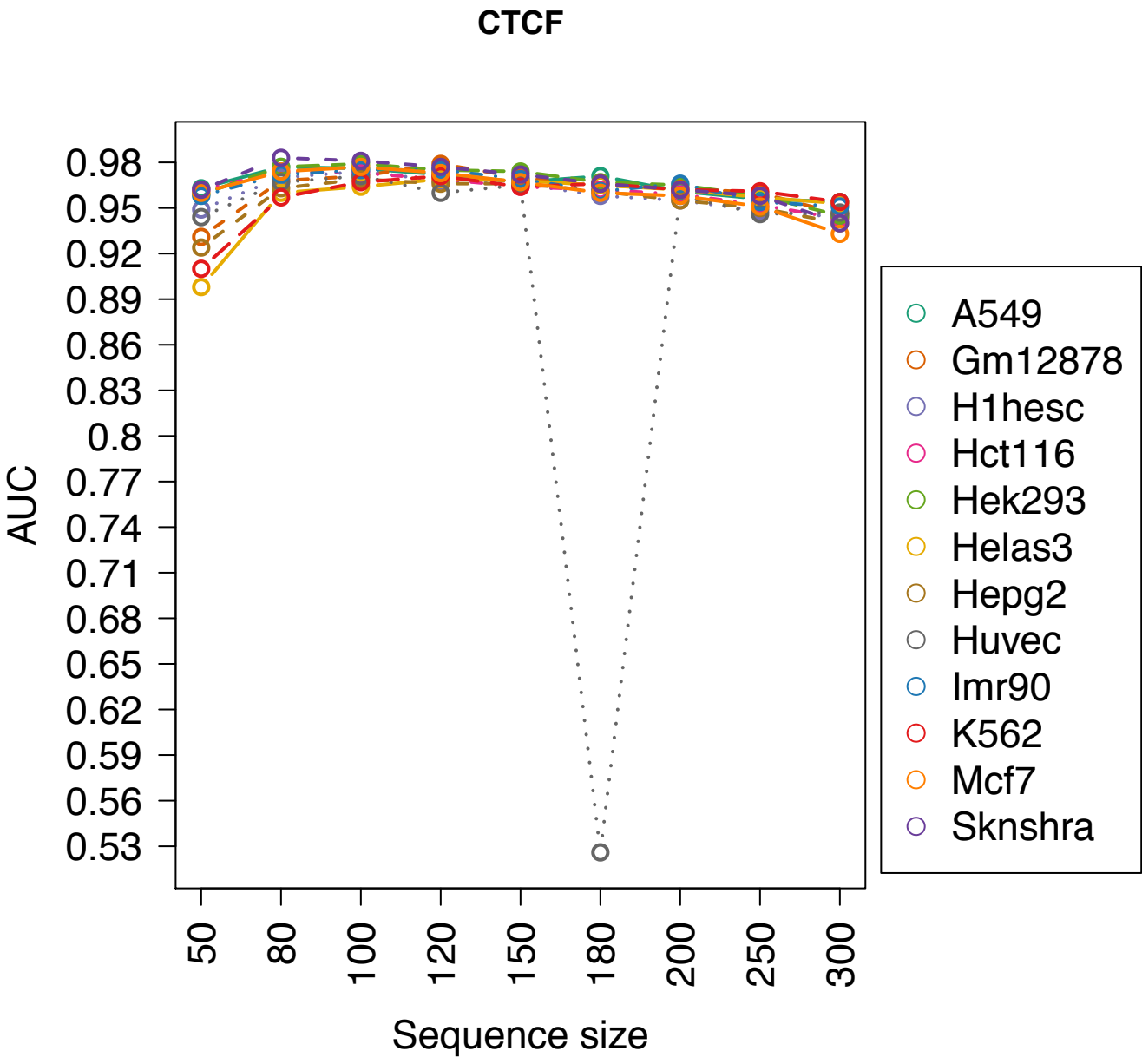
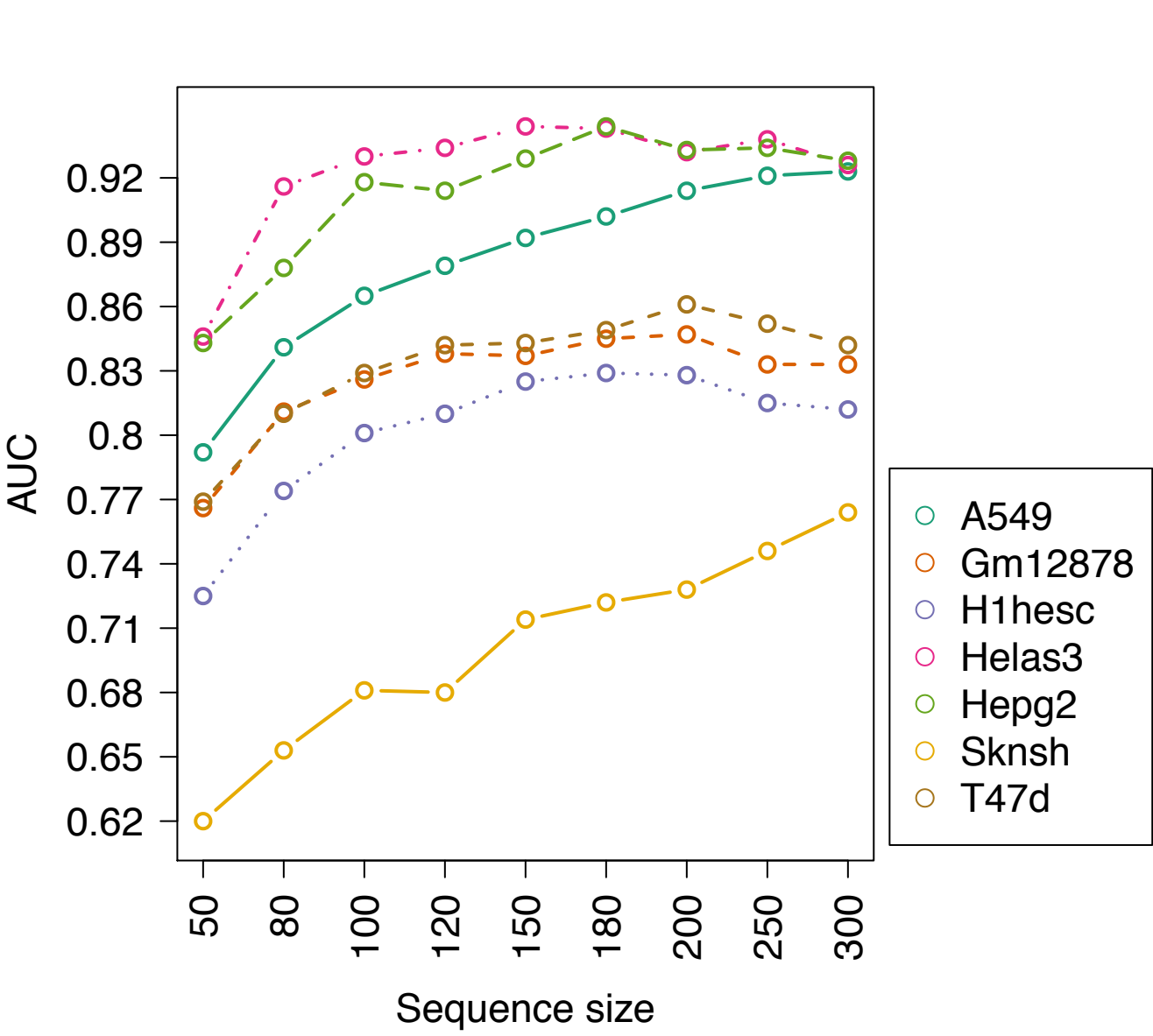
Figure S8. Relationship between model accuracy and sequence size. In each plot, color is used to indicate models from different cell lines.

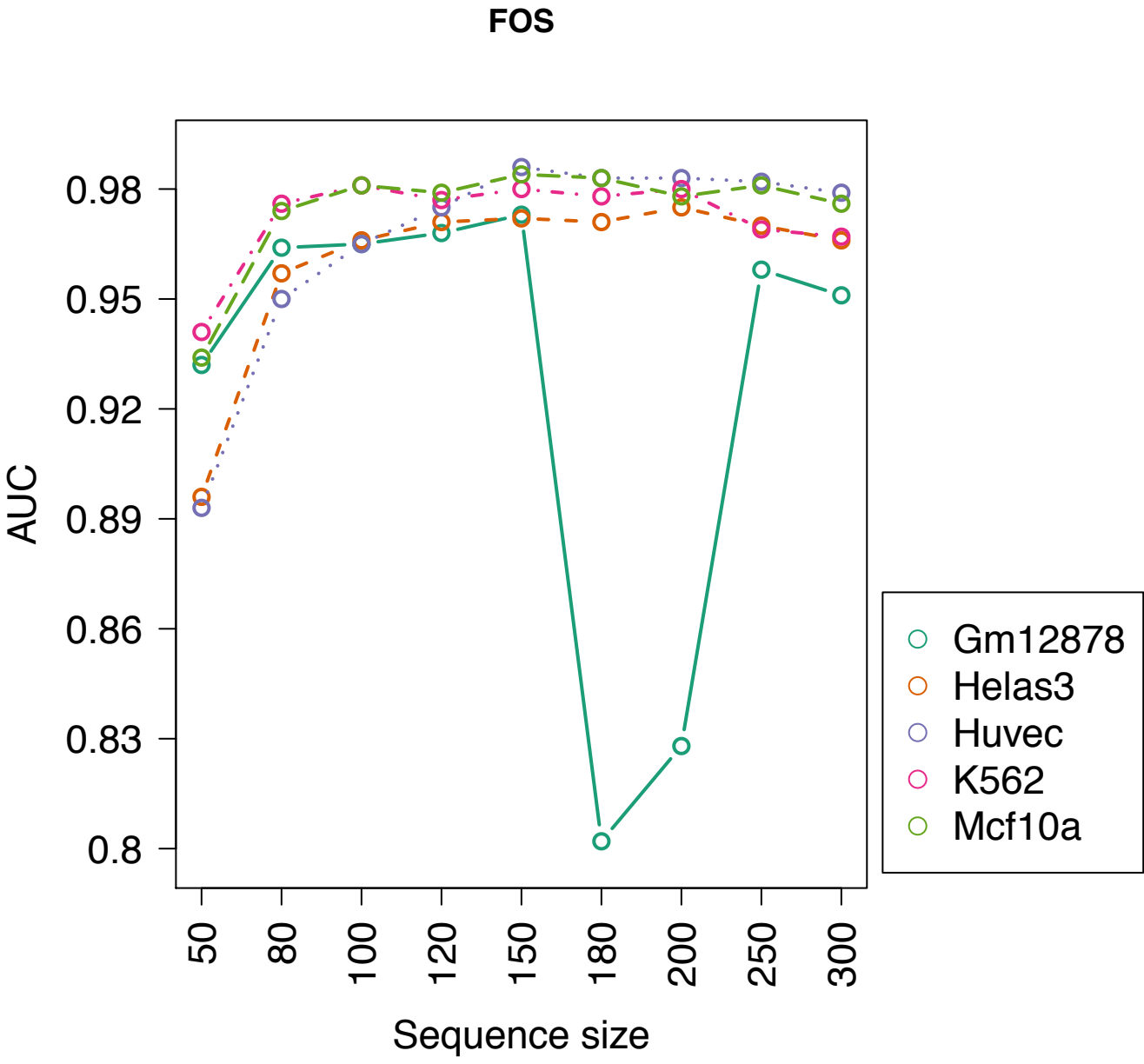

ATF3
AUC
0.94
0.91
0.88
0.85
0.82
0.79
0.76
0.73
50
80
100
120
150
180
200
250
300
Sequence size
A549
H1hesc
Hepg2
K562

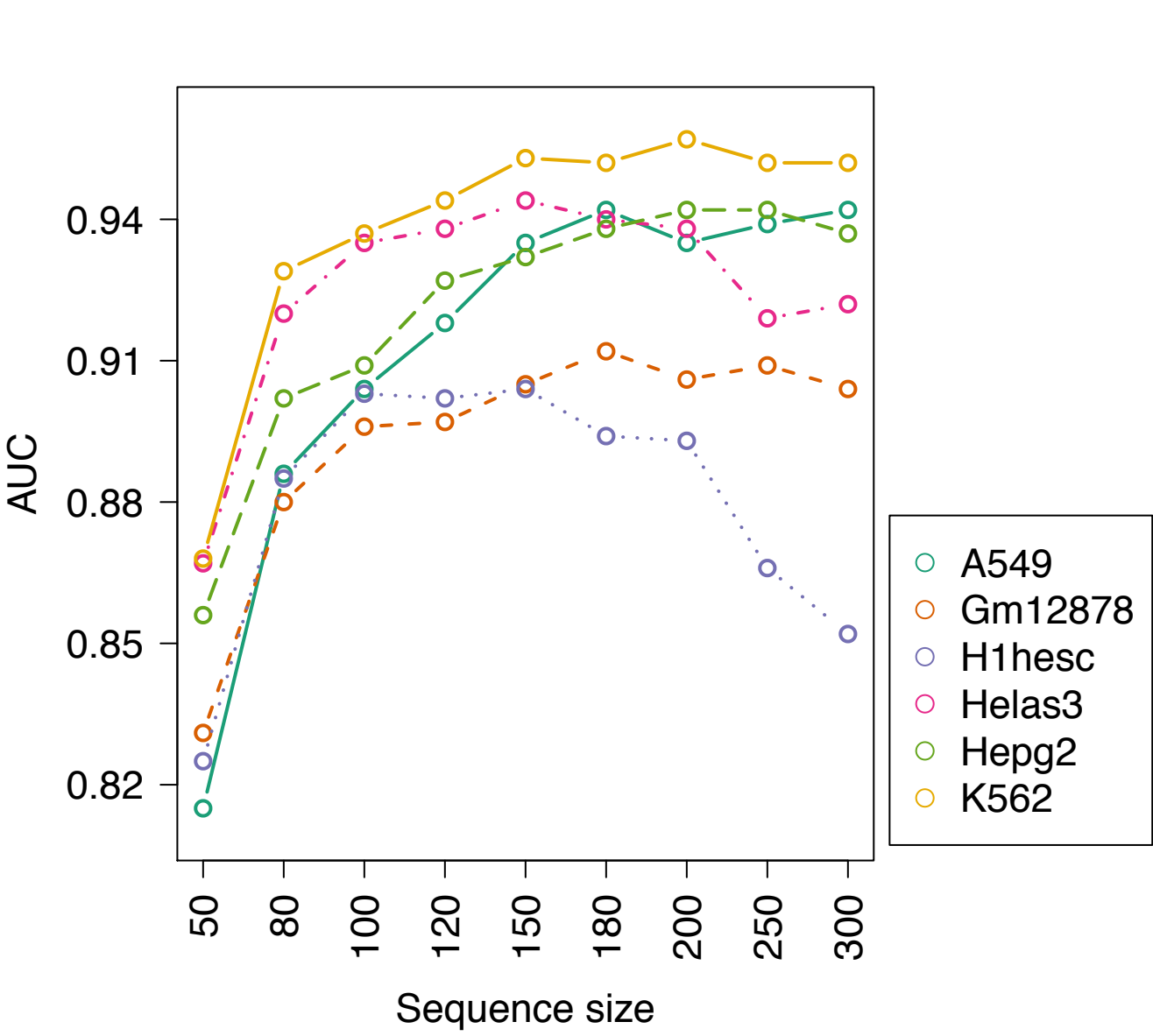

BHLHE40
AUC
0.93
0.9
0.87
0.84
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
Hepg2
K562

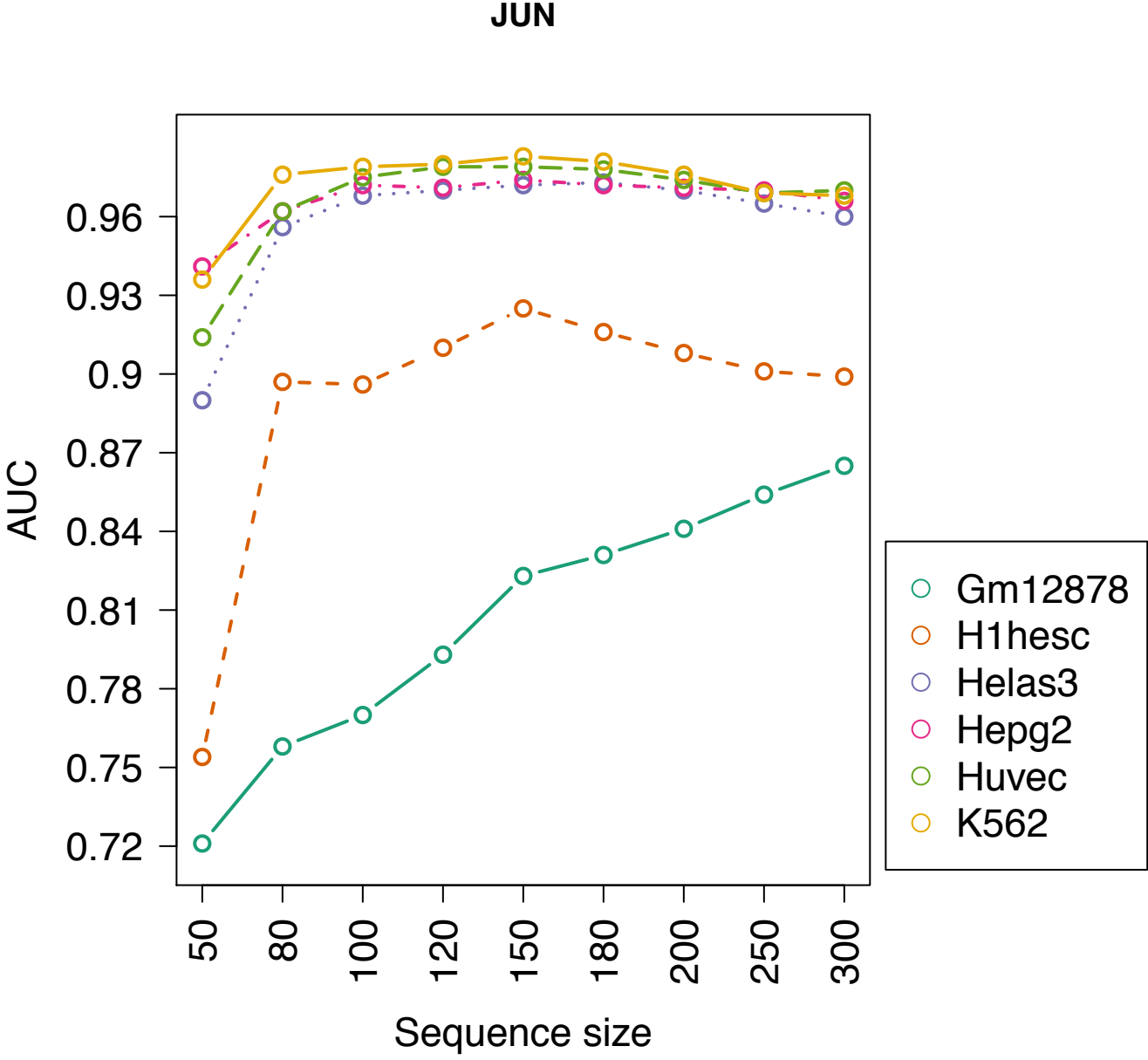

CEBPB
AUC
0.99
0.96
0.93
0.9
0.87
0.84
0.81
0.78
0.75
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Helas3
Hepg2
Imr90
K562

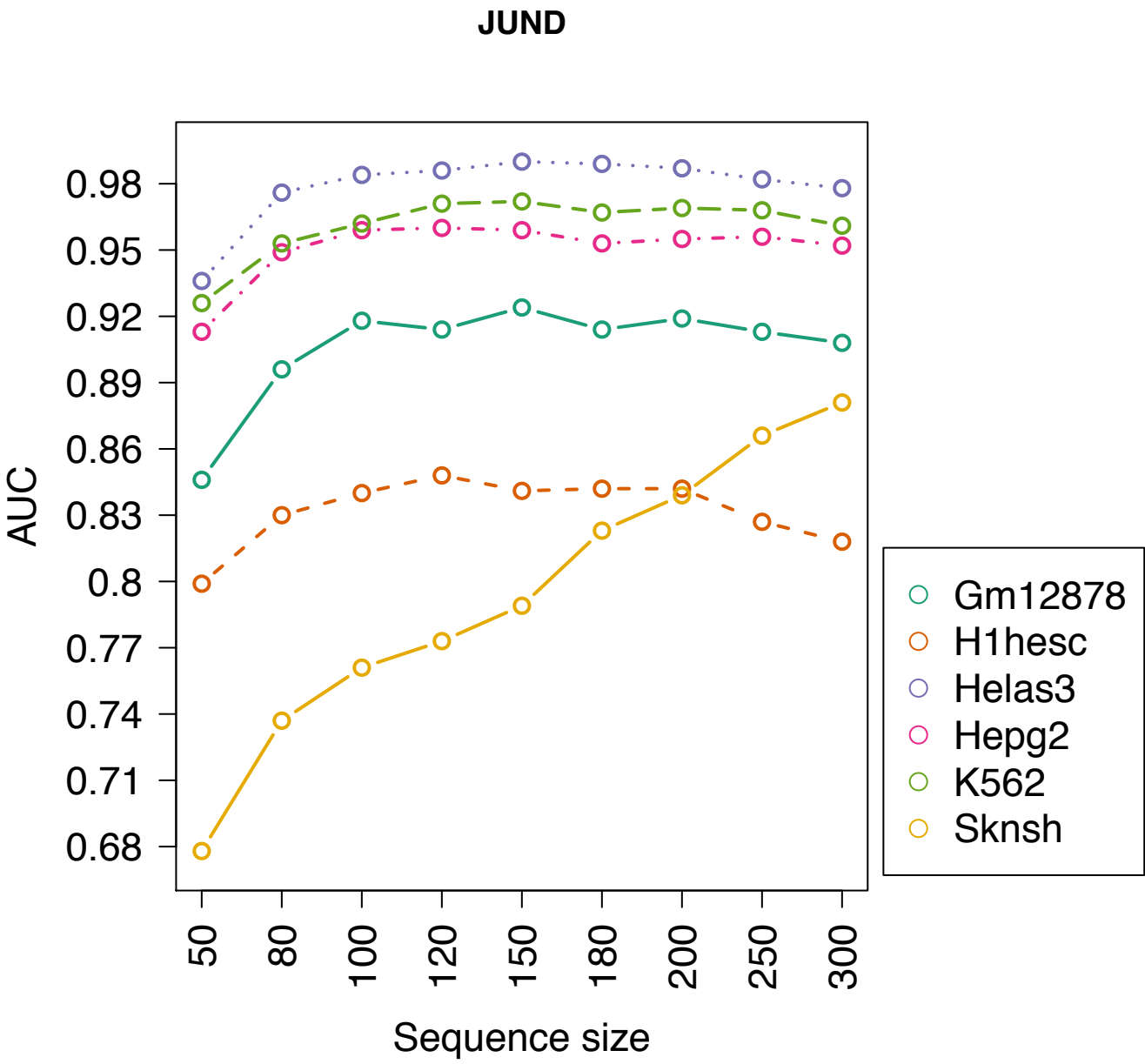

CTCF
AUC
0.98
0.95
0.92
0.89
0.86
0.83
0.8
0.77
0.74
0.71
0.68
0.65
0.62
0.59
0.56
0.53
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Hct116
Hek293
Helas3
Hepg2
Huvec
Imr90
K562
Mcf7
Sknshra

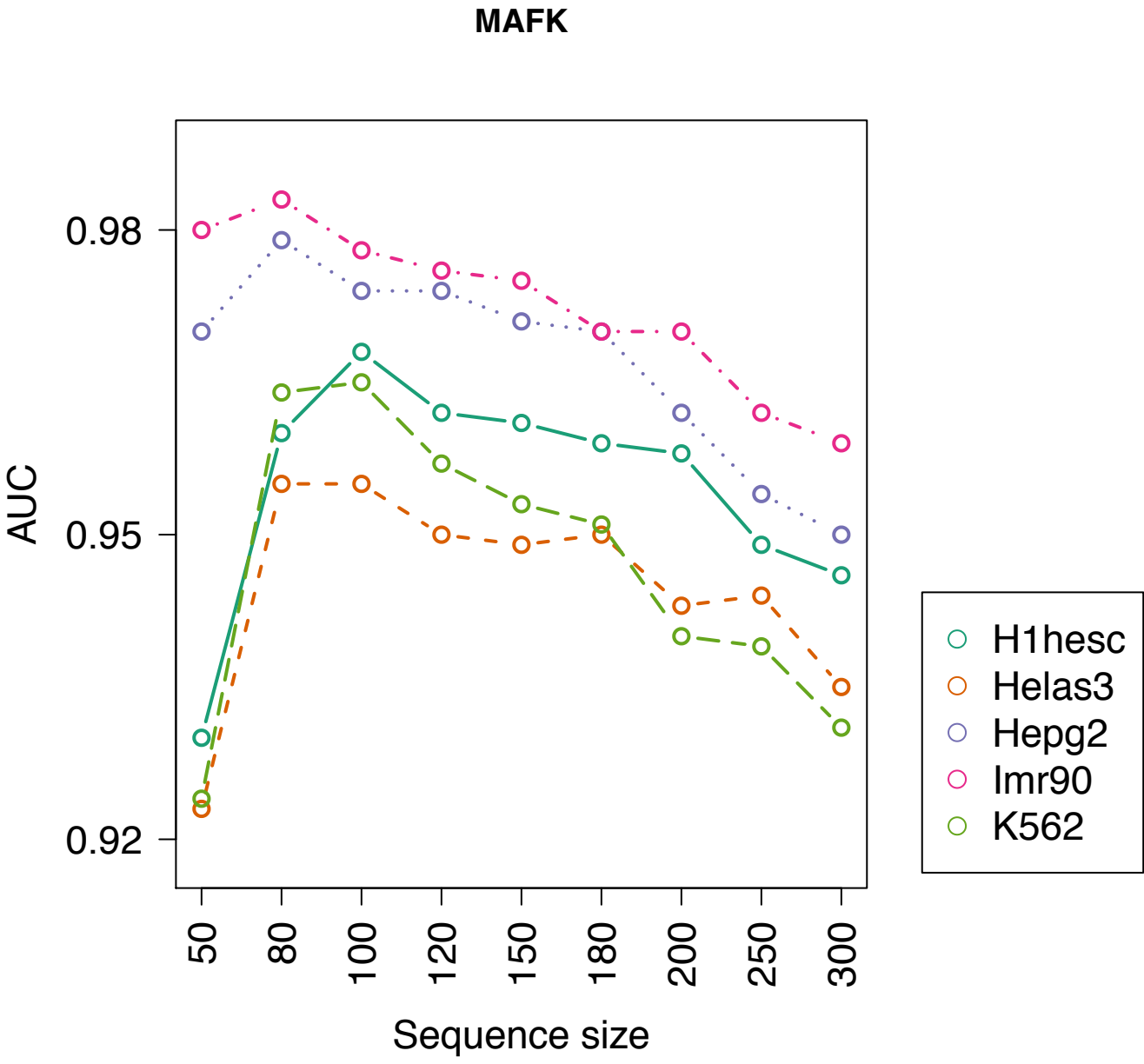

EP300
AUC
0.92
0.89
0.86
0.83
0.8
0.77
0.74
0.71
0.68
0.65
0.62
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Helas3
Hepg2
Sknsh
T47d

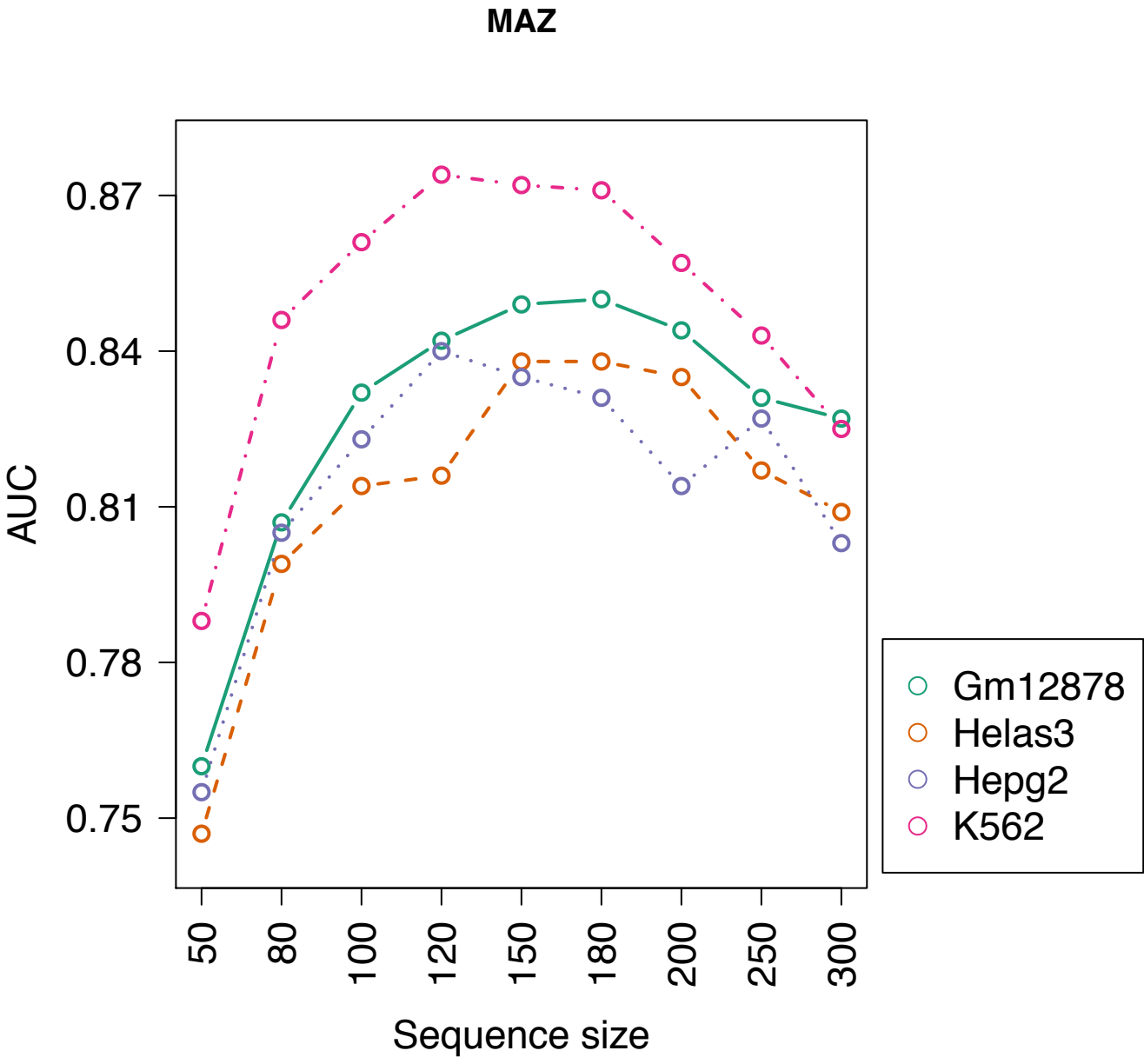

FOS
AUC
0.98
0.95
0.92
0.89
0.86
0.83
0.8
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
Helas3
Huvec
K562
Mcf10a


GABPA
AUC
0.94
0.91
0.88
0.85
0.82
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Helas3
Hepg2
K562

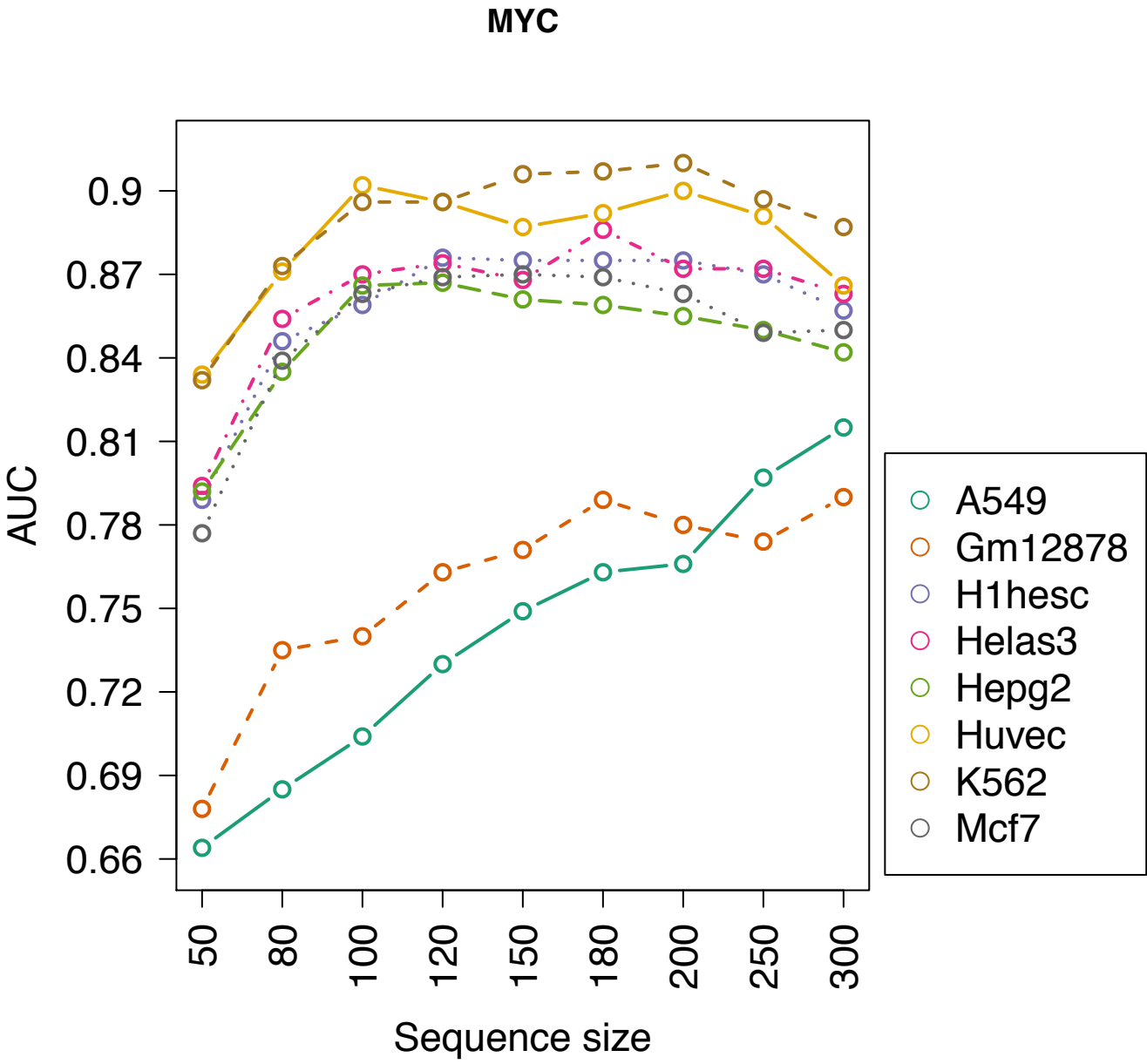

JUN
AUC
0.96
0.93
0.9
0.87
0.84
0.81
0.78
0.75
0.72
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
H1hesc
Helas3
Hepg2
Huvec
K562

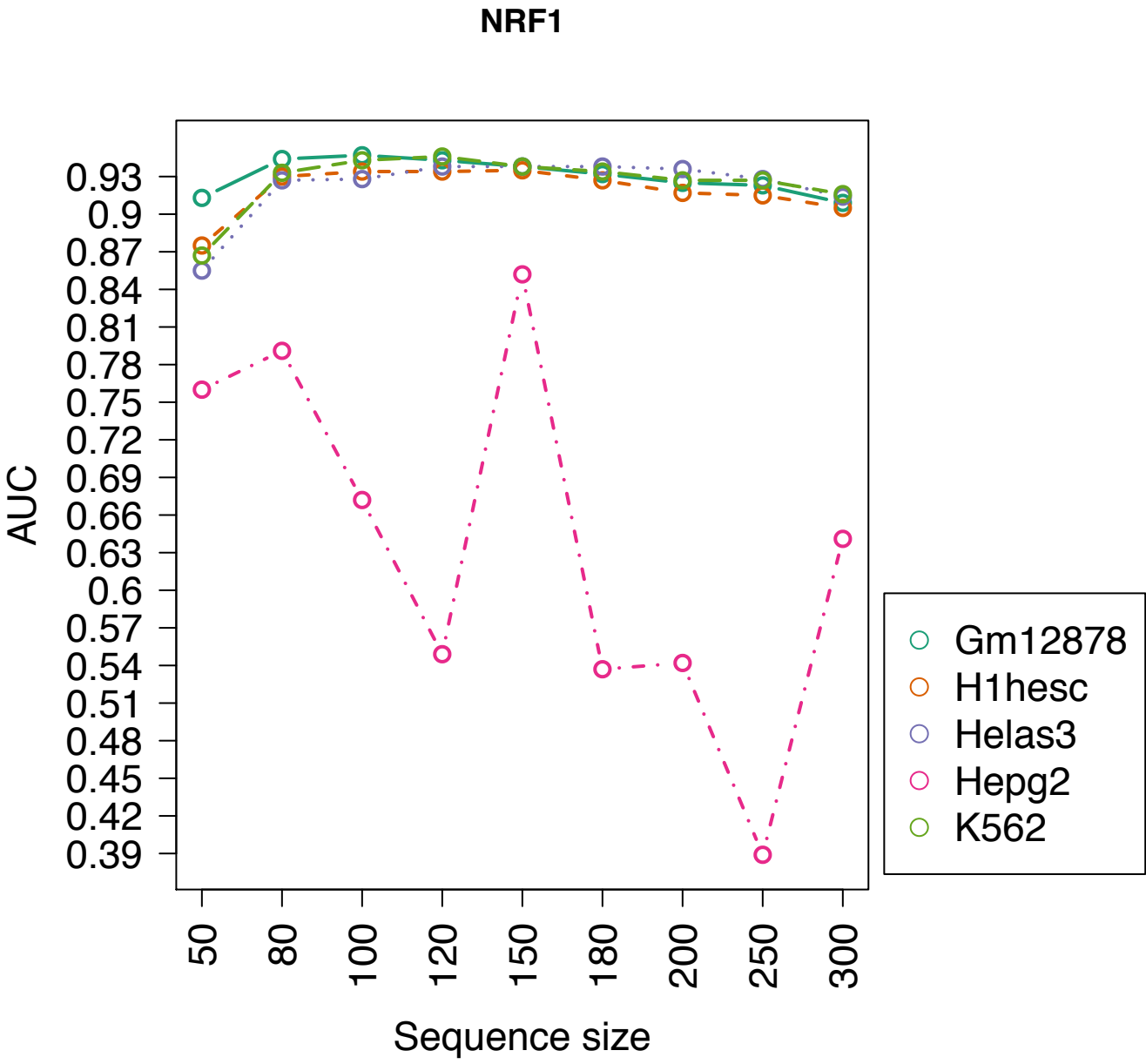

JUND
AUC
0.98
0.95
0.92
0.89
0.86
0.83
0.8
0.77
0.74
0.71
0.68
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
H1hesc
Helas3
Hepg2
K562
Sknsh

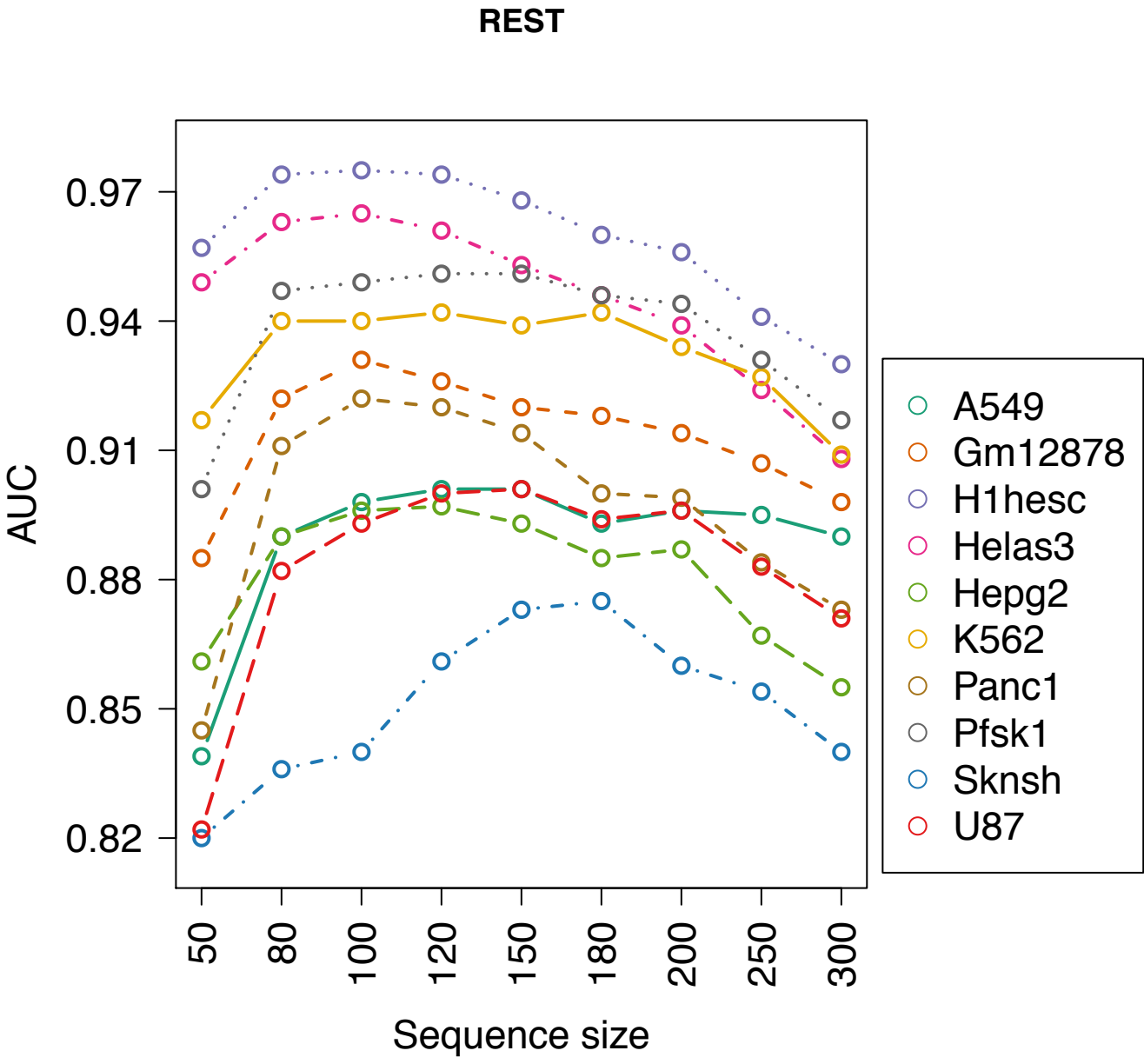

MAFK
AUC
0.98
0.95
0.92
50
80
100
120
150
180
200
250
300
Sequence size
H1hesc
Helas3
Hepg2
Imr90
K562

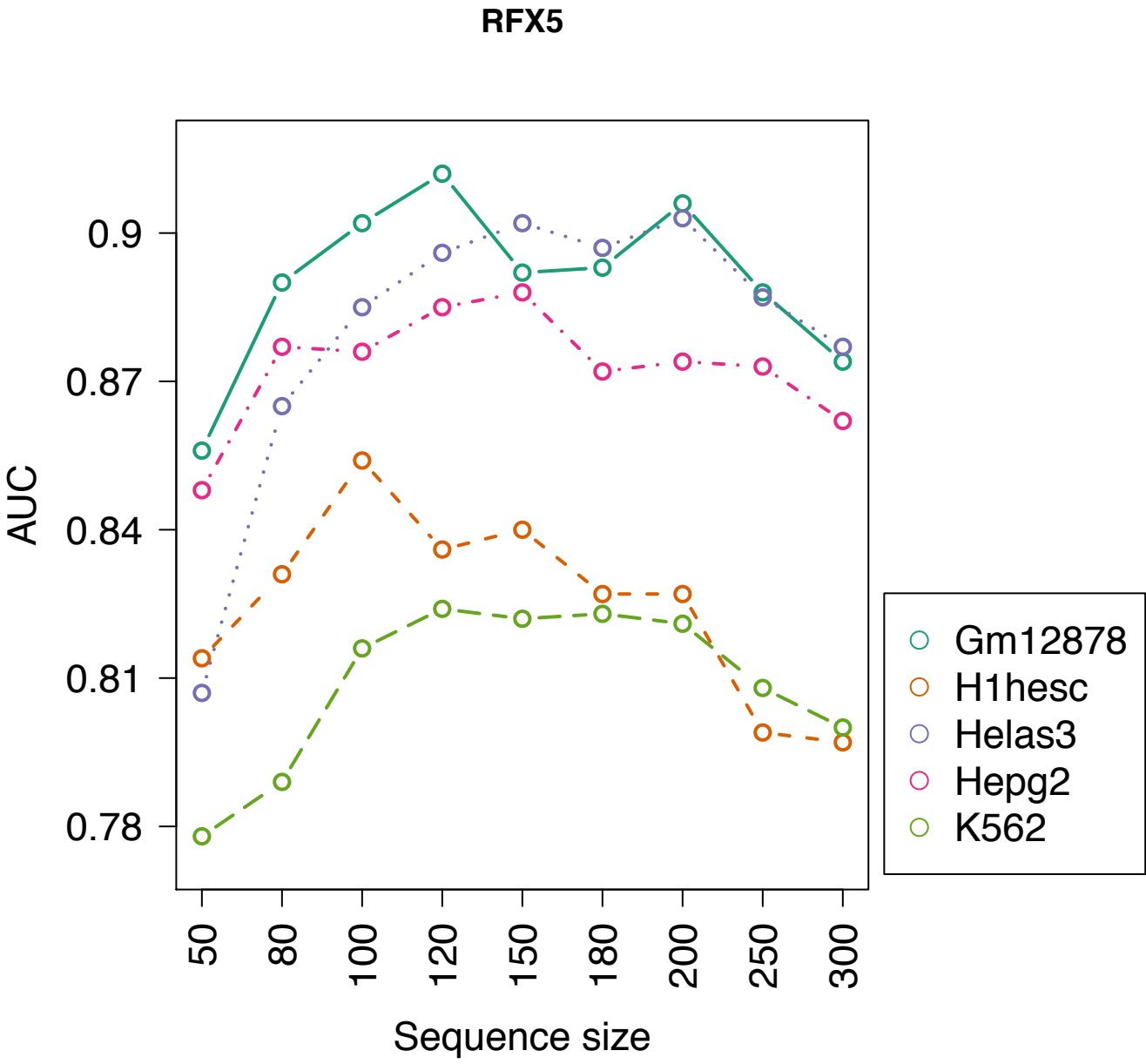

MAZ
AUC
0.87
0.84
0.81
0.78
0.75
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
Helas3
Hepg2
K562

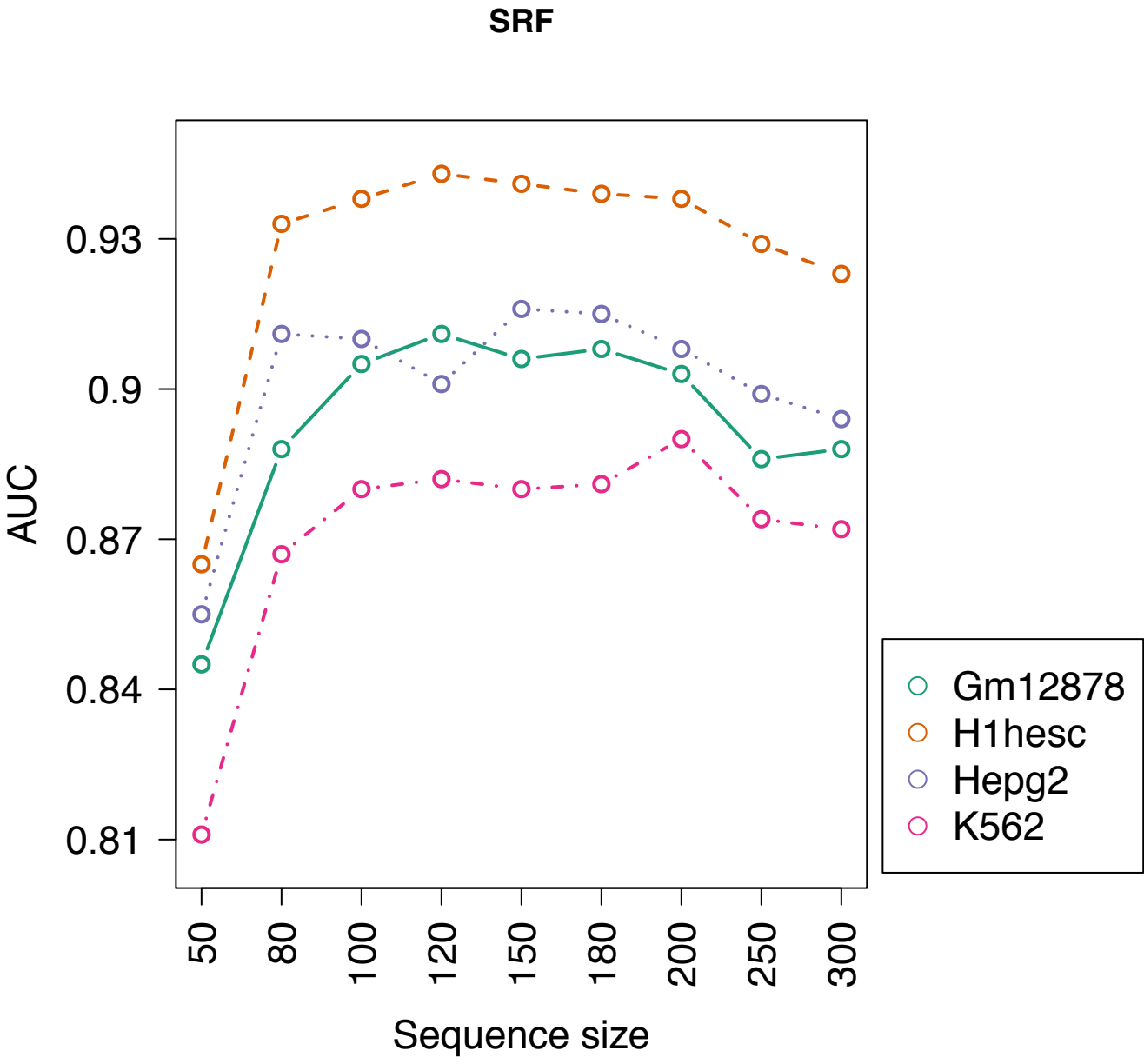
MXI1
AUC
0.88
0.85
0.82
0.79
0.76
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
H1hesc
Helas3
Hepg2
K562

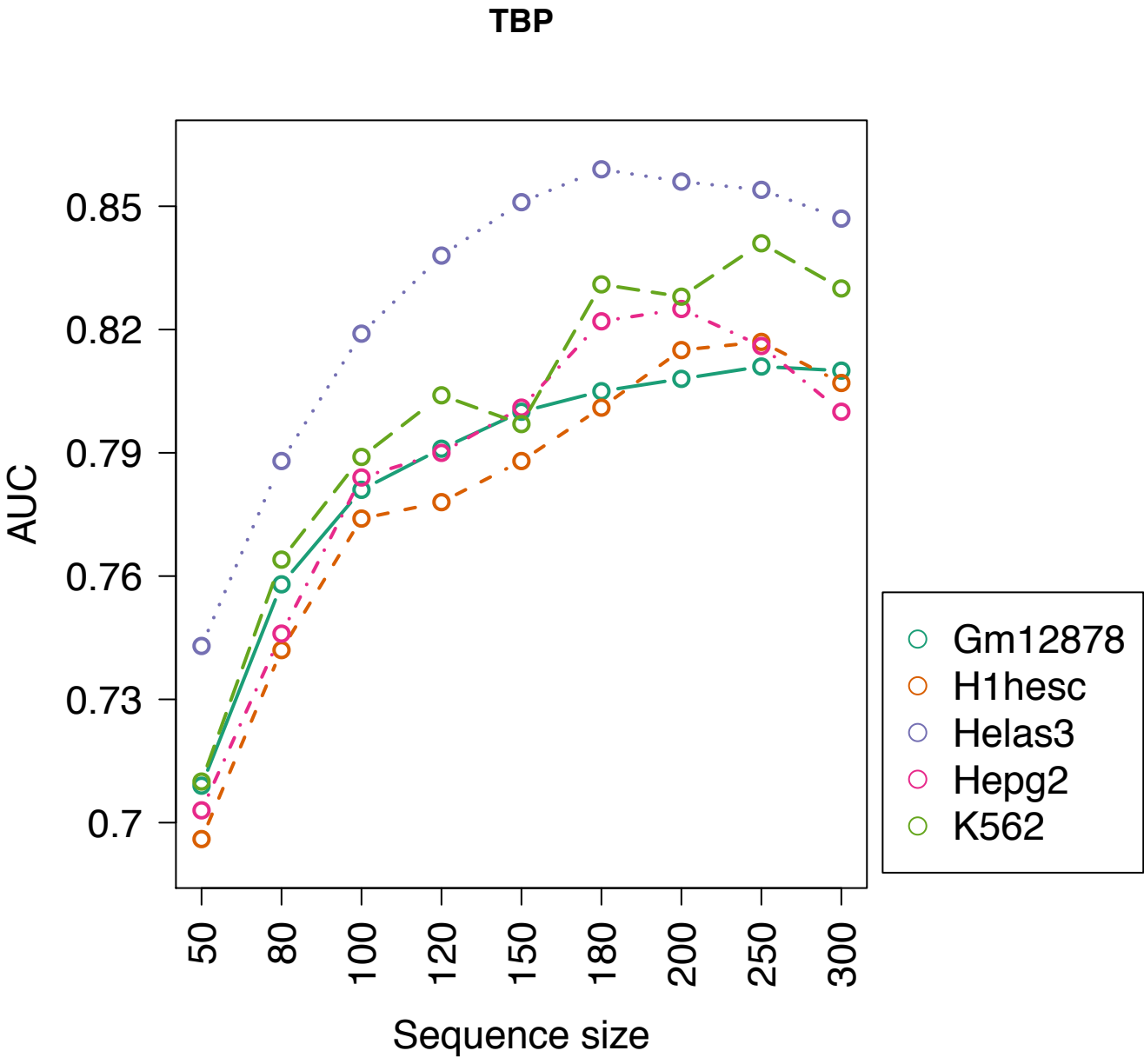

MYC
AUC
0.9
0.87
0.84
0.81
0.78
0.75
0.72
0.69
0.66
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Helas3
Hepg2
Huvec
K562
Mcf7

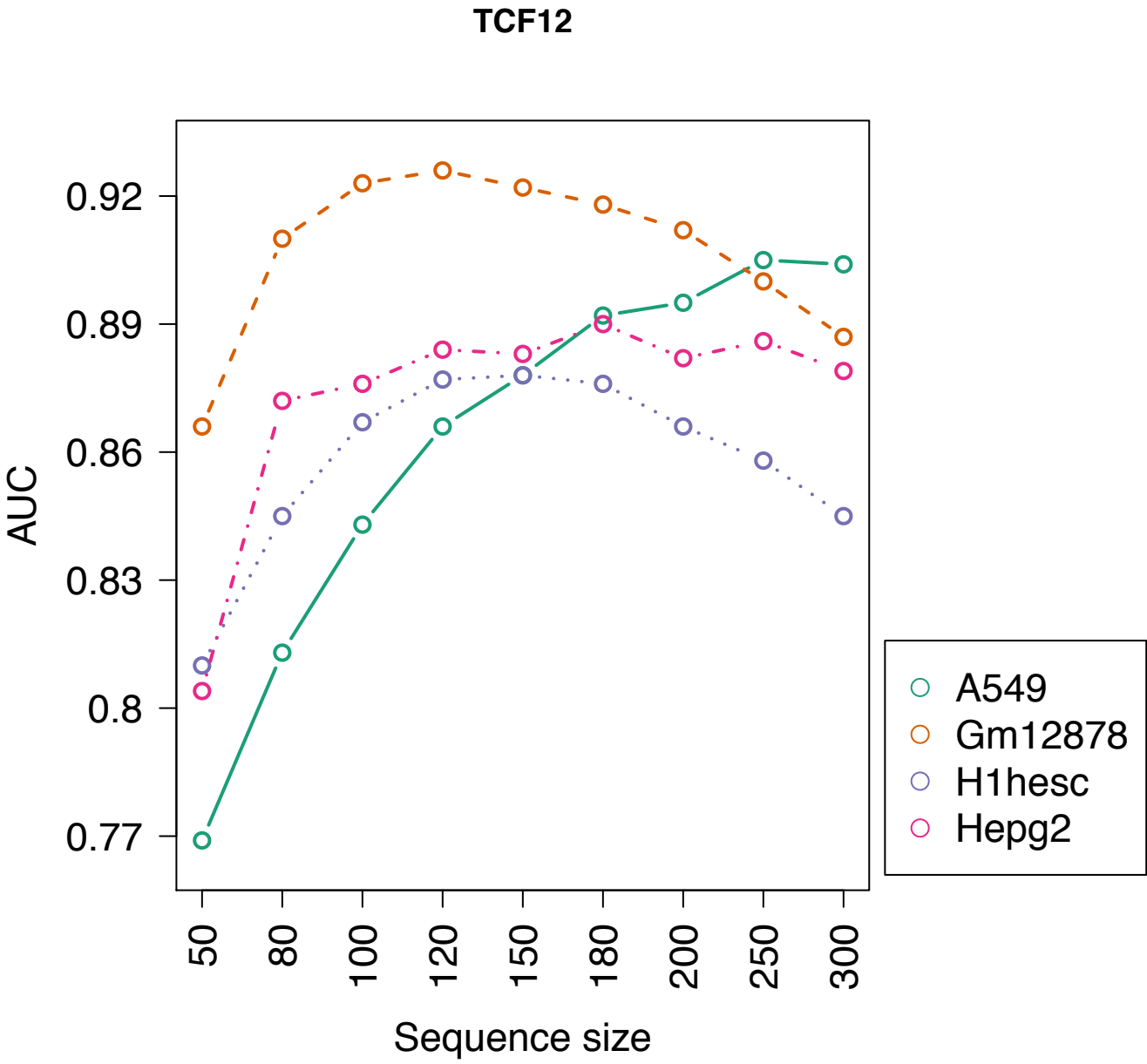

NRF1
AUC
0.93
0.9
0.87
0.84
0.81
0.78
0.75
0.72
0.69
0.66
0.63
0.6
0.57
0.54
0.51
0.48
0.45
0.42
0.39
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
H1hesc
Helas3
Hepg2
K562

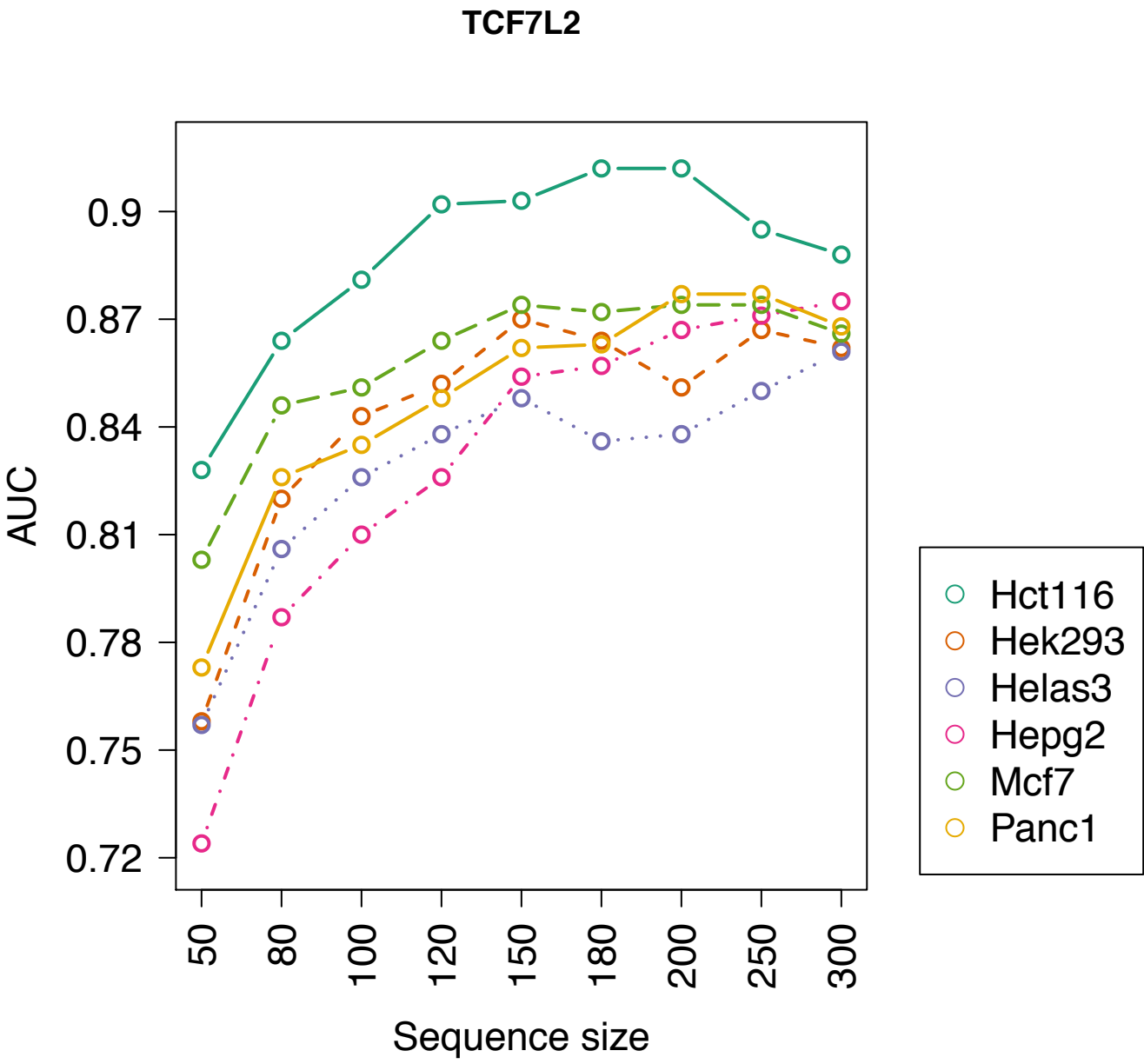

REST
AUC
0.97
0.94
0.91
0.88
0.85
0.82
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Helas3
Hepg2
K562
Panc1
Pfsk1
Sknsh
U87

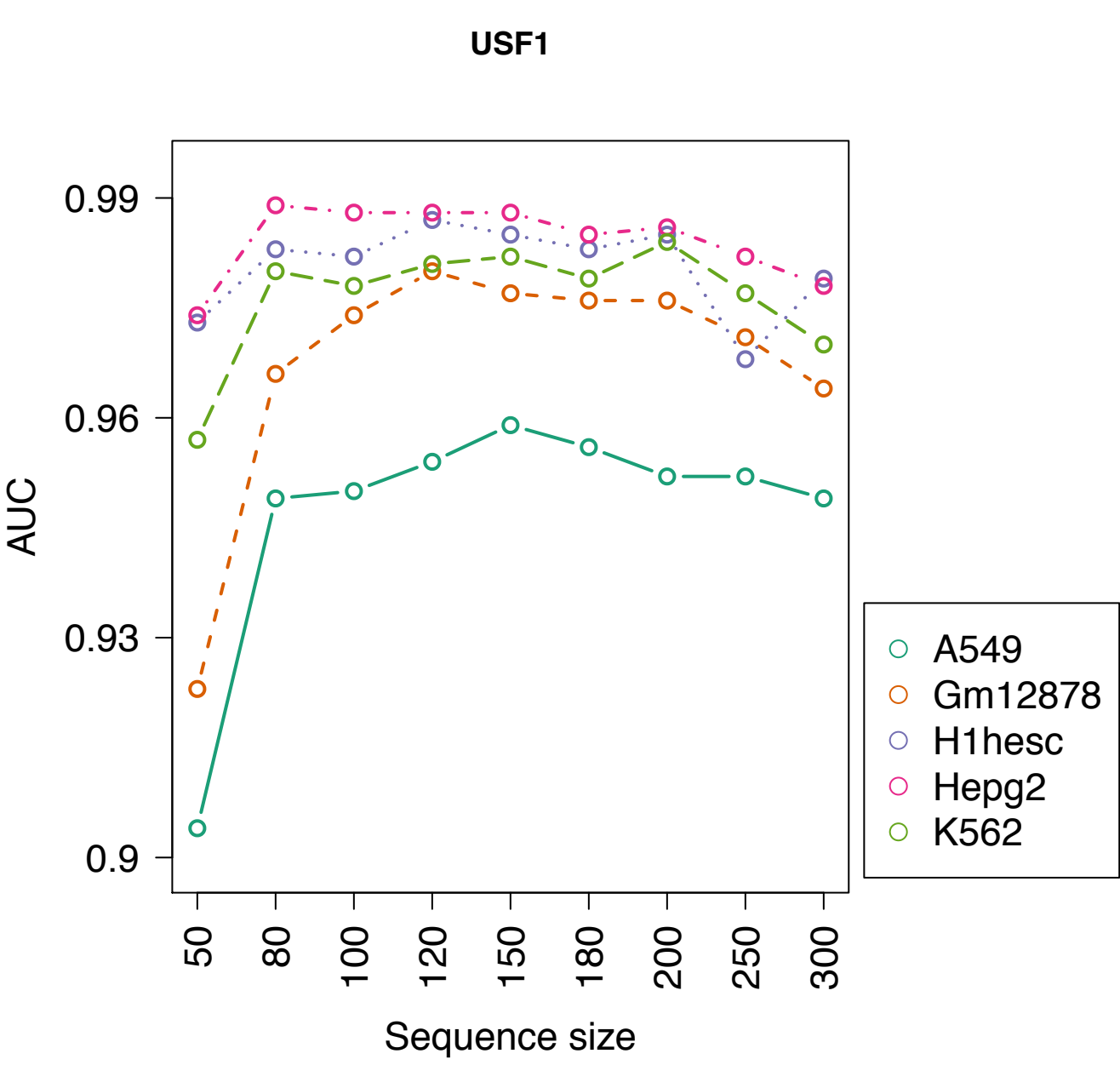

RFX5
AUC
0.9
0.87
0.84
0.81
0.78
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
H1hesc
Helas3
Hepg2
K562

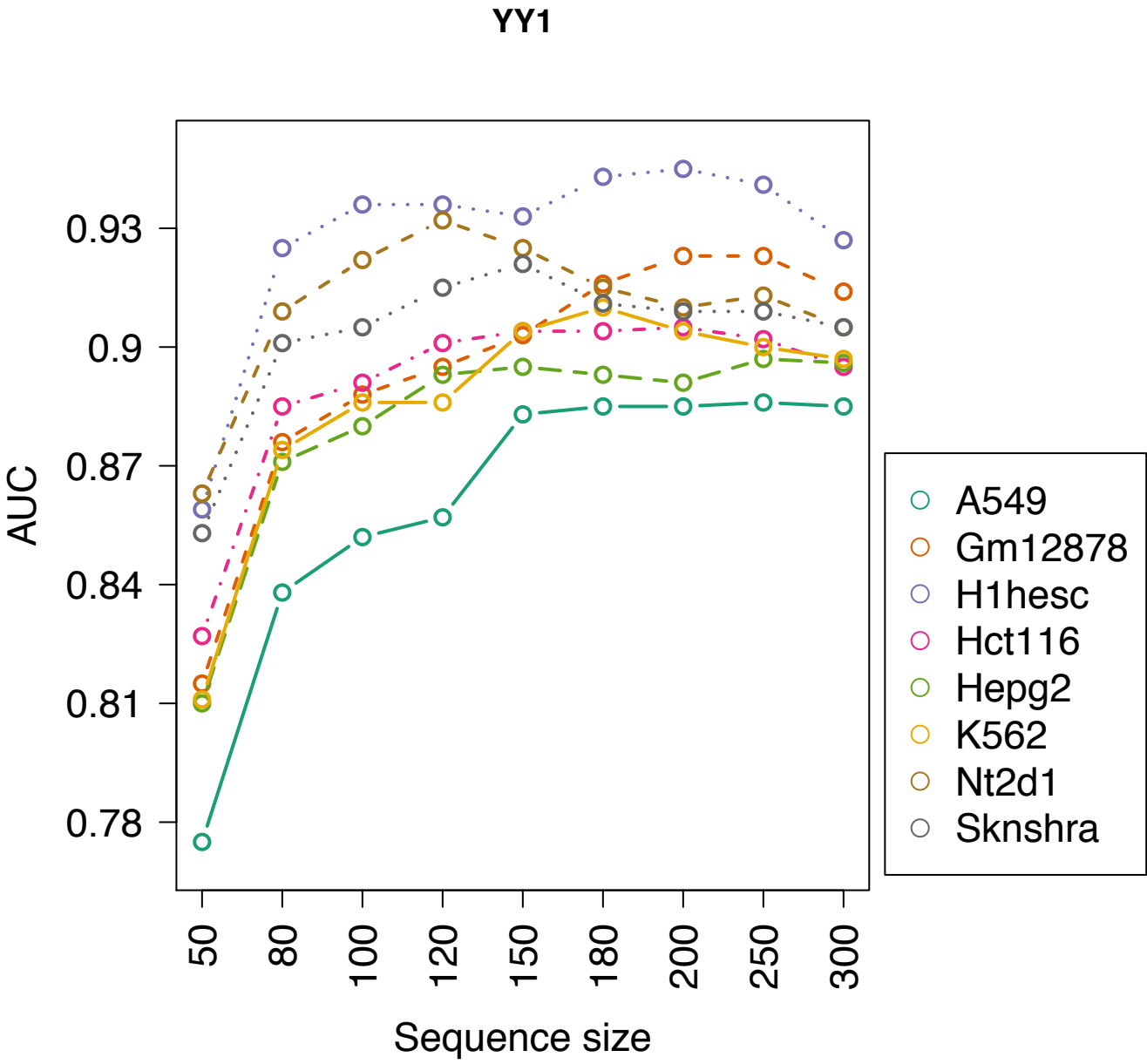

SRF
AUC
0.93
0.9
0.87
0.84
0.81
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
H1hesc
Hepg2
K562


TBP
AUC
0.85
0.82
0.79
0.76
0.73
0.7
50
80
100
120
150
180
200
250
300
Sequence size
Gm12878
H1hesc
Helas3
Hepg2
K562


TCF12
AUC
0.92
0.89
0.86
0.83
0.8
0.77
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Hepg2


TCF7L2
AUC
0.9
0.87
0.84
0.81
0.78
0.75
0.72
50
80
100
120
150
180
200
250
300
Sequence size
Hct116
Hek293
Helas3
Hepg2
Mcf7
Panc1


USF1
0.99
0.96
0.93
0.9
AUC
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Hepg2
K562


YY1
AUC
0.93
0.9
0.87
0.84
0.81
0.78
50
80
100
120
150
180
200
250
300
Sequence size
A549
Gm12878
H1hesc
Hct116
Hepg2
K562
Nt2d1
Sknshra