

Primer PICKR: Literature-Mined Scoring Platform for Robust RT-qPCR Primers

Thomas G. Molley^{1,3*}, Abhin

aba Banerjee^{1,3*}, Alis Balayan^{2,3*}, Jun K. Robbins¹, and Adam J. Engler^{1-3†}

¹Shu Chien-Gene Lay Department of Bioengineering and ²Biomedical Sciences Program,
University of California, San Diego; La Jolla, CA 92093 USA

³Sanford Consortium for Regenerative Medicine; La Jolla, CA 92037 USA

*These Authors contributed equally to this work.

†Corresponding Author: aengler@ucsd.edu

Phone: 858-246-0678

Fax: 858-534-5722

Supplementary Text

The full math of the PICKR score can be seen here:

$$\text{PICKR score: } = 100 (W_E e + W_B b + W_S s)$$

$$(W_E, W_B, W_S) = (0.50, 0.30, 0.20)$$

$$\text{Evidence score: } e_{\text{raw}} = s_{\text{shared}} + s_{\text{reg}} + s_{\text{tot}}$$

$$e = \min(e_{\text{raw}}, 1)$$

$$s_{\text{shared}} = 0.5 \min\left(\frac{\log_{10}(\text{shared}+1)}{\log_{10}(P_{\text{max}}+1)}, 1\right),$$

$$s_{\text{reg}} = 0.5 \frac{1}{2} \left[\min\left(\frac{\log_{10}(f_{\text{reg}}+1)}{\log_{10}(R_{\text{max}}+1)}, 1\right) + \min\left(\frac{\log_{10}(r_{\text{reg}}+1)}{\log_{10}(R_{\text{max}}+1)}, 1\right) \right],$$

$$s_{\text{tot}} = 0.5 \frac{1}{2} \left[\min\left(\frac{\log_{10}(f_{\text{tot}}+1)}{\log_{10}(T_{\text{max}}+1)}, 1\right) + \min\left(\frac{\log_{10}(r_{\text{tot}}+1)}{\log_{10}(T_{\text{max}}+1)}, 1\right) \right]$$

$$f_{\text{tot}} = f_{\text{reg}} + f_{\text{inv}}, \quad r_{\text{tot}} = r_{\text{reg}} + r_{\text{inv}}$$

$$\text{Biophysics score: } b = 0.5 (Q_f + Q_r), \quad Q_j = \frac{f_L(t_j) + f_G(GC_j) + f_T(T_{m,j})}{3};$$

For each primer side $p \in \{f, r\}$

$$X_p = \prod_{a \in A_p} \sigma(d_a, n_a^d), \quad \sigma(d, n^d) = \min\left(1, \beta_d + (1 - \beta_d) \rho \frac{n^d}{d}\right).$$

$$\beta_1 = 0.10, \quad \beta_2 = 0.35, \quad \beta_3 = 0.65, \quad \beta_4 = 0.90, \quad \rho = 0.60$$

$$\text{Synergy score: } s = \max\left(0, 1 - |T_m^f - T_m^r|/3 - \max(0, PC - D_c D_c)\right)$$

$$\text{Parameters: } P_{\text{max}} = 25, \quad R_{\text{max}} = 100, \quad T_{\text{max}} = 500$$

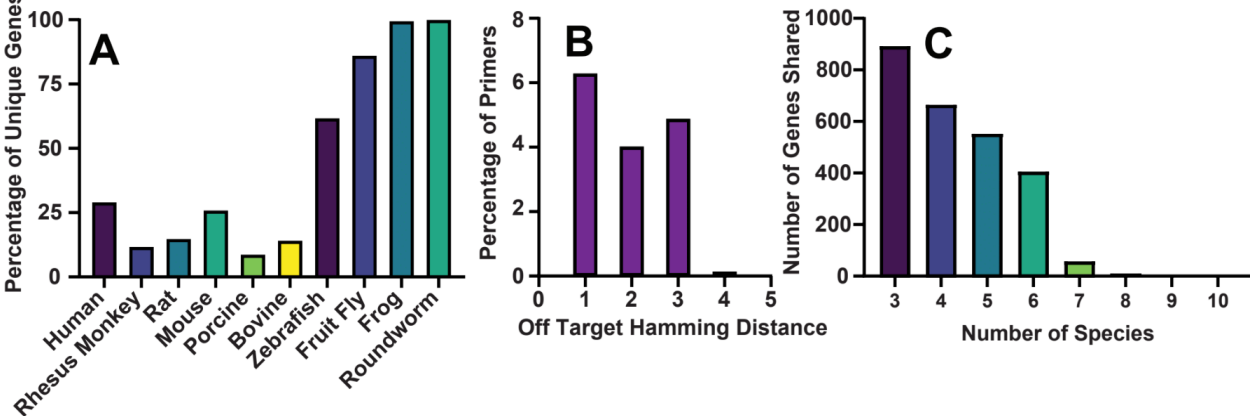
Literature citation
the strength of prior
experimental use

Amplicon length,
GC content,
Melt temperature

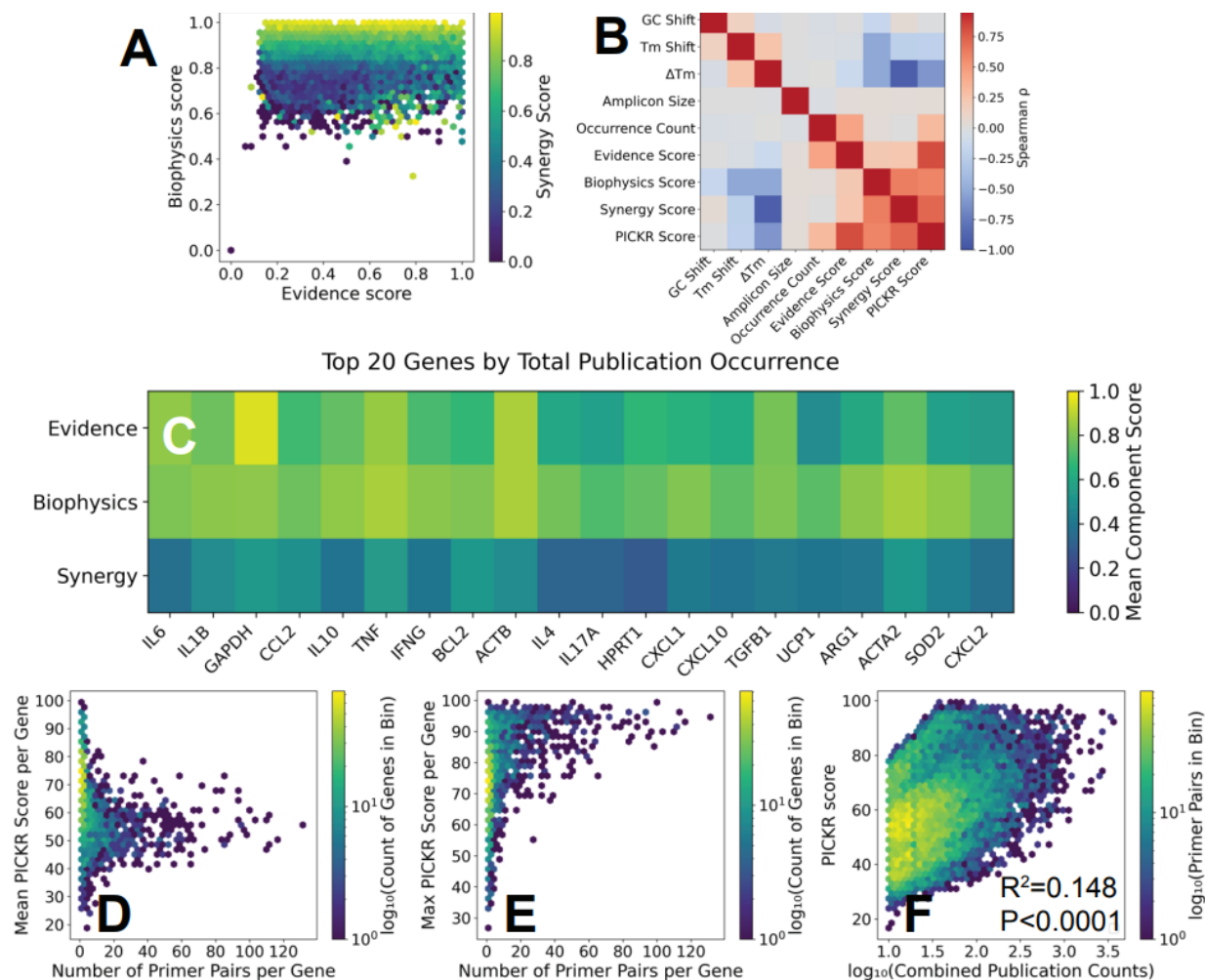
Differences in melt temp,
Complementarity

Here, e is the evidence score, b is the biophysics score, and s is the synergy score. W_E , W_B , and W_S are the weights for those three scores respectively. s_{shared} is the evidence component score for the number of pmcids that are shared between a forward and reverse primer. s_{reg} is the individual pmcids associated with each orientation at that exact direction, and s_{tot} is s_{reg} but also includes the matched reverse complements for a primer, i.e. a forward sequence for one person was inverted and used as a reverse for another person. Q_f and Q_r are the component biophysics score for the forward and reverse primers respectively. f_L is the amplicon length function, f_G is for GC content, and f_T is for melt temperature. X_p calculated the penalty for cross specificity matches, summing up the products of mismatches, with the penalties for mismatches at x hamming distances being equal to B_x . For the synergy score, PC is the pair complementarity score, and $D_c D_c$ is set to 6. T_m is melt temperature.

Supplementary Figures

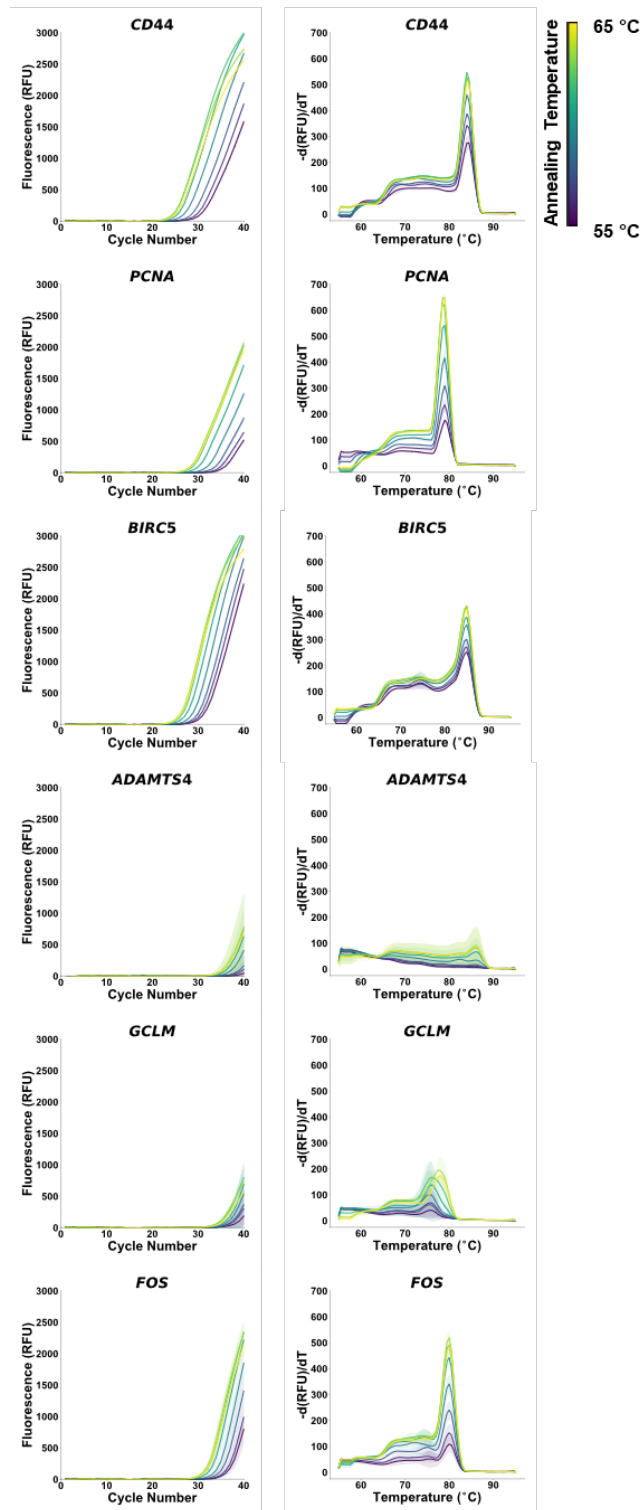


Supplementary Figure 1: Primer uniqueness. (A) Percentage of primers for each species that are unique only to that species. (B) Percentage of primers with off-target hits at different Hamming distances. (C) number of genes that are shared across multiple species. All graphs generated with the 290,071 unique primers used at least 3 times in the literature from the 399,394 papers scraped. Source data are provided as a Source Data file. The code used in the creation of these data are published under a CC-BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>).



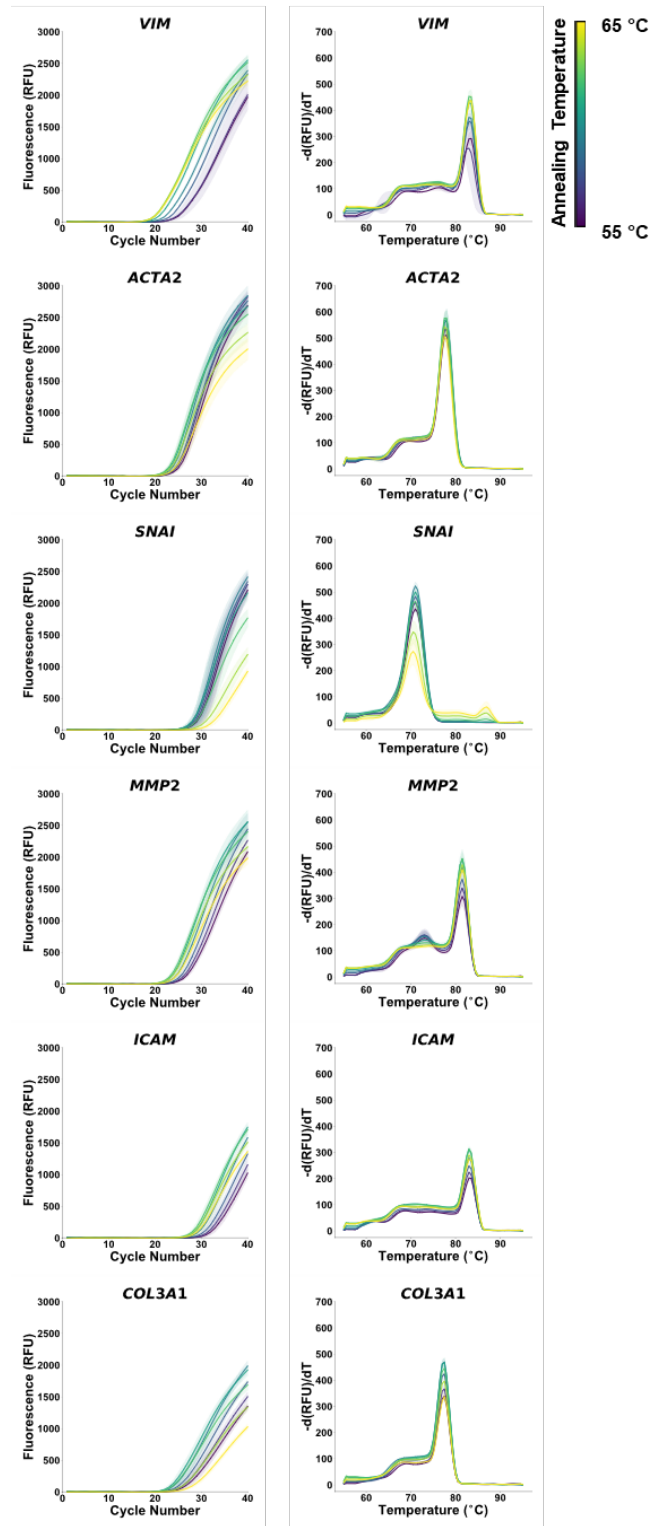
Supplementary Figure 2: Primer score characteristics for Mouse. (A) Hexbin plot of the three component scores (evidence, biophysics, and synergy). (B) Pearson correlation plot of PICKR scores, the component scores, and the meta data used to calculate the scores. (C) Component scores of the average primer pair for the twenty most used genes. Hexbin plots of mean (D) and maximum PICKR (E) score for all human genes. (F) Hexbin plot of PICKR score against log transformed publication counts for primer pairs. Sample sizes for all graphs were the 16,529 generated primer pairs derived from the 290,071 primers pulled from the literature. Source data are provided as a Source Data file. The code used in the creation of these data are published under a CC-BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>).

Set 1 from 90-100



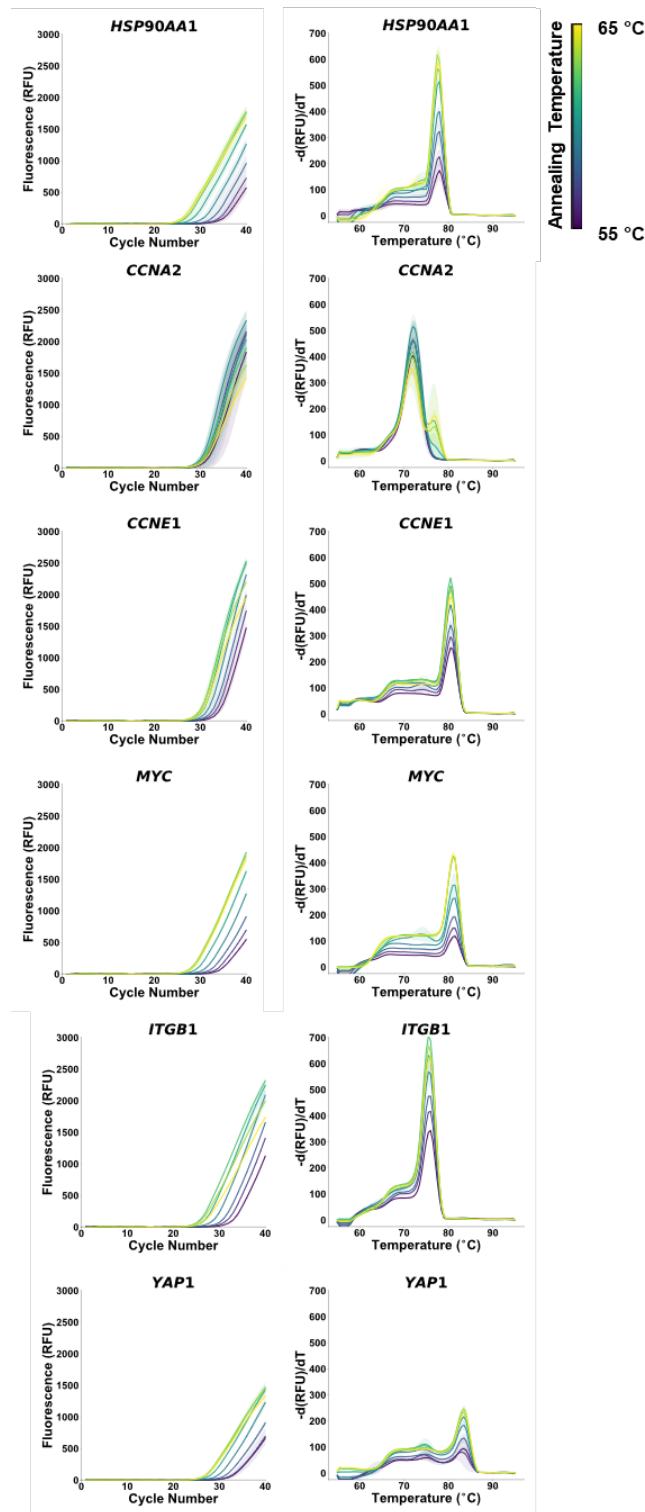
Supplementary Figure 3. Amplification curves and melt curves for 90-100 bin set 1. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 2 from 90-100



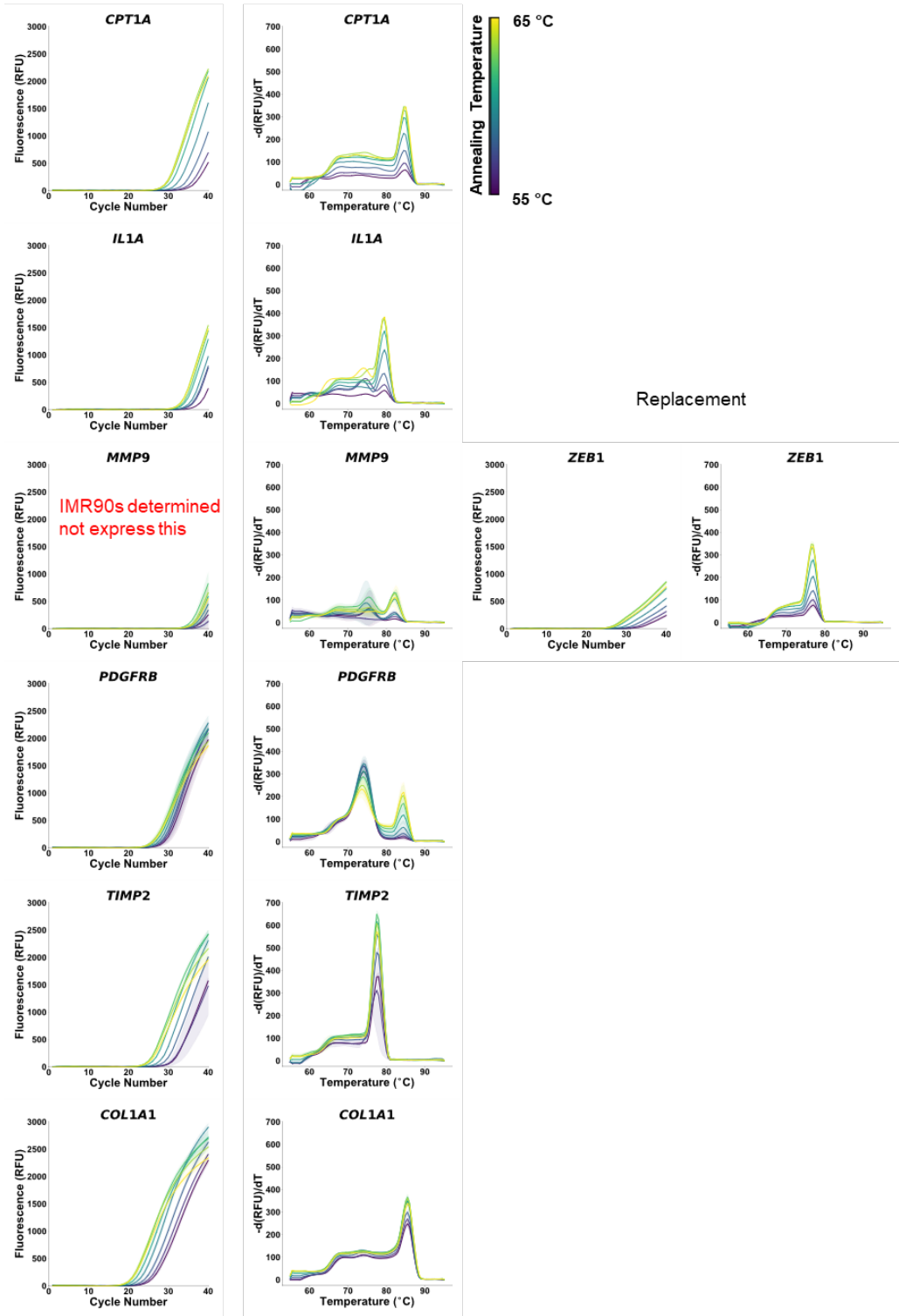
Supplementary Figure 4. Amplification curves and melt curves for 90-100 bin set 2. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 3 from 90-100

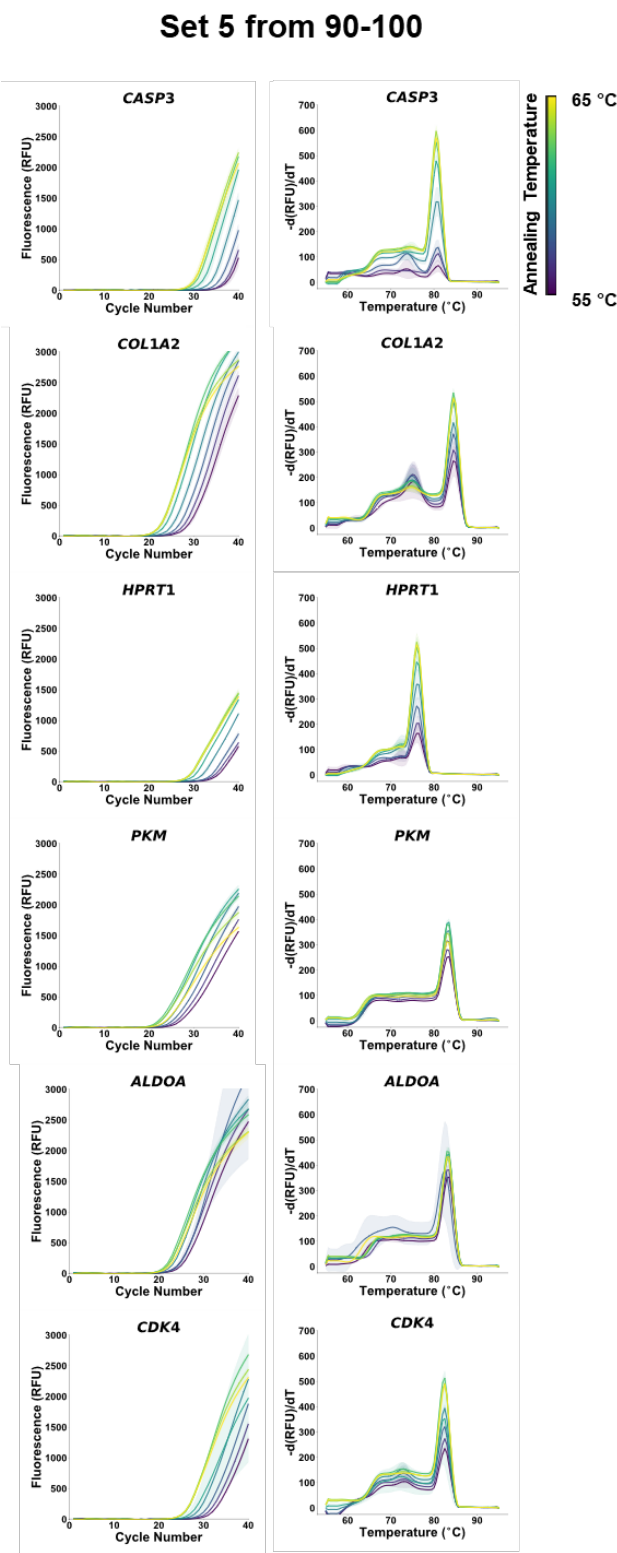


Supplementary Figure 5. Amplification curves and melt curves for 90-100 bin set 3. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 4 from 90-100



Supplementary Figure 6. Amplification curves and melt curves for 90-100 bin set 4. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



63

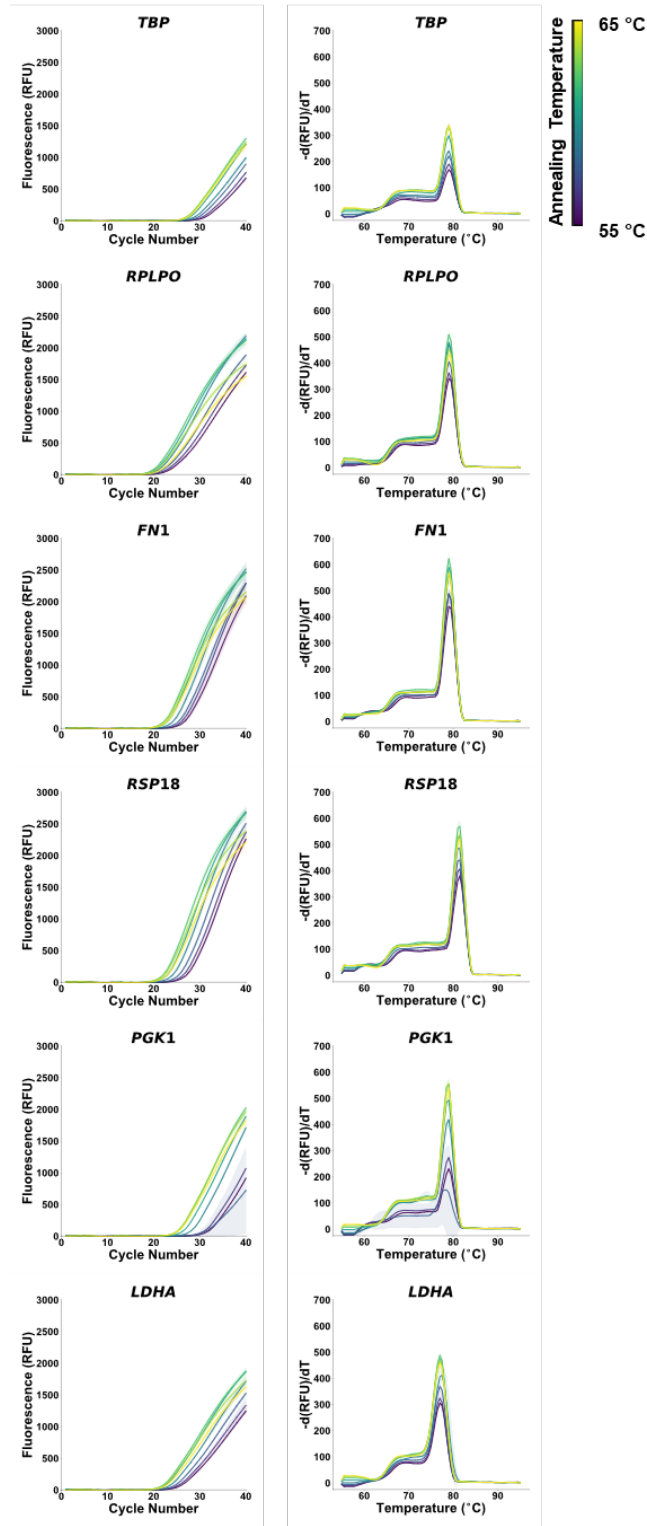
64

65

66

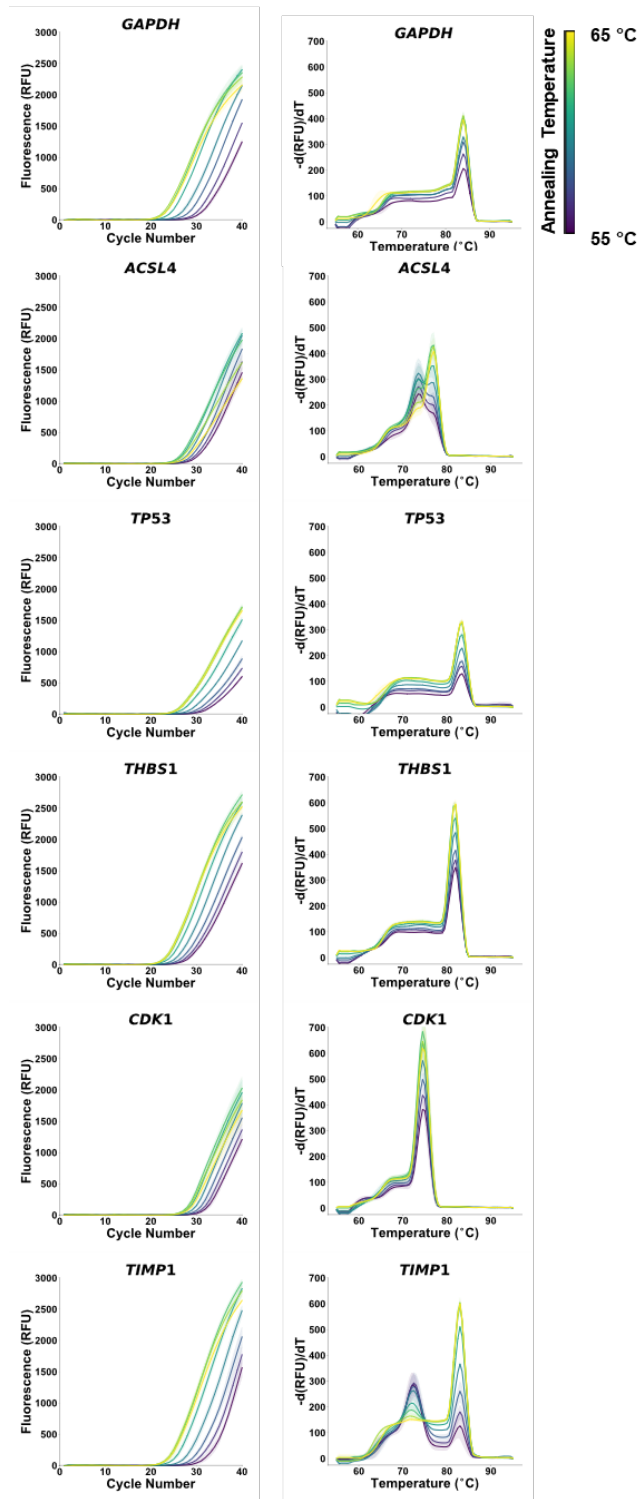
Supplementary Figure 7. Amplification curves and melt curves for 90-100 bin set 5. 8
different temperatures were chosen to create amplification curves (left) and melt temperature
curves (right). Source data are provided as a Source Data file.

Set 6 from 90-100



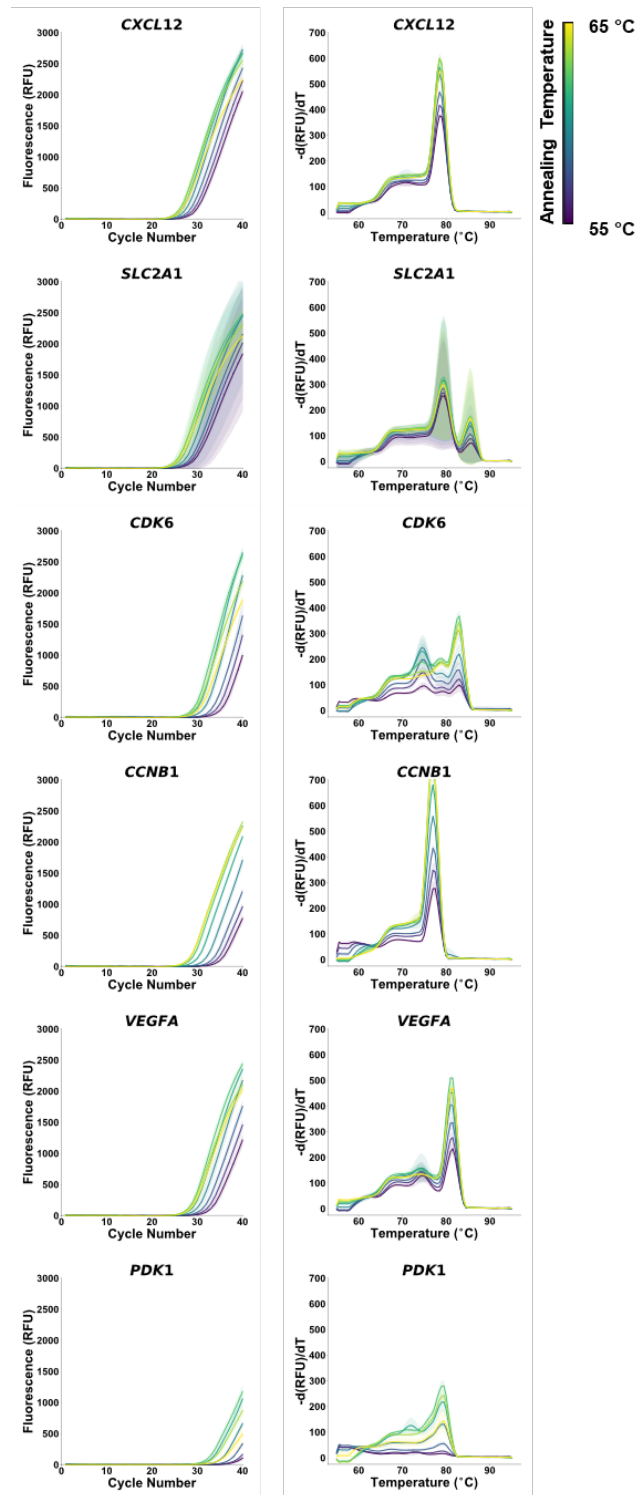
Supplementary Figure 8. Amplification curves and melt curves for 90-100 bin set 6. 8
different temperatures were chosen to create amplification curves (left) and melt temperature
curves (right). Source data are provided as a Source Data file.

Set 7 from 90-100



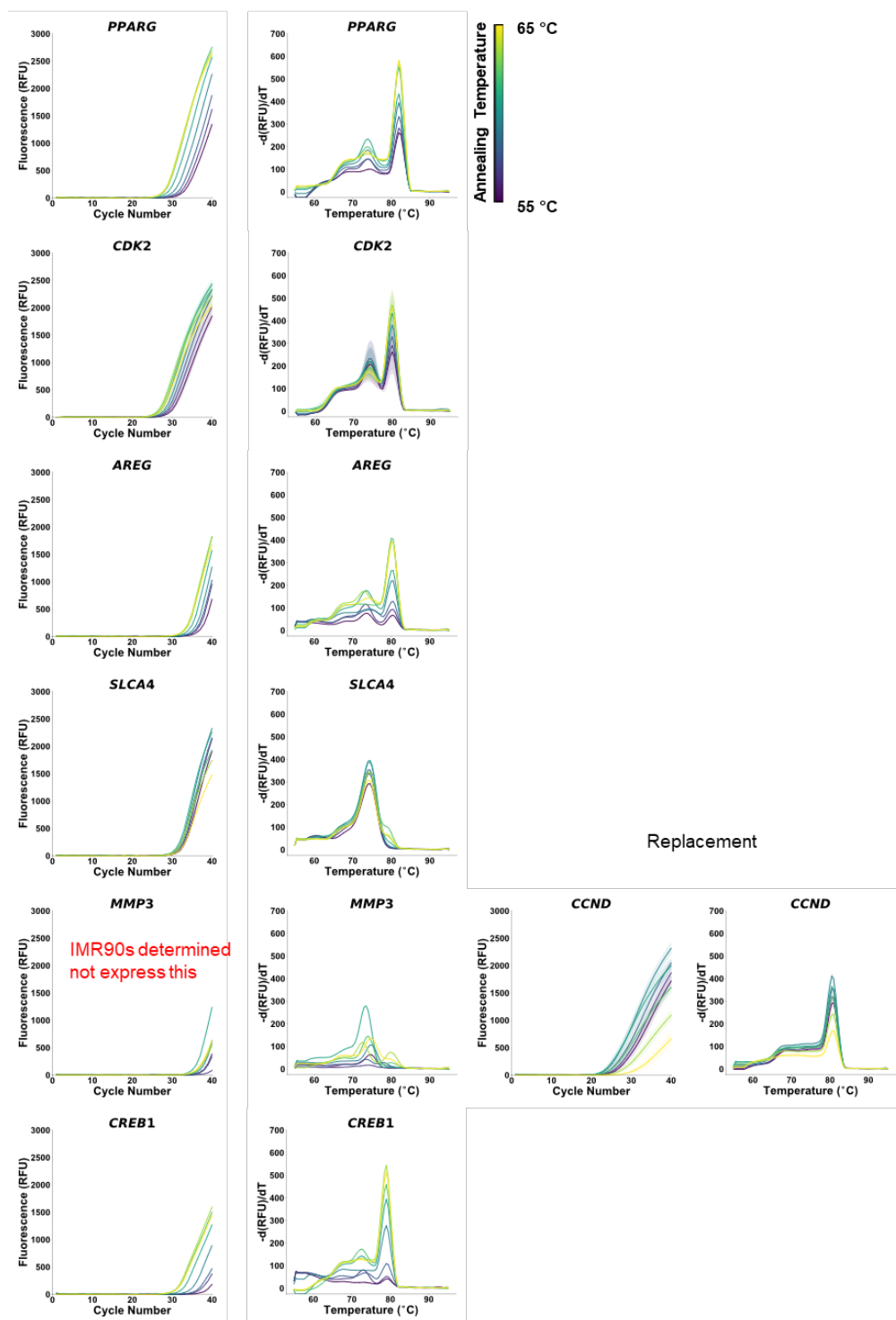
Supplementary Figure 9. Amplification curves and melt curves for 90-100 bin set 7.8
different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 8 from 90-100



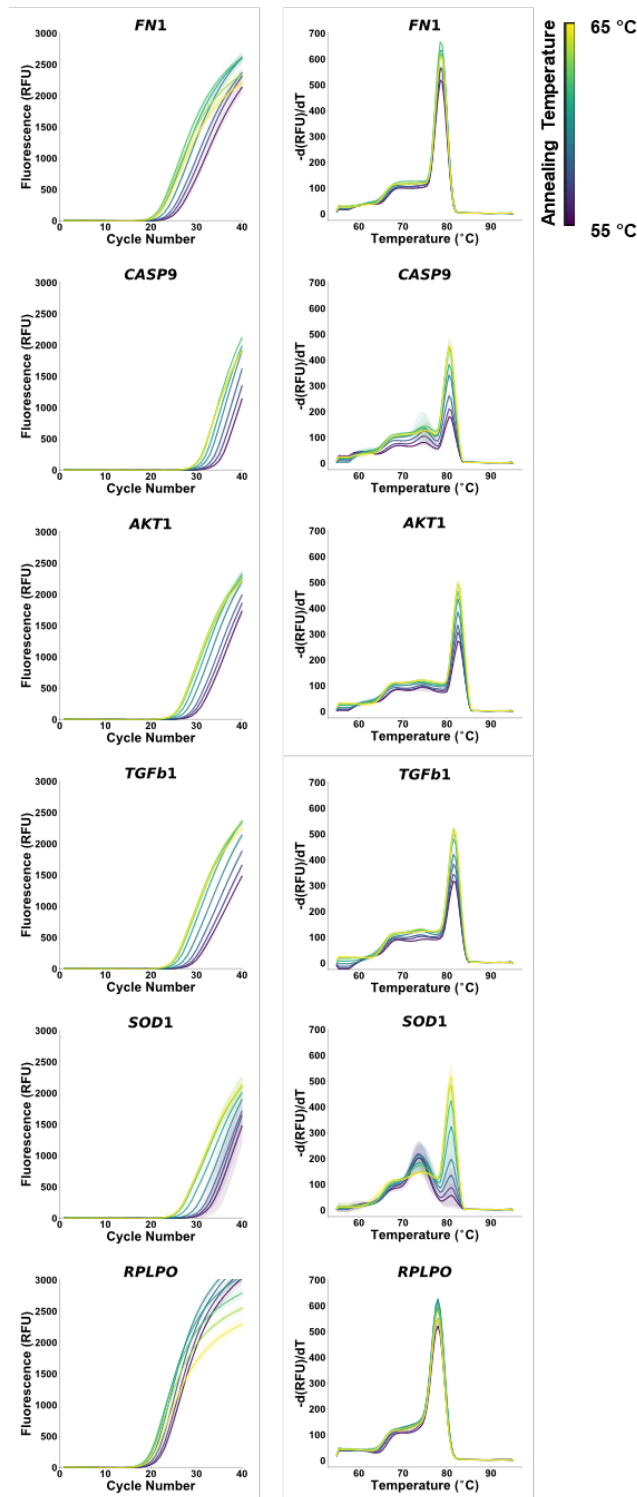
Supplementary Figure 10. Amplification curves and melt curves for 90-100 bin set 8. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 1 from 80-90



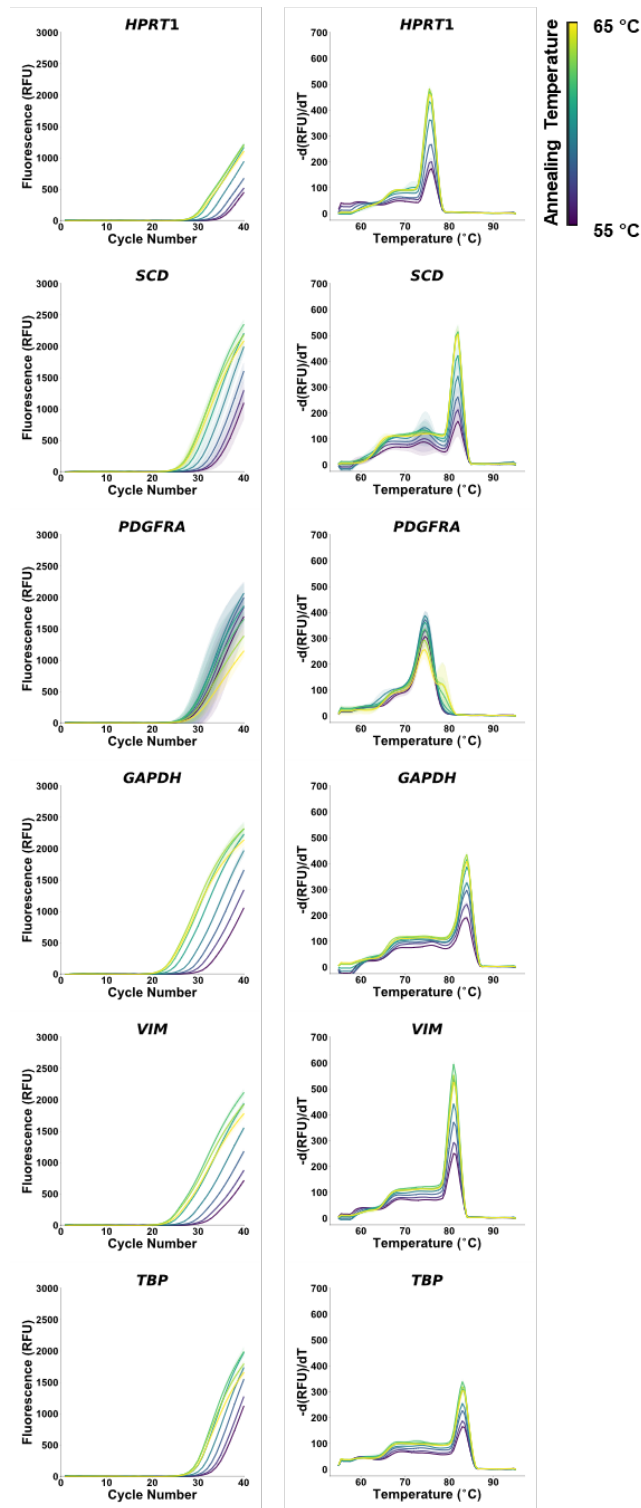
Supplementary Figure 11. Amplification curves and melt curves for 80-90 bin set 1. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 2 from 80-90



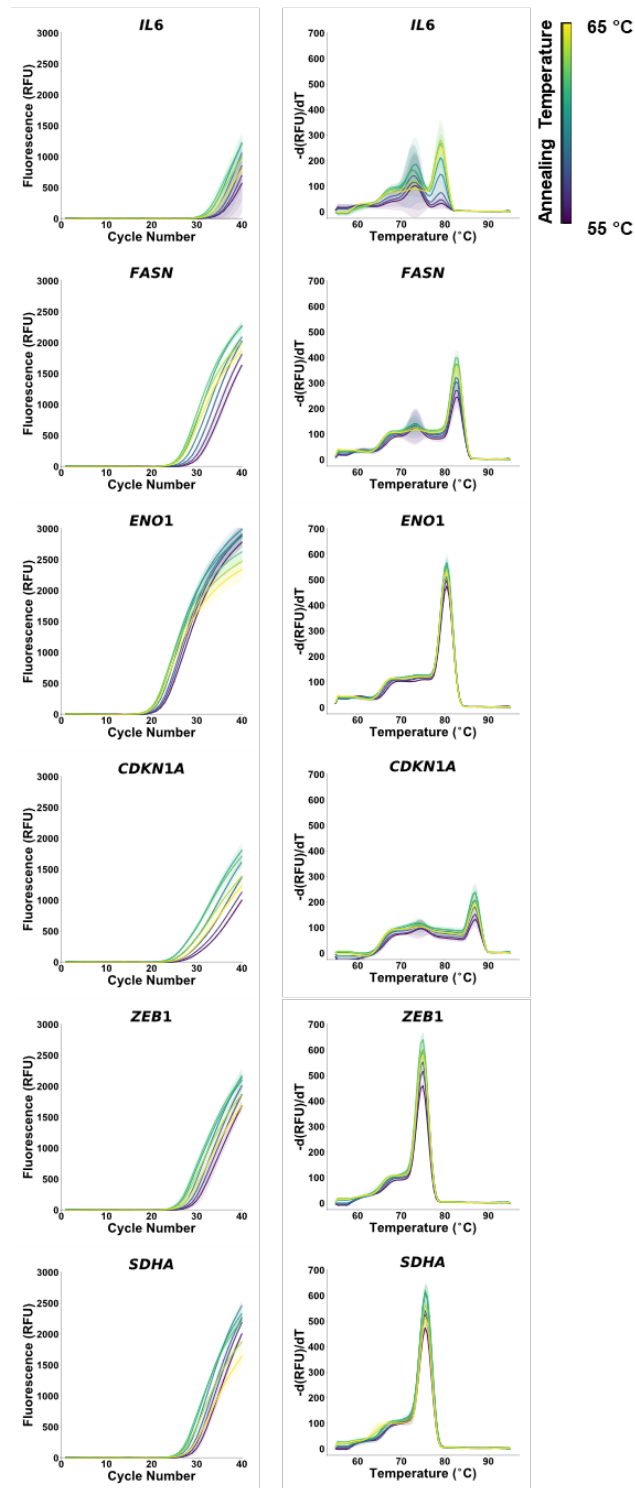
Supplementary Figure 12. Amplification curves and melt curves for 80-90 bin set 2. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 3 from 80-90



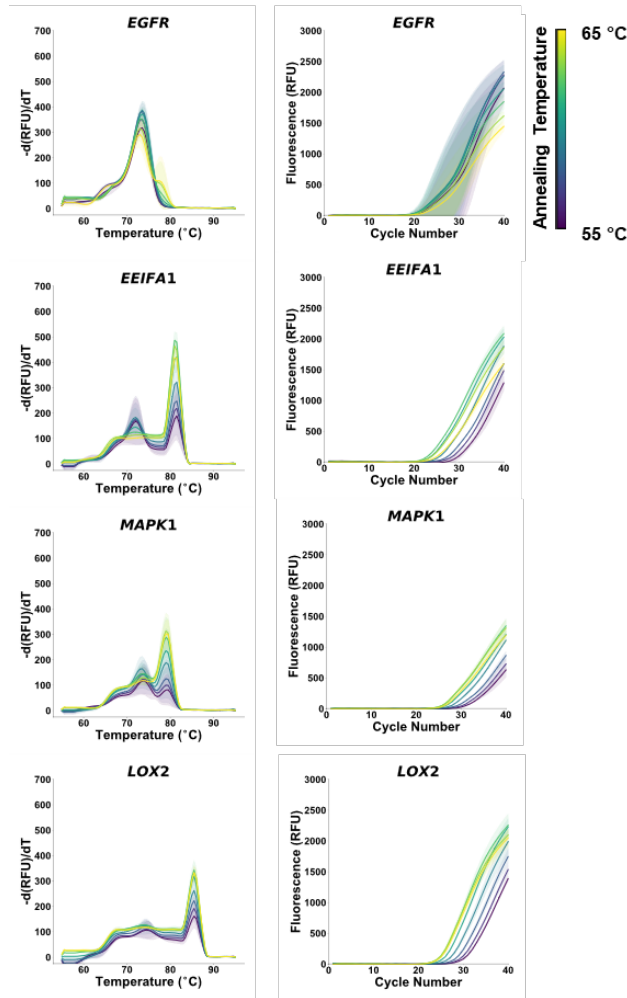
Supplementary Figure 13. Amplification curves and melt curves for 80-90 bin set 3. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 4 from 80-90



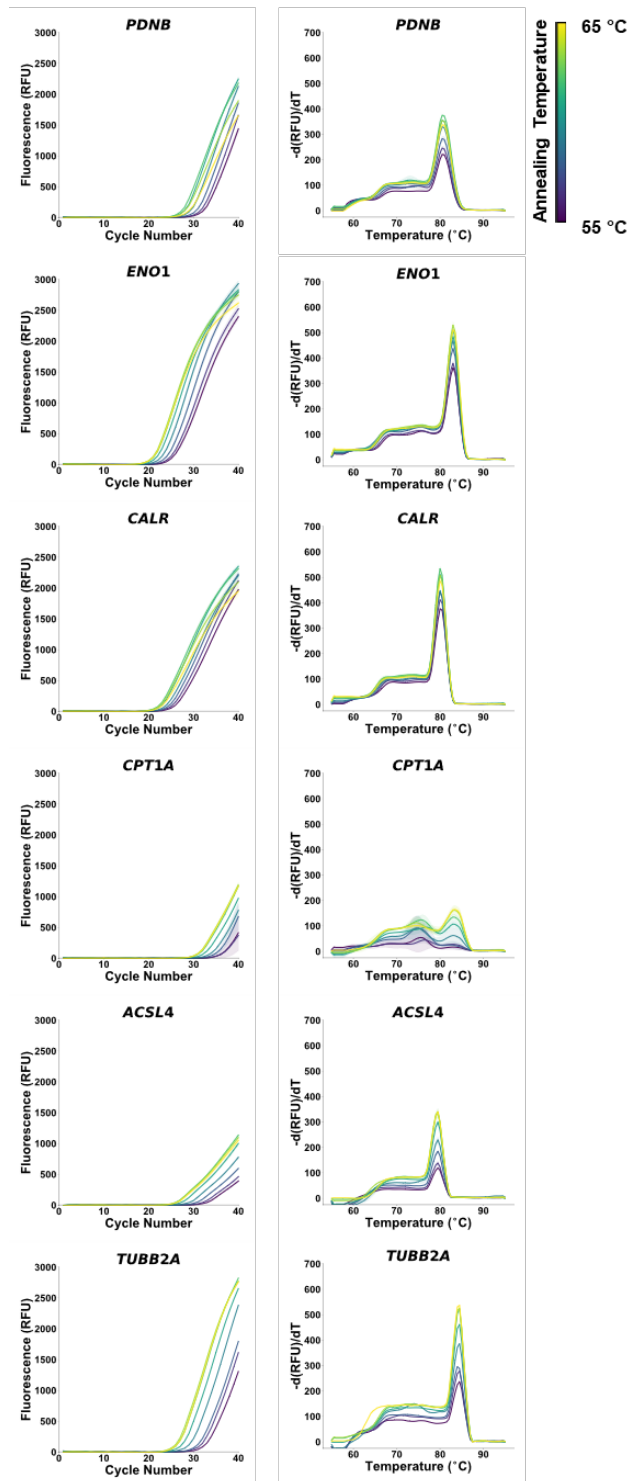
Supplementary Figure 14. Amplification curves and melt curves for 80-90 bin set 4. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 5 from 80-90



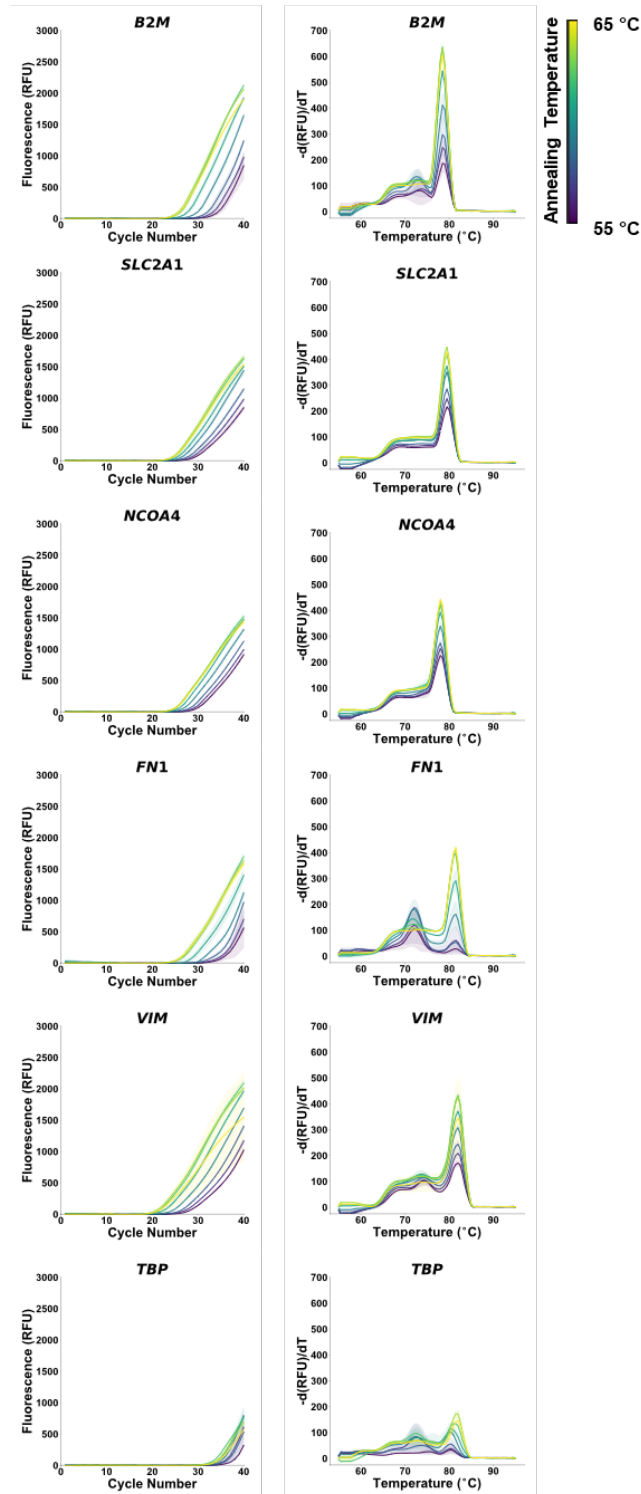
Supplementary Figure 15. Amplification curves and melt curves for 80-90 bin set 5. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 1 from 70-80



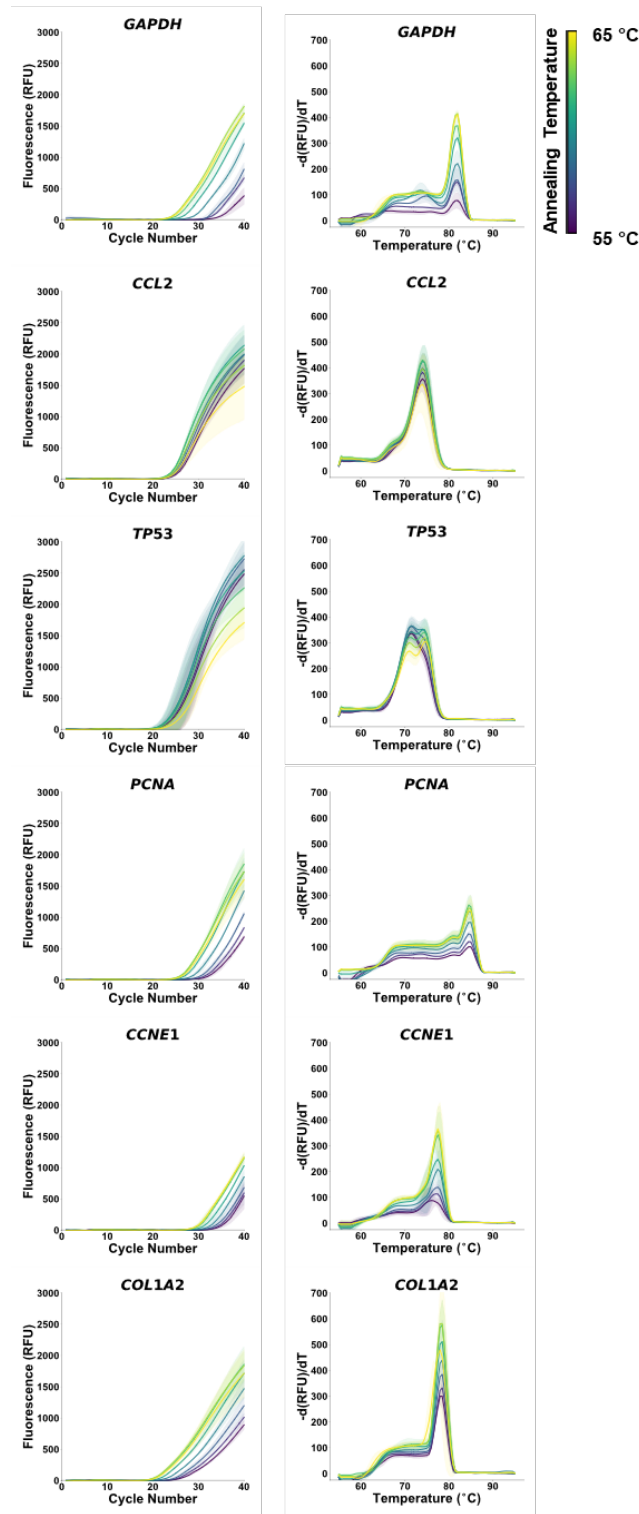
Supplementary Figure 16. Amplification curves and melt curves for 70-80 bin set 1. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 2 from 70-80



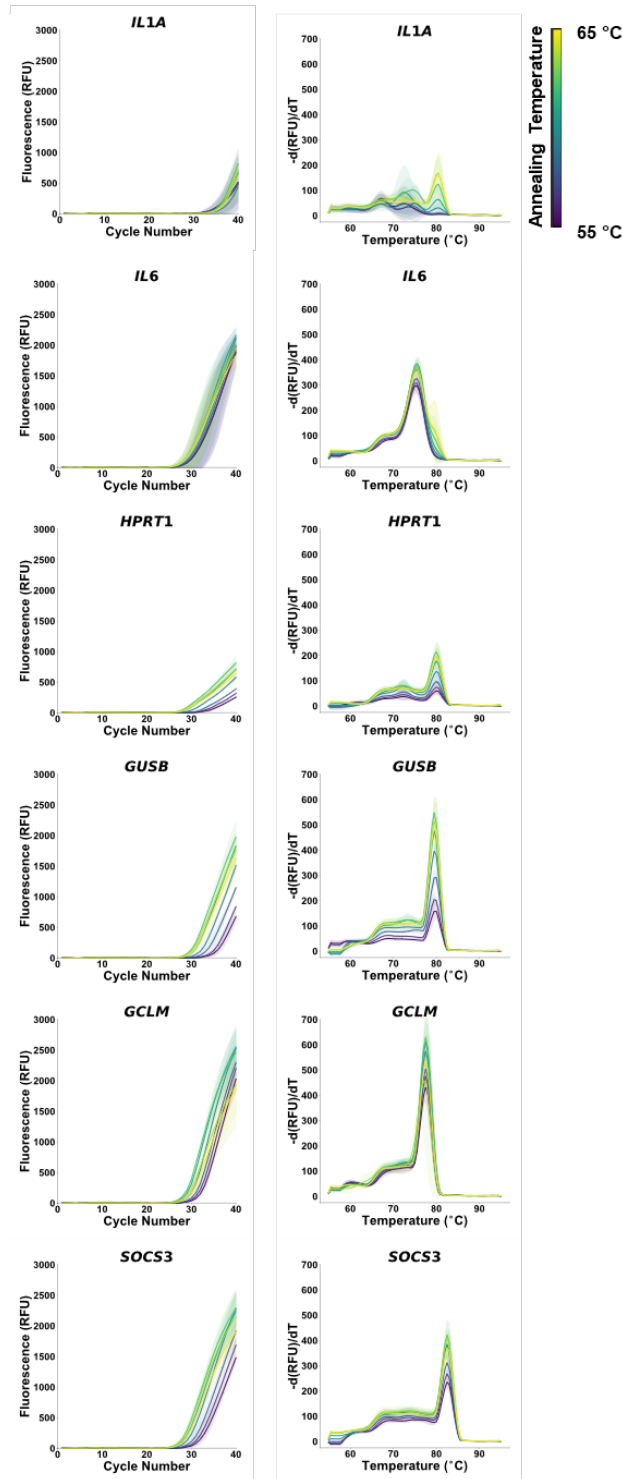
Supplementary Figure 17. Amplification curves and melt curves for 70-80 bin set 2. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 3 from 70-80



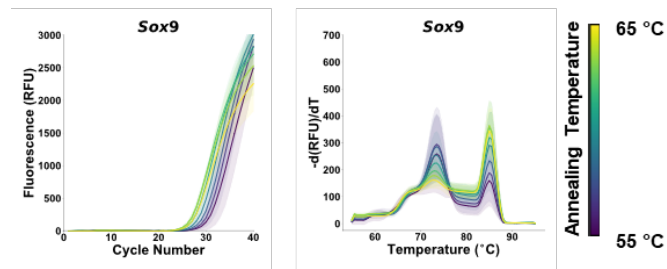
Supplementary Figure 18. Amplification curves and melt curves for 70-80 bin set 3. 8
different temperatures were chosen to create amplification curves (left) and melt temperature
curves (right). Source data are provided as a Source Data file.

Set 4 from 70-80



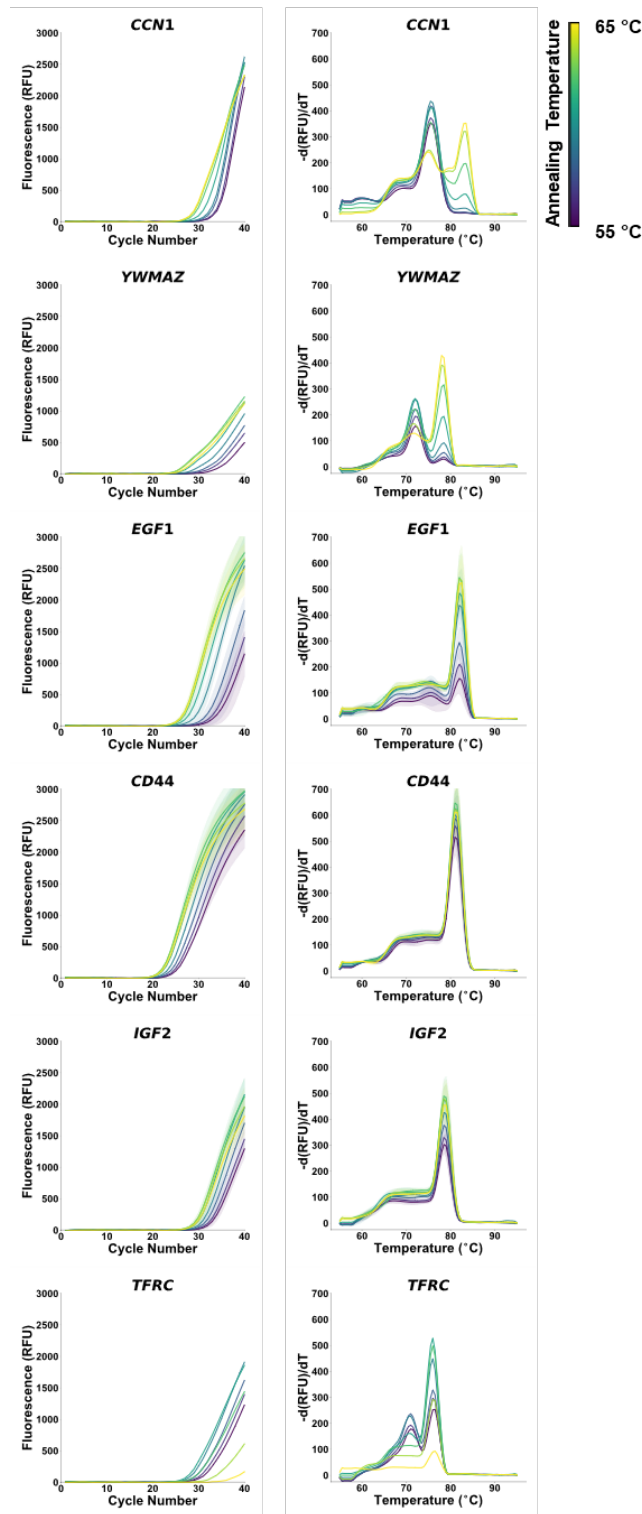
Supplementary Figure 19. Amplification curves and melt curves for 70-80 bin set 4. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 5 from 70-80



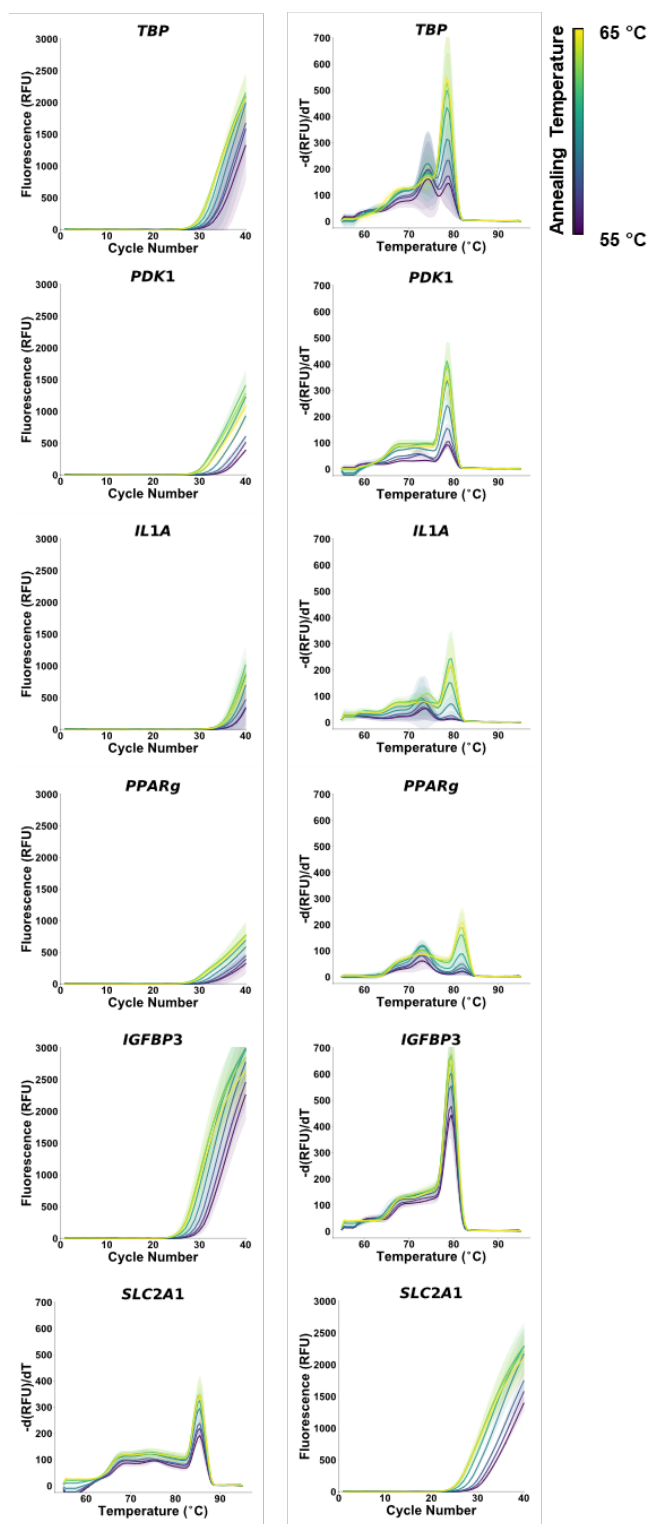
Supplementary Figure 20. Amplification curves and melt curves for 70-80 bin set 5. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 1 from 60-70



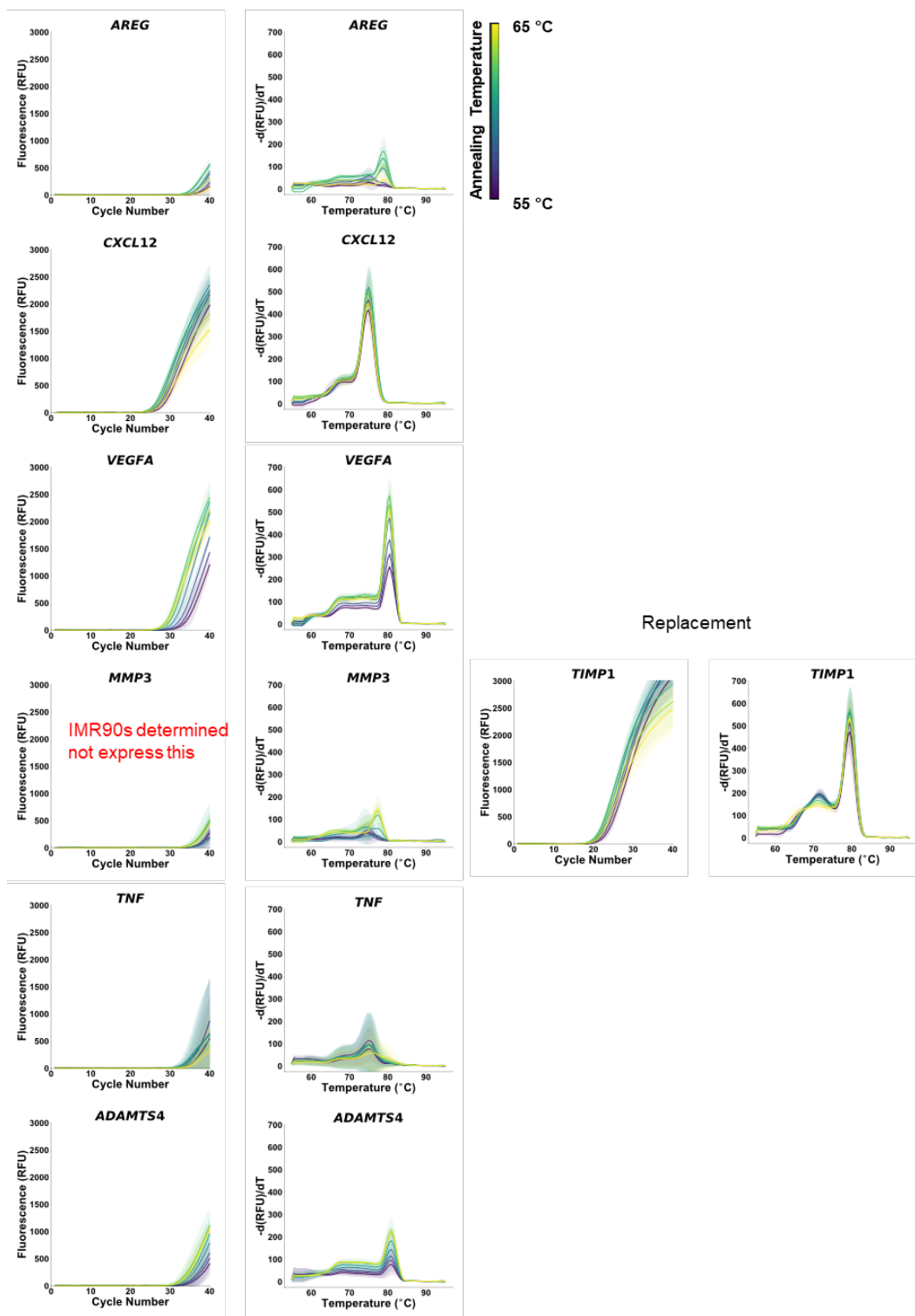
Supplementary Figure 21. Amplification curves and melt curves for 60-70 bin set 1. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 2 from 60-70



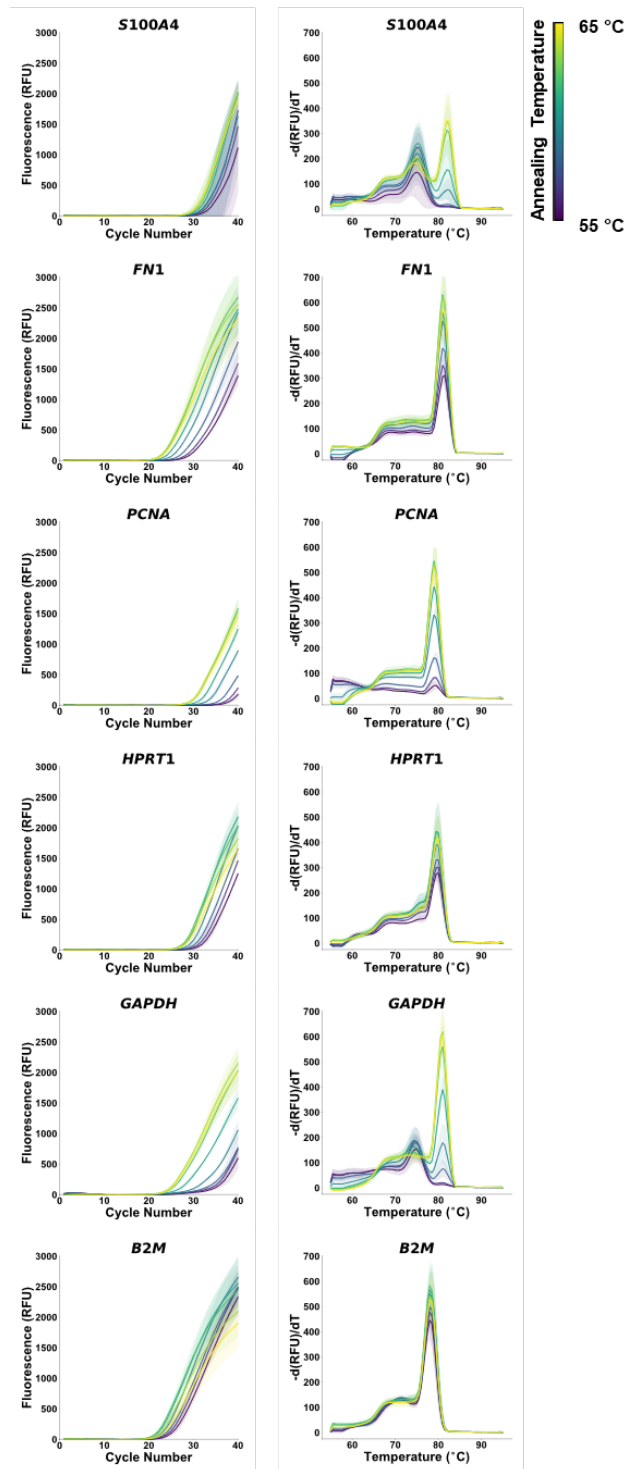
Supplementary Figure 22. Amplification curves and melt curves for 60-70 bin set 2. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 3 from 60-70



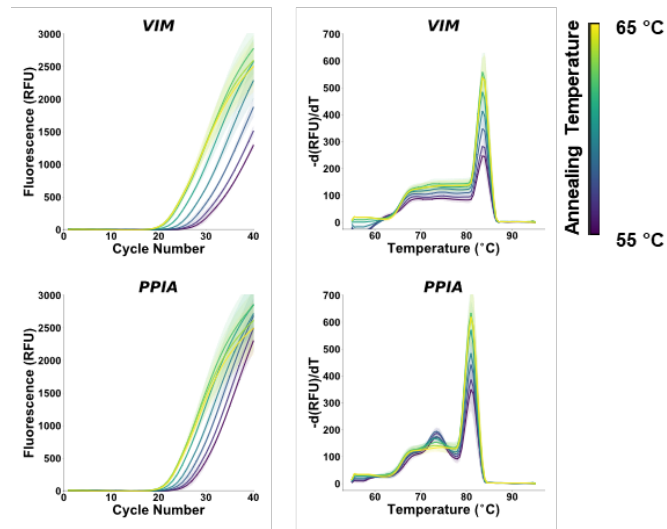
Supplementary Figure 23. Amplification curves and melt curves for 60-70 bin set 3. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 4 from 60-70



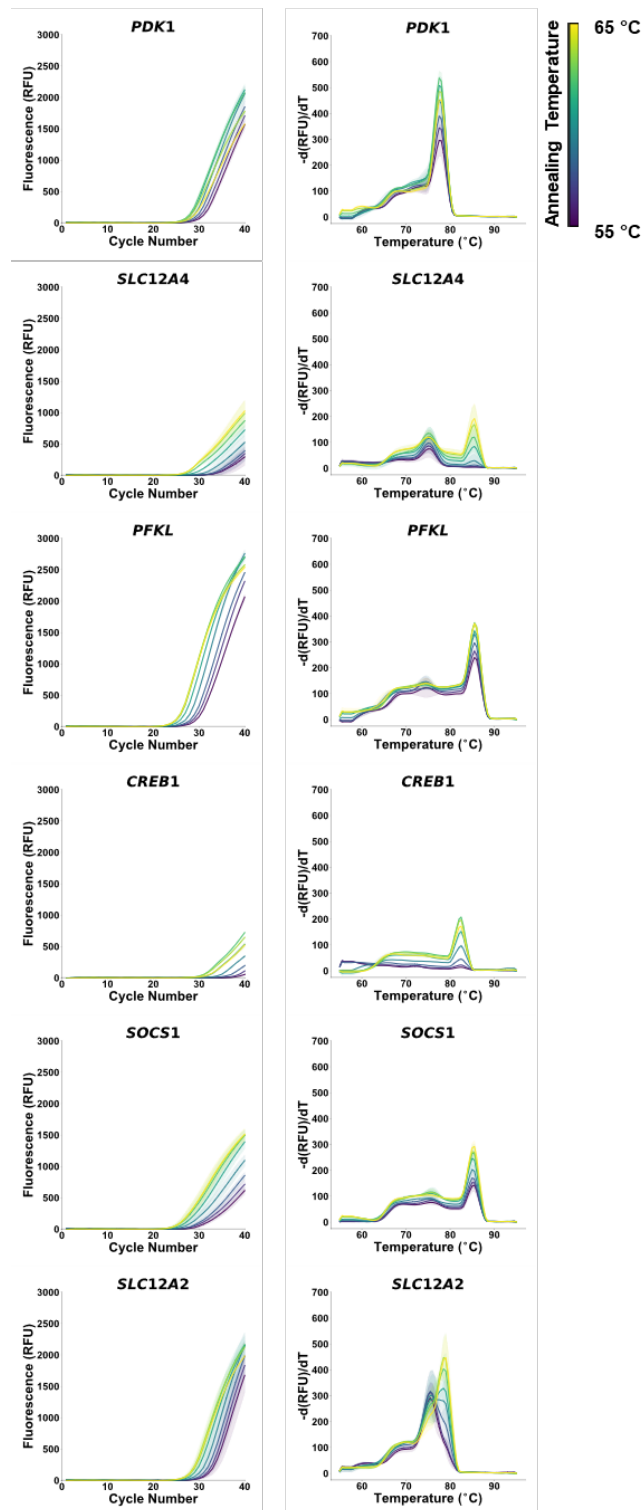
Supplementary Figure 24. Amplification curves and melt curves for 60-70 bin set 4. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 5 from 60-70



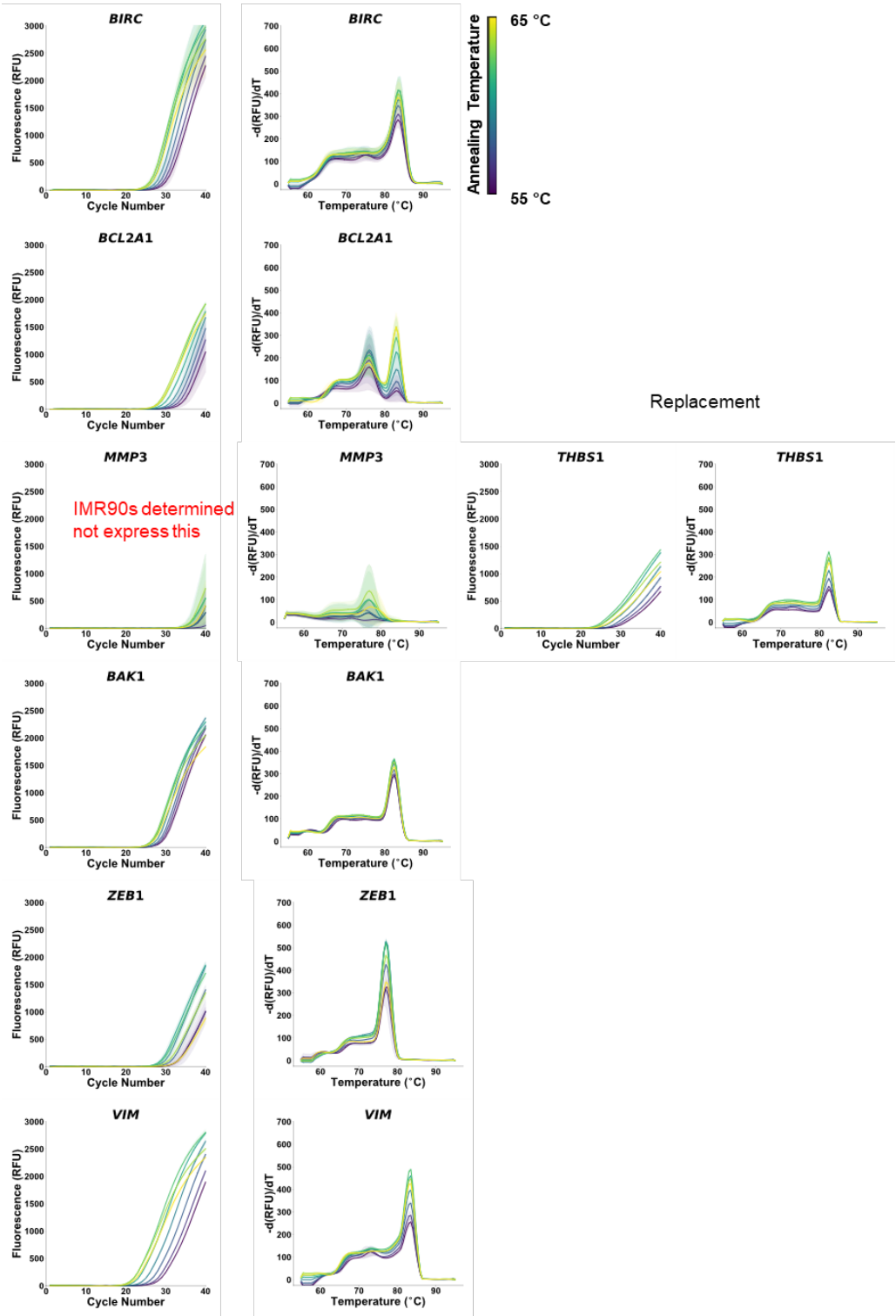
Supplementary Figure 25. Amplification curves and melt curves for 60-70 bin set 5. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

Set 1 from 50-60

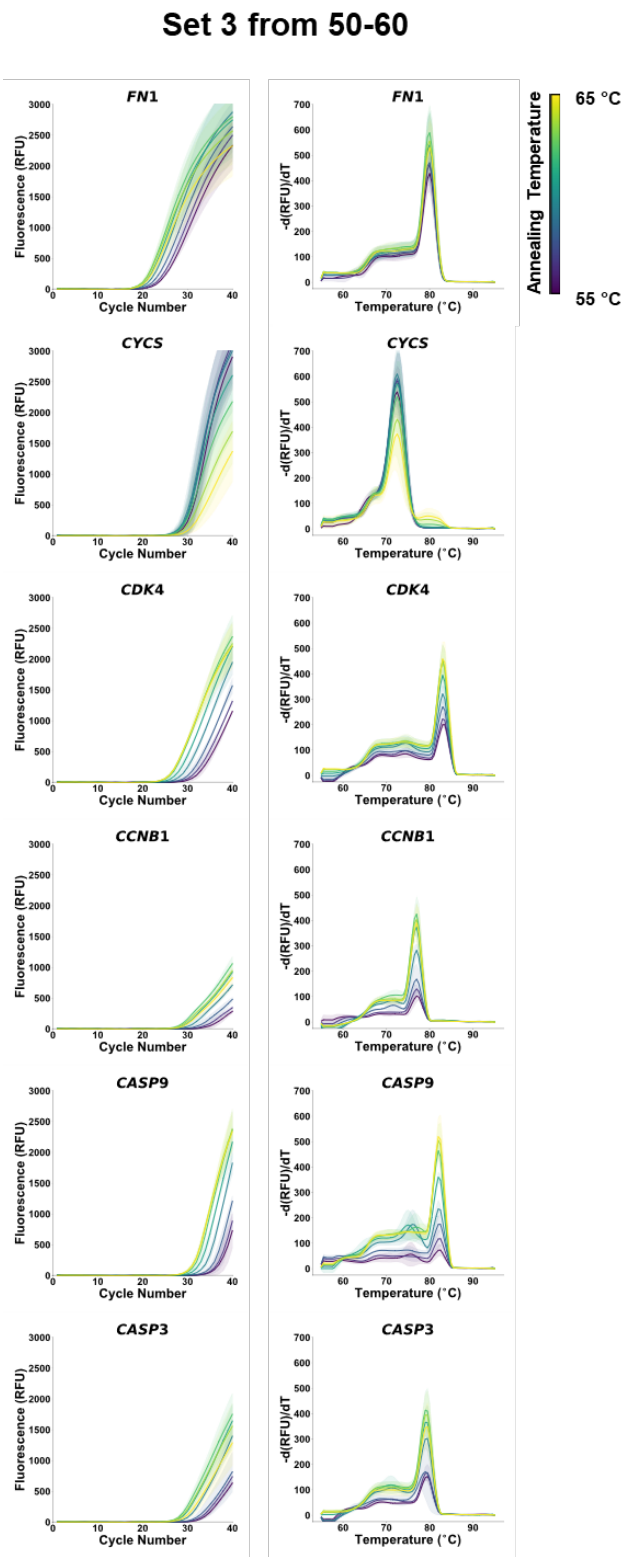


Supplementary Figure 26. Amplification curves and melt curves for 50-60 bin set 1. 8
different temperatures were chosen to create amplification curves (left) and melt temperature
curves (right). Source data are provided as a Source Data file.

Set 2 from 50-60



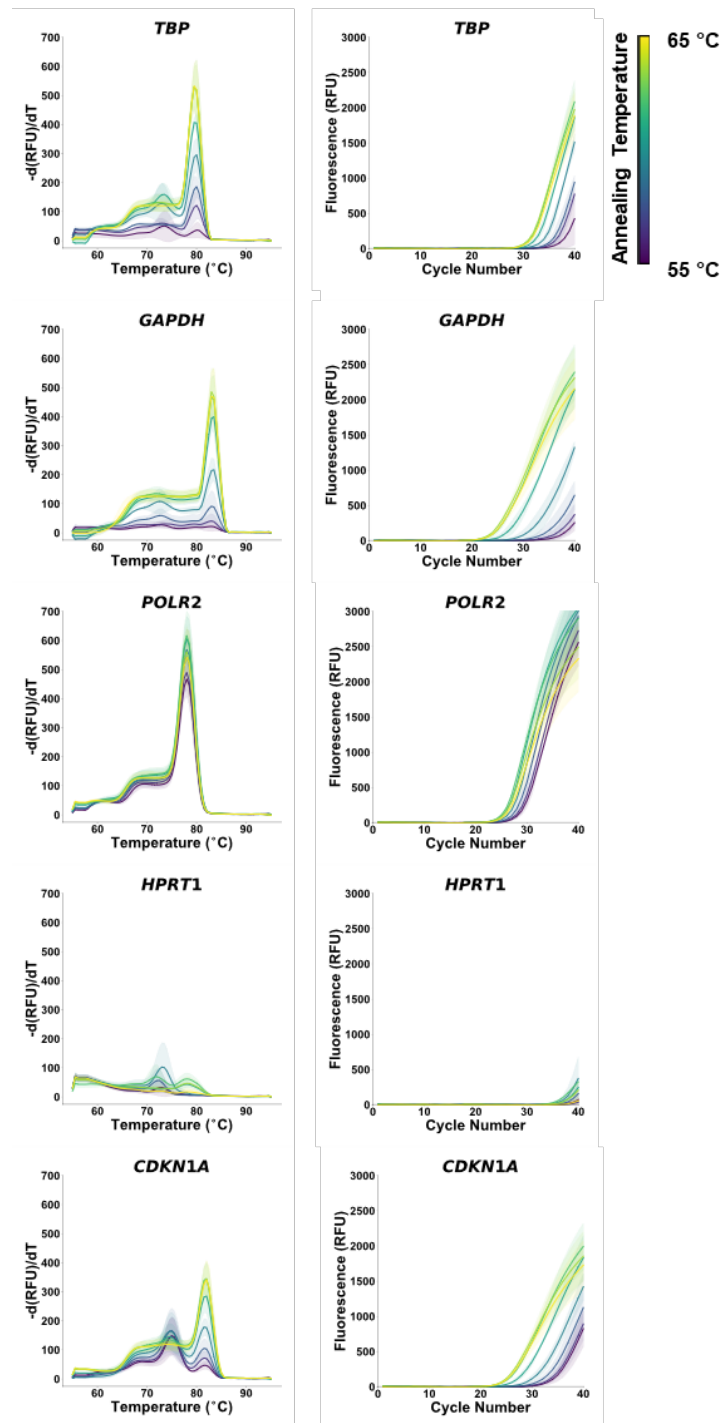
Supplementary Figure 27. Amplification curves and melt curves for 50-60 bin set 2. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



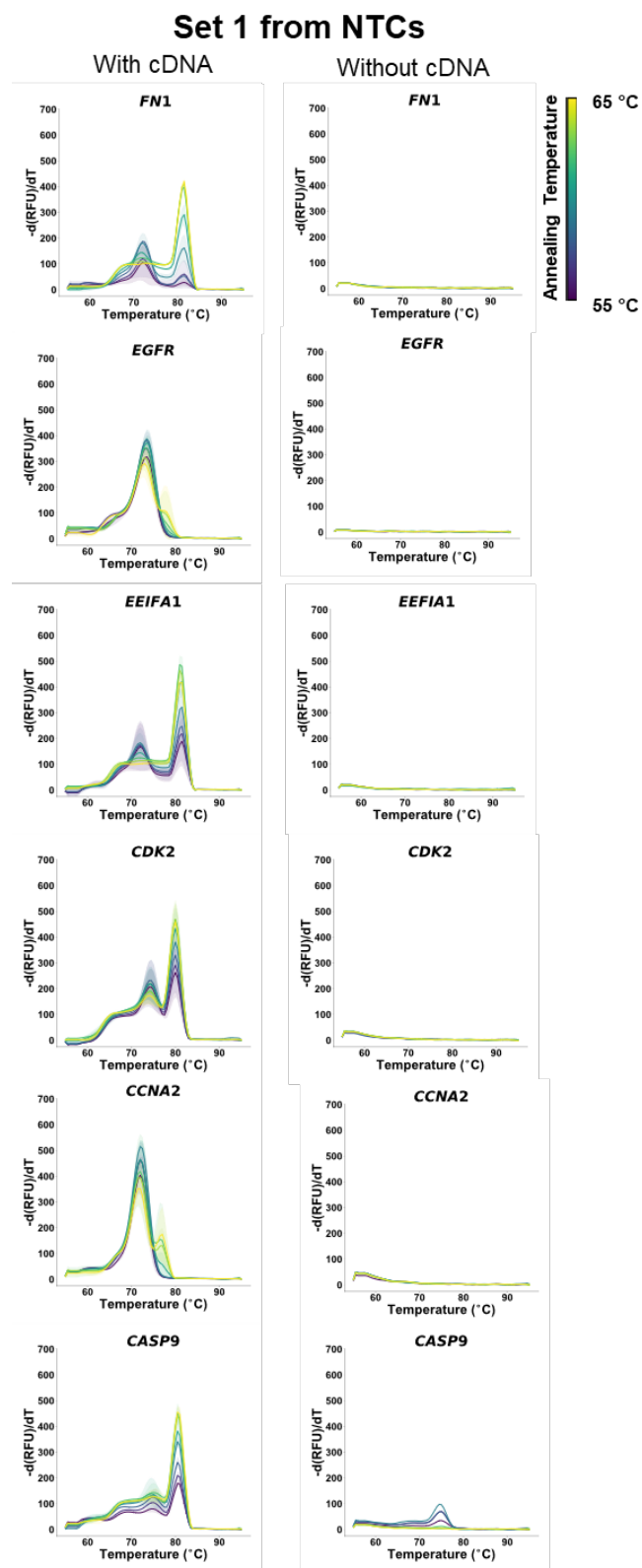
150
151
152
153

Supplementary Figure 28. Amplification curves and melt curves for 50-60 bin set 3. 8
different temperatures were chosen to create amplification curves (left) and melt temperature
curves (right). Source data are provided as a Source Data file.

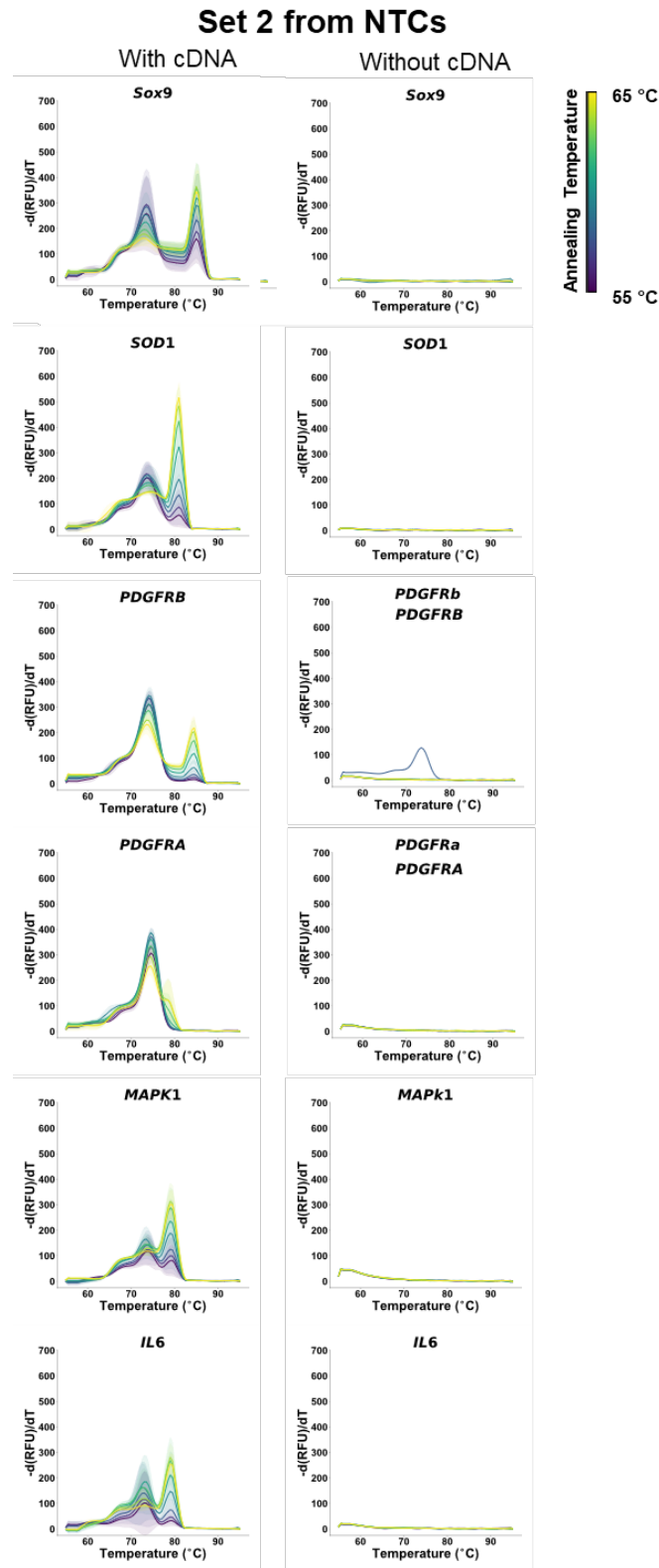
Set 4 from 50-60



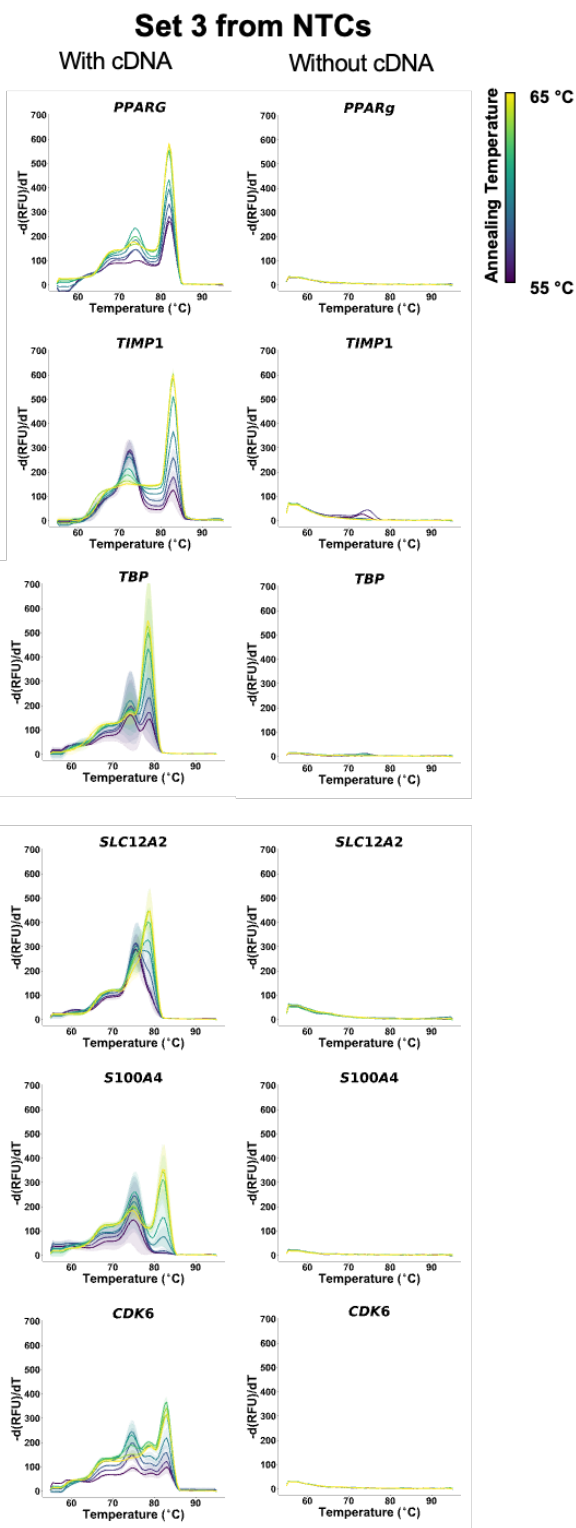
Supplementary Figure 29. Amplification curves and melt curves for 50-60 bin set 4. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



Supplementary Figure 30. Melt curves for cDNA and NTCs of double peak genes set 1. 8 different temperatures were chosen to create melt temperature curves with (left) and without cDNA template. Source data are provided as a Source Data file.



Supplementary Figure 31. Melt curves for cDNA and NTCs of double peak genes set 2. 8 different temperatures were chosen to create melt temperature curves with (left) and without cDNA template. Source data are provided as a Source Data file.

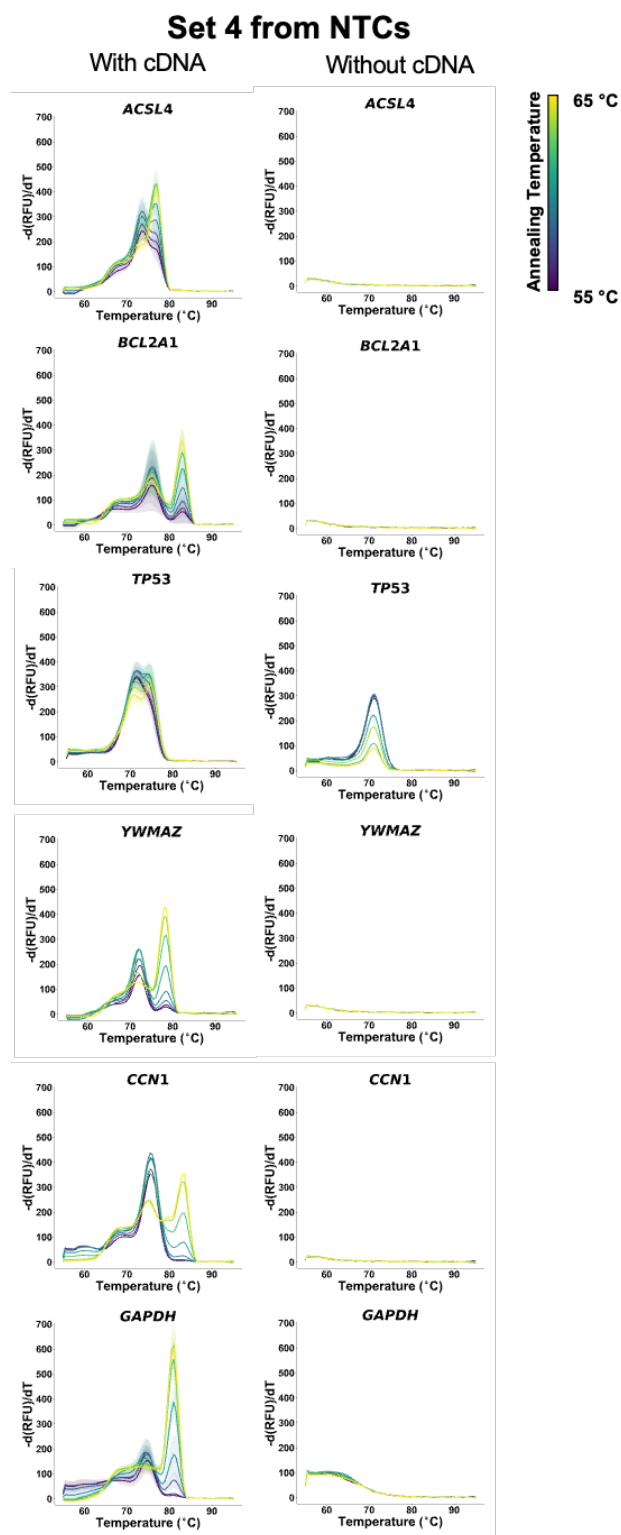


168

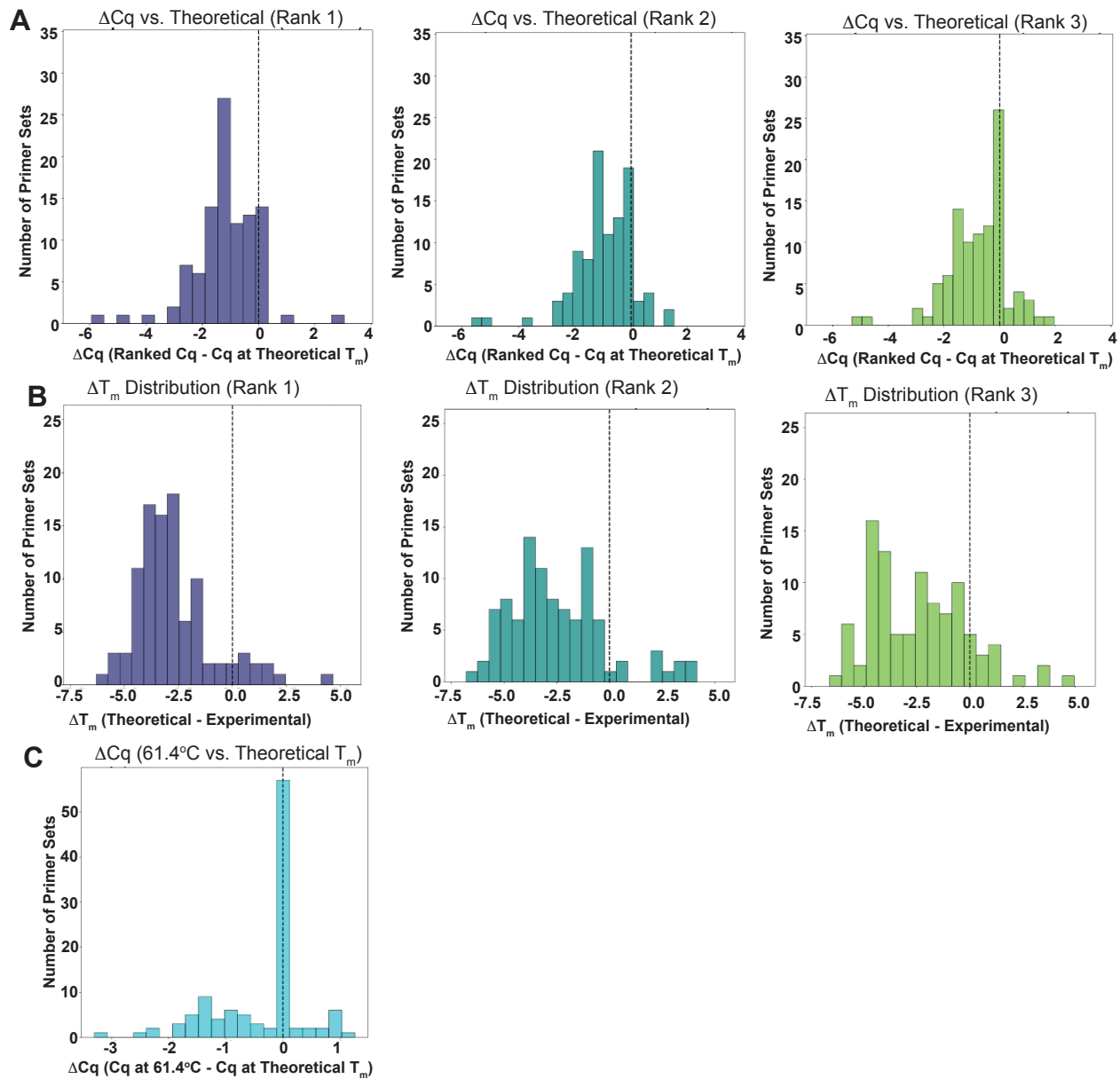
169

170

Supplementary Figure 32. Melt curves for cDNA and NTCs of double peak genes set 3. 8
different temperatures were chosen to create melt temperature curves with (left) and without
cDNA template. Source data are provided as a Source Data file.



Supplementary Figure 33. Melt curves for cDNA and NTCs of double peak genes set 4. 8 different temperatures were chosen to create melt temperature curves with (left) and without cDNA template. Source data are provided as a Source Data file.



176

177

178

179

180

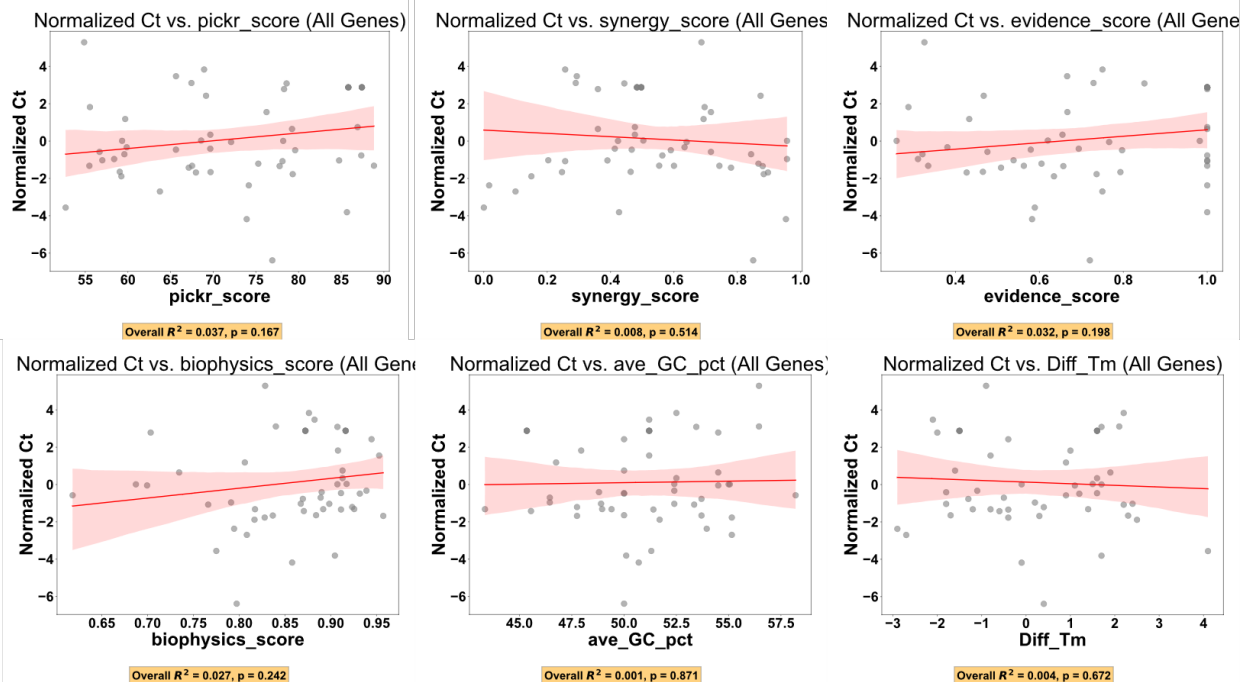
181

182

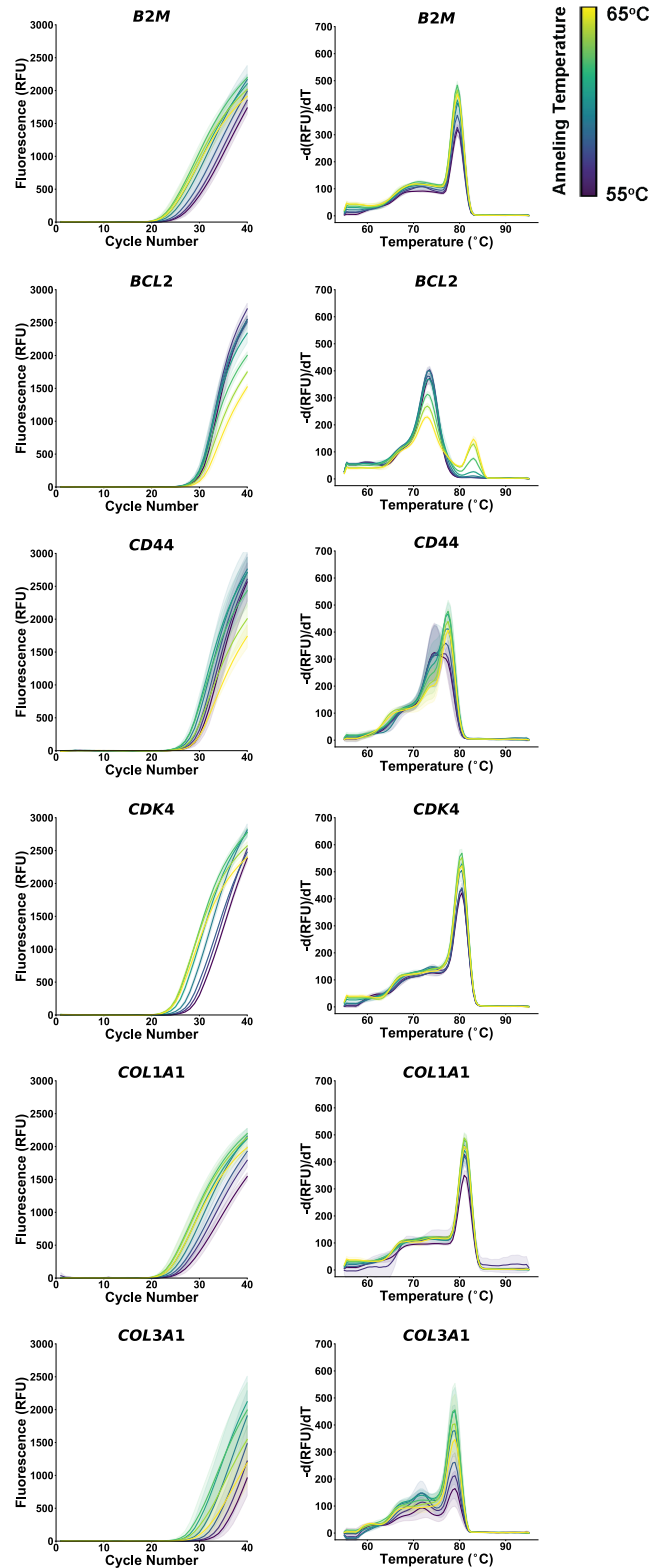
183

184

Supplementary Figure 34. Ranking of optimal melt temps. (A) the difference in CT value from the top three ranked melt temperatures as compared to the theoretical. (B) the difference in melt temperature from the top three ranked melt temperatures as compared to the theoretical. (C) Difference in theoretical melt temperatures for each primer pair as compared to 61.4 °C. Sample sizes are (A) n = 132, 132, and 132 for ranks 1, 2, and 3, respectively. Sample sizes are (B) n = 132, 132, and 132 for ranks 1, 2, and 3, respectively. Sample size in (C) n = 132. Source data are provided as a Source Data file.

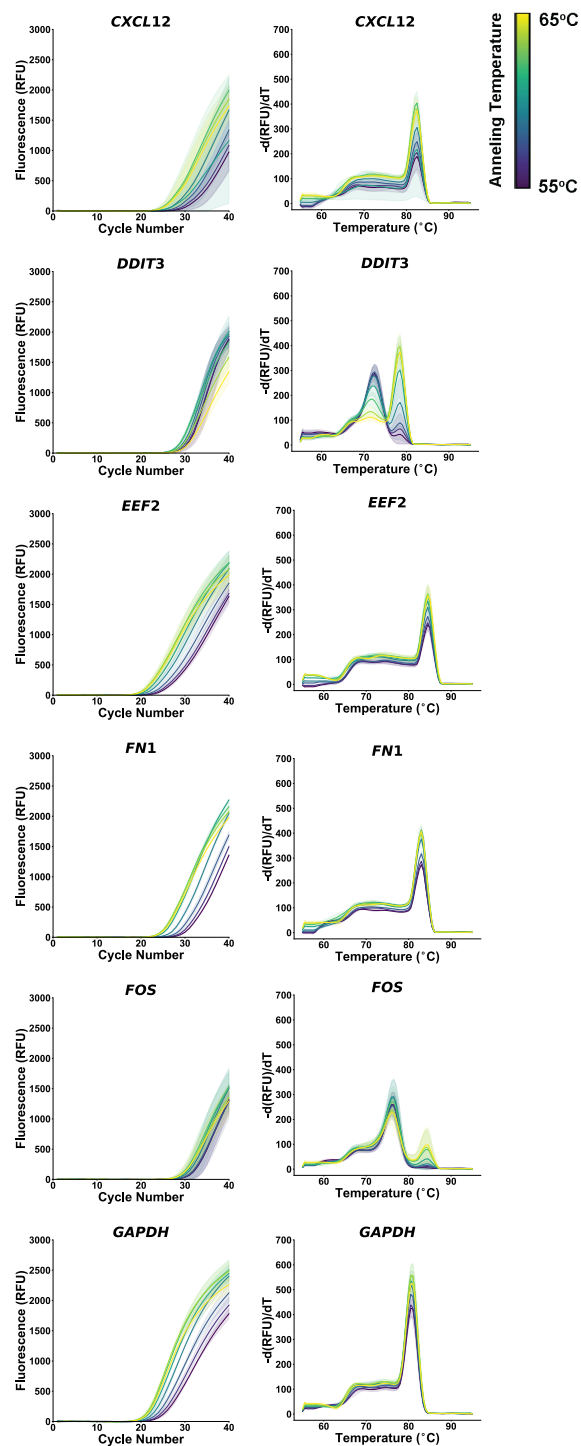


Supplementary Figure 35. CT versus score sub-types. Scatter plots of various scores and primer pair characteristics against normalized CT difference between a primer pair for a gene and the best CT for a pair for the same gene. Sample sizes $n=112$ for all the comparisons. Zero values were removed from the graphs. Source data are provided as a Source Data file. The code used in the creation of these data are published under a CC-BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>).



Supplementary Figure 36. Amplification curves and melt curves for 80-100 score mouse primers. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

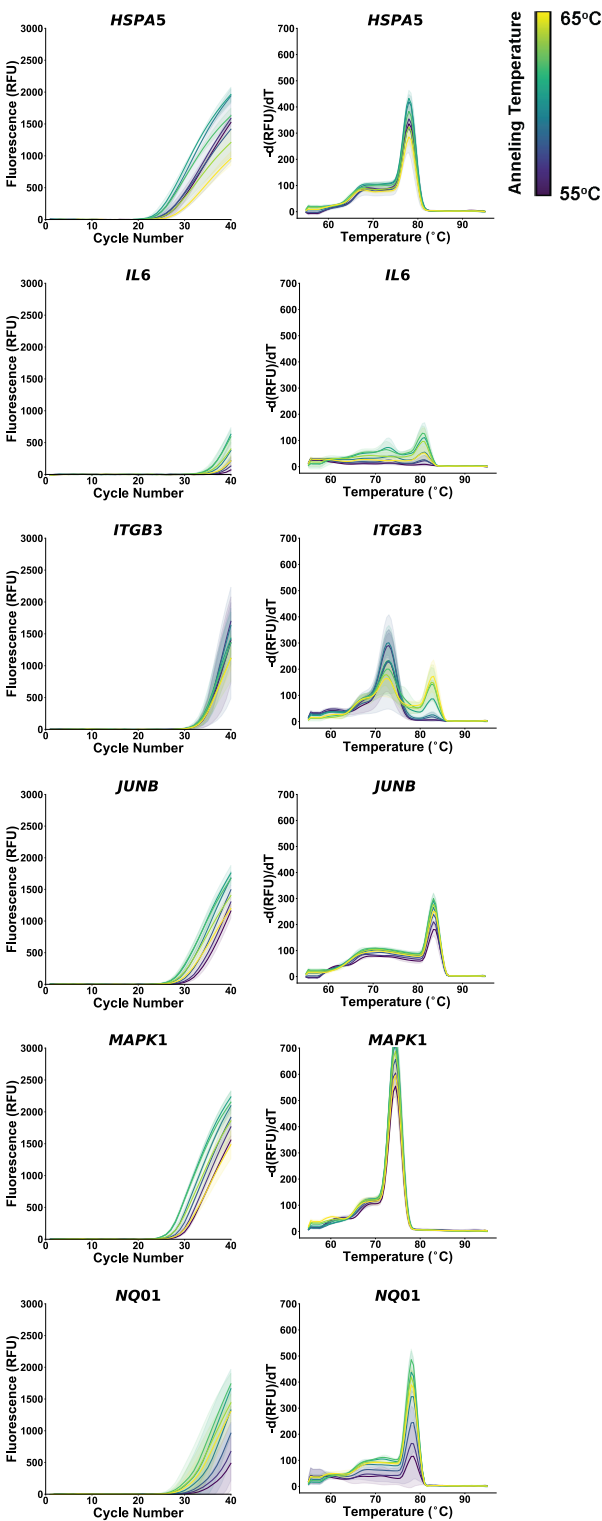
196
197



198

199 **Supplementary Figure 37. Amplification curves and melt curves for 80-100 score mouse**
 200 **primers.** 8 different temperatures were chosen to create amplification curves (left) and melt
 201 temperature curves (right). Source data are provided as a Source Data file.

202



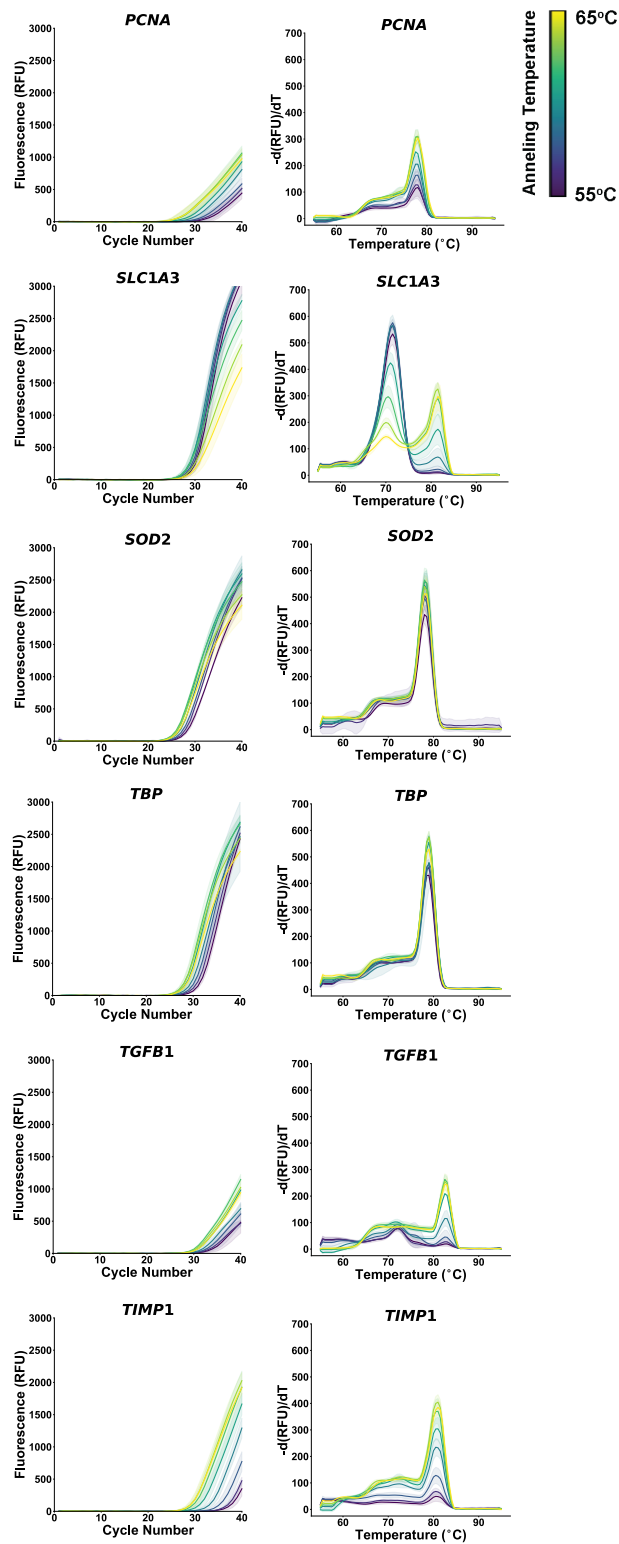
203

204

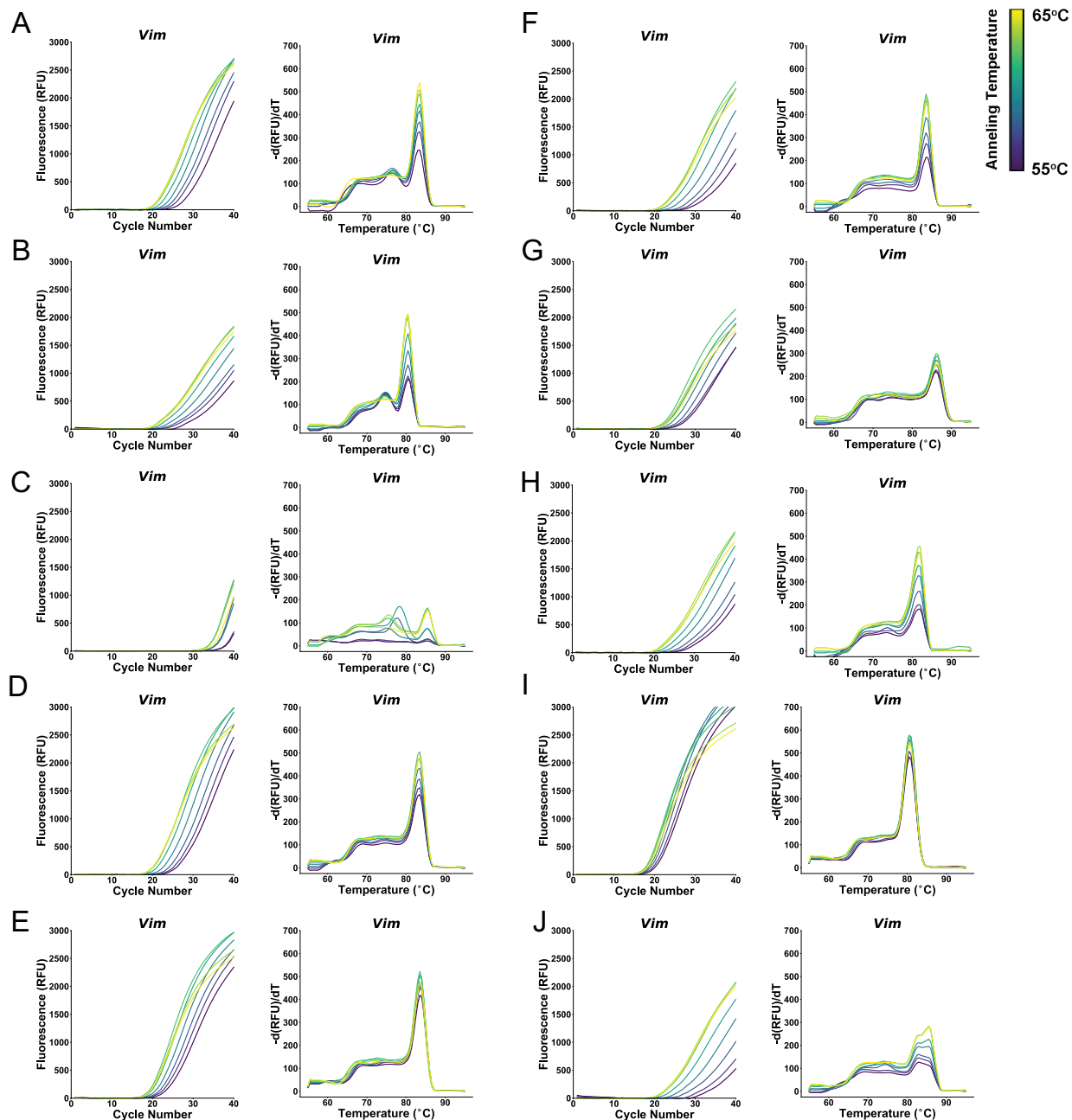
205

206

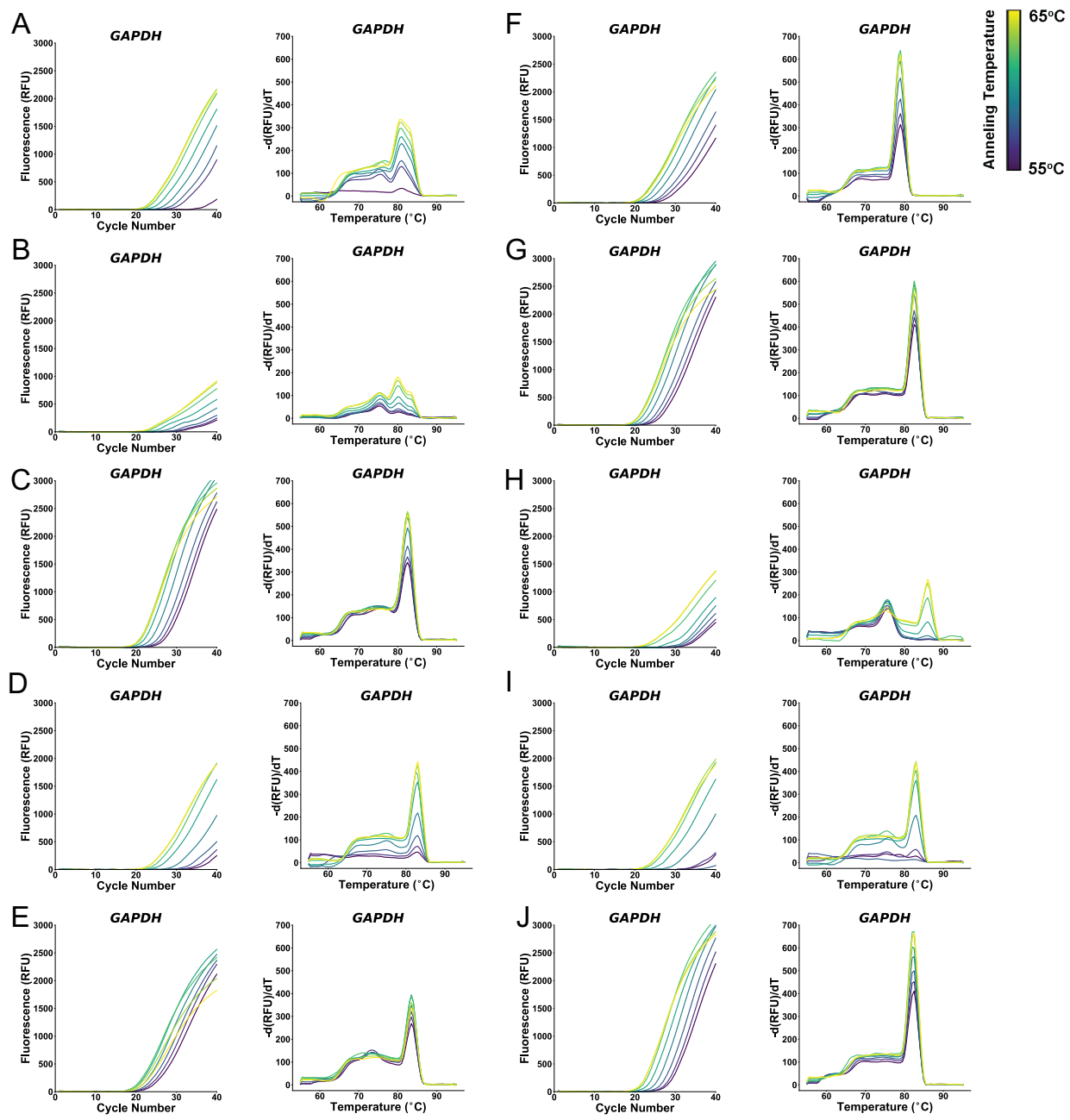
Supplementary Figure 38. Amplification curves and melt curves for 80-100 score mouse primers. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



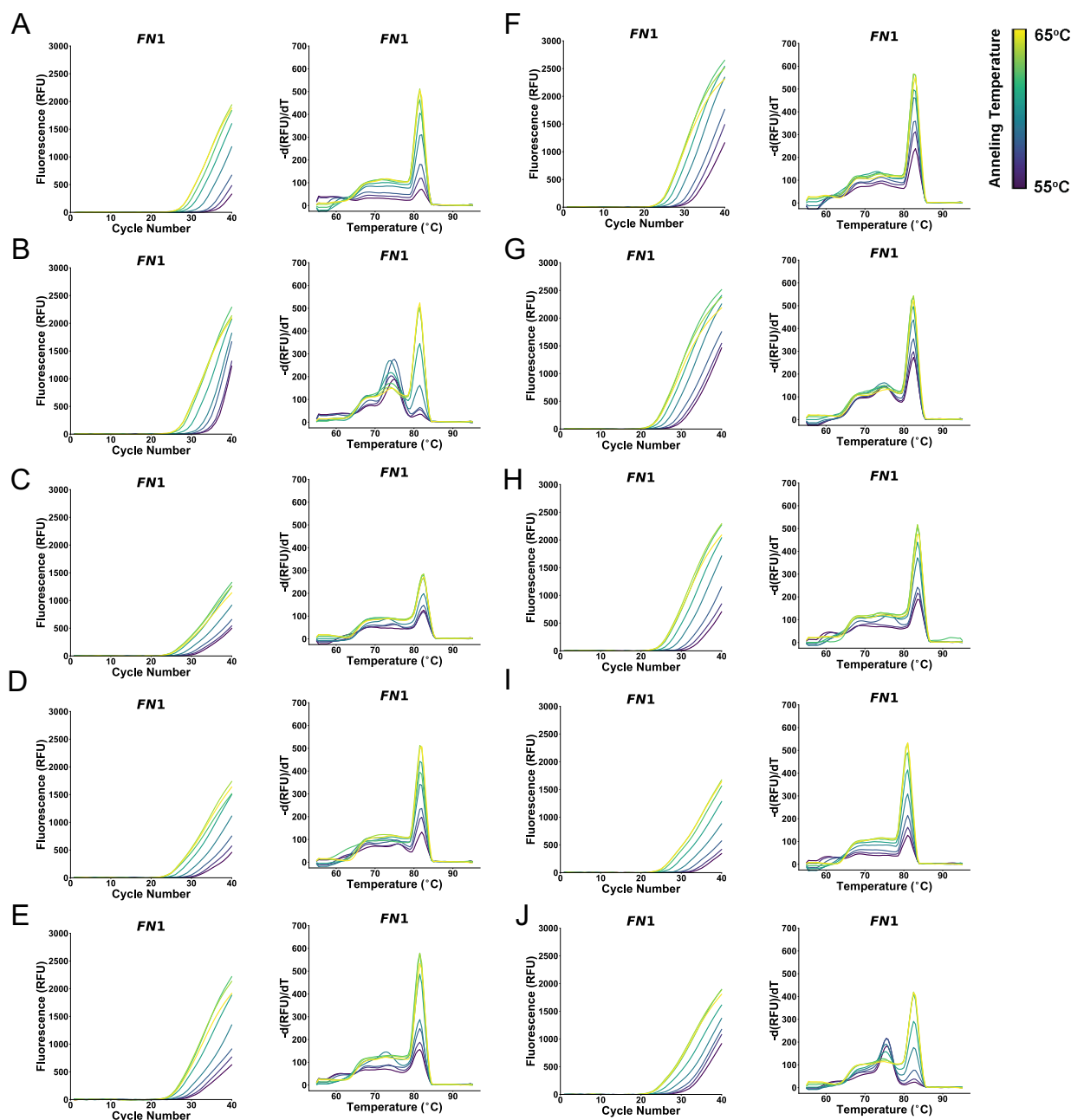
Supplementary Figure 39. Amplification curves and melt curves for 80-100 score mouse primers. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



Supplementary Figure 40. Amplification curves and melt curves for 10 primer pairs generated for human Vimentin (VIM) using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

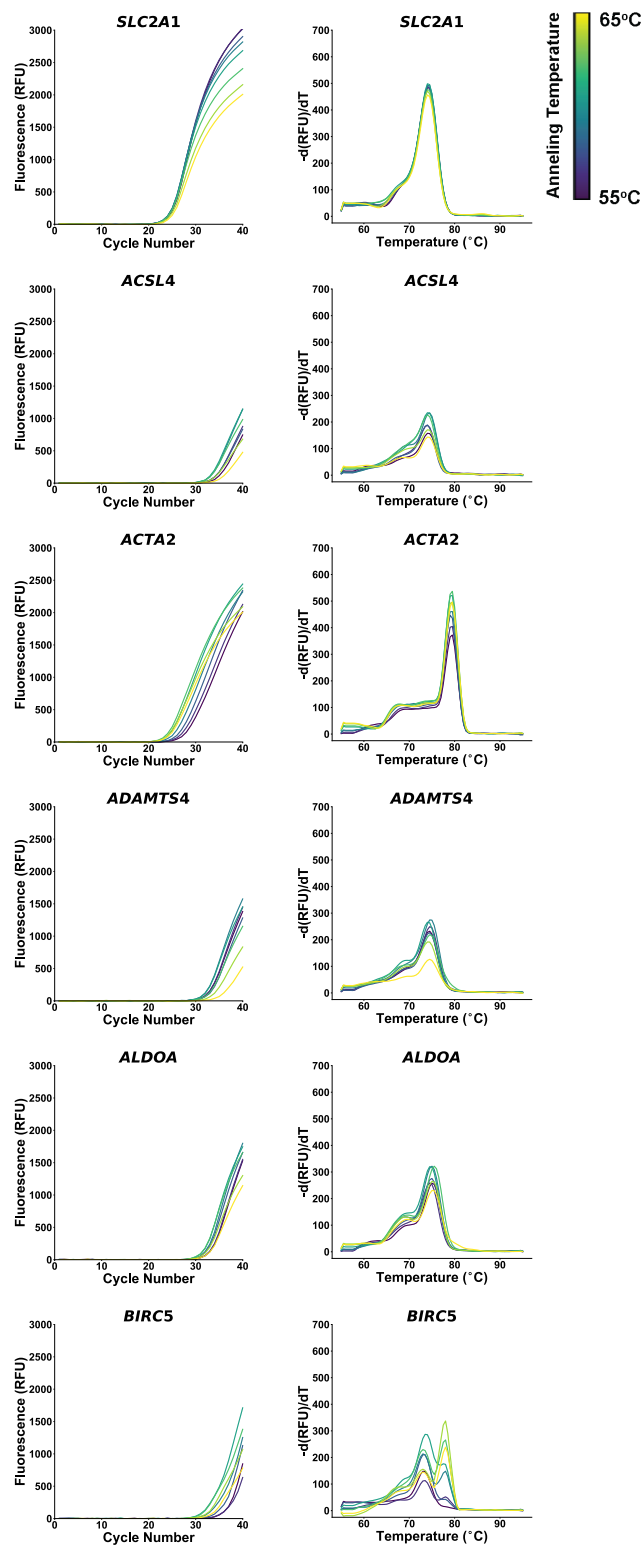


219 **Supplementary Figure 41. Amplification curves and melt curves for 10 primer pairs**
220 **generated for human GAPDH using Primer Blast. 8 different temperatures were chosen to**
221 **create amplification curves (left) and melt temperature curves (right). Source data are provided**
222 **as a Source Data file.**



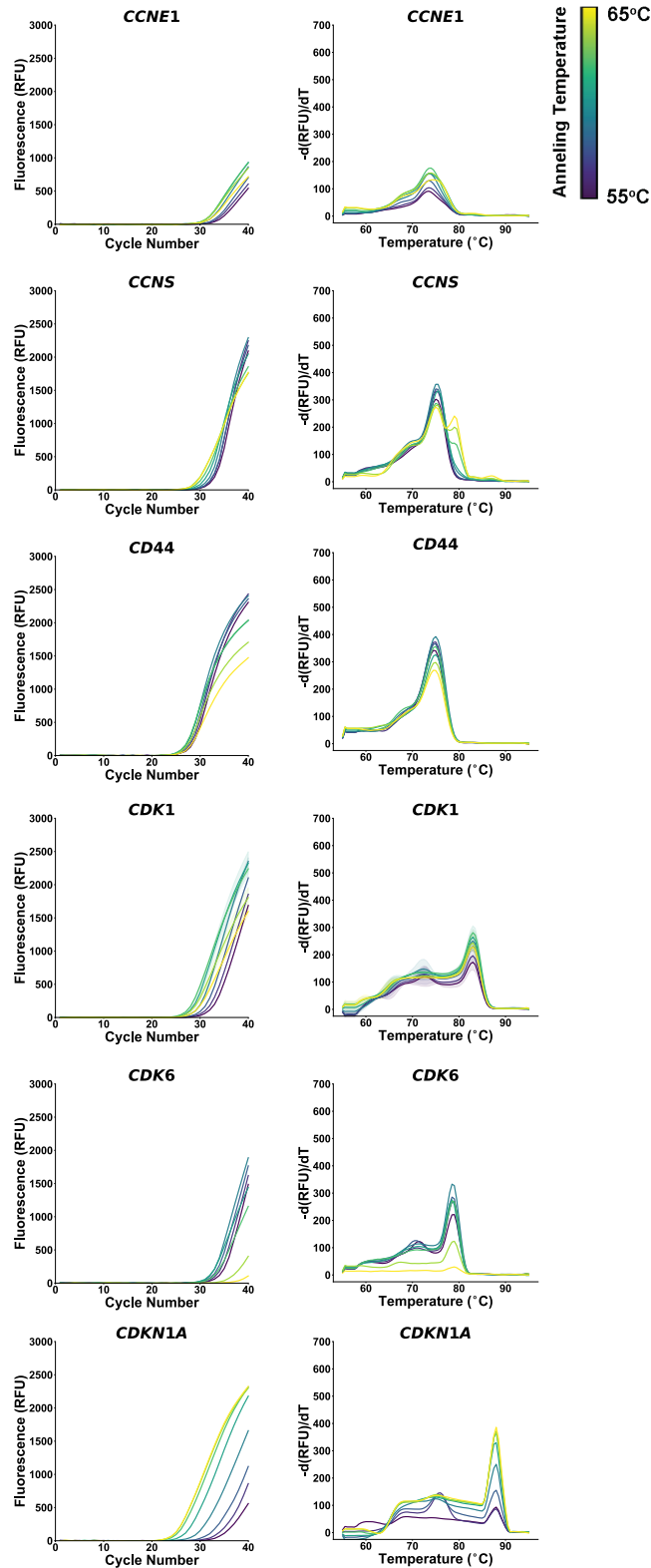
225
226
227
228
229

Supplementary Figure 42. Amplification curves and melt curves for 10 primer pairs generated for human Fibronectin 1 (FN1) using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

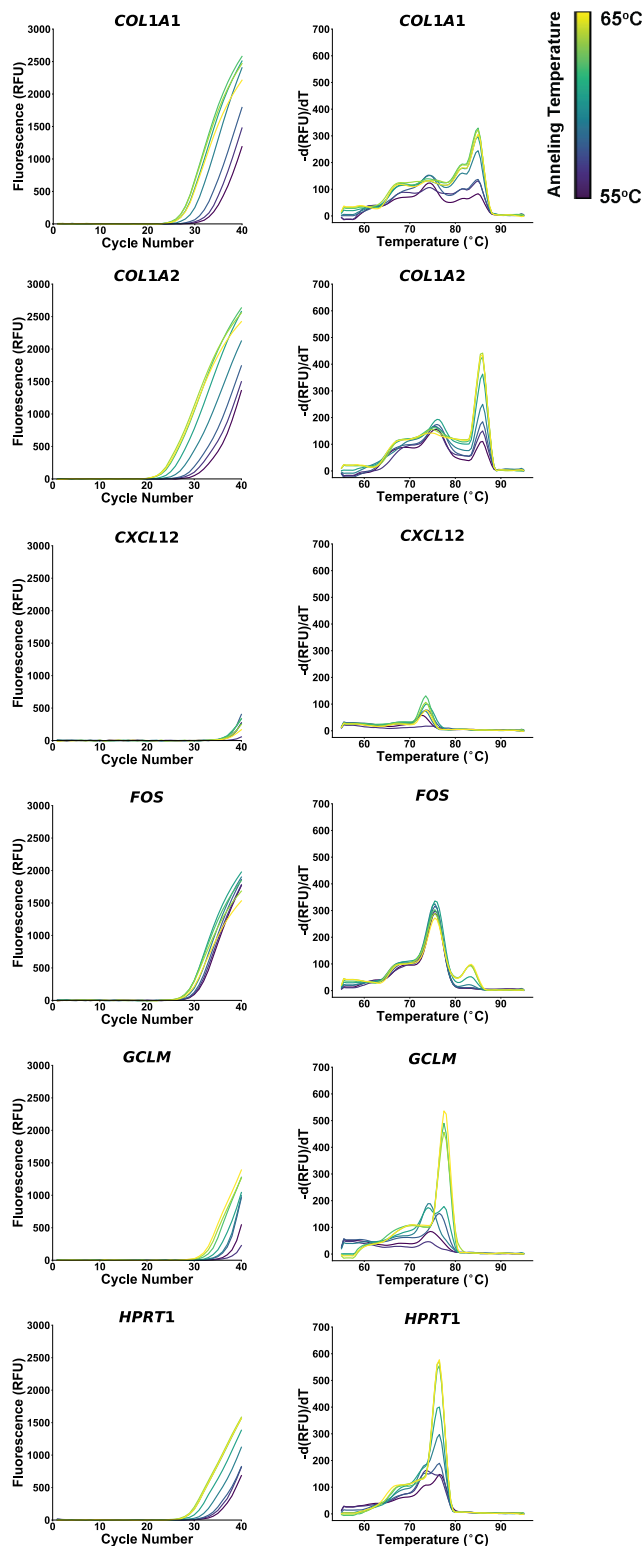


232
233
234

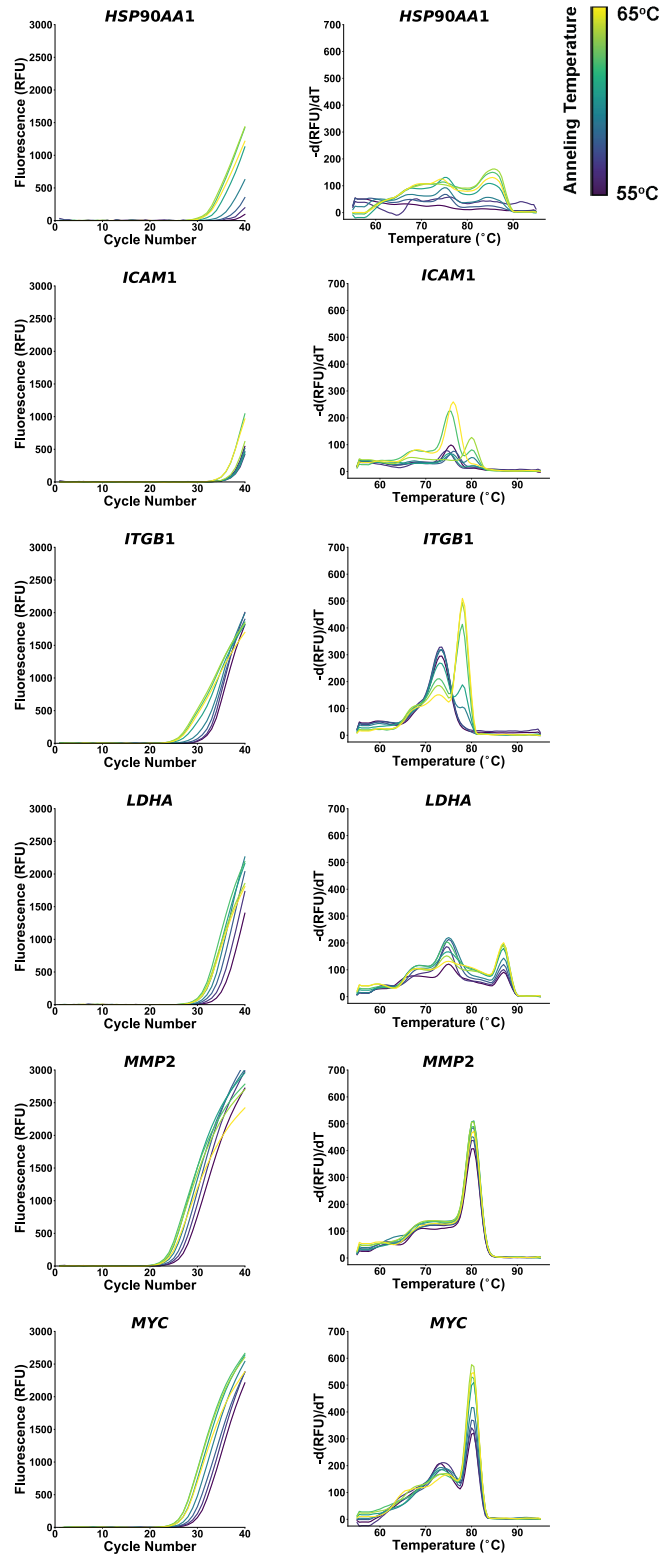
Supplementary Figure 43. Amplification curves and melt curves for primer pairs generated using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



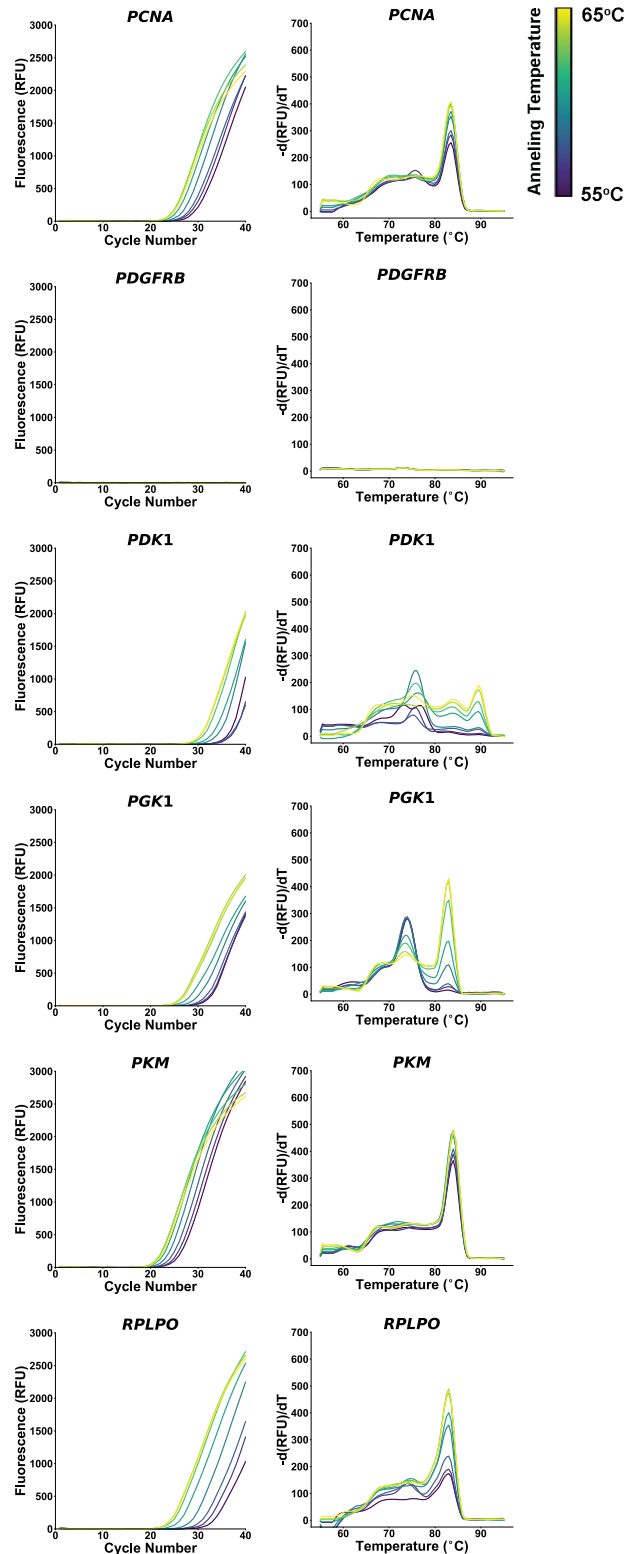
Supplementary Figure 44. Amplification curves and melt curves for primer pairs generated using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



Supplementary Figure 45. Amplification curves and melt curves for primer pairs generated using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



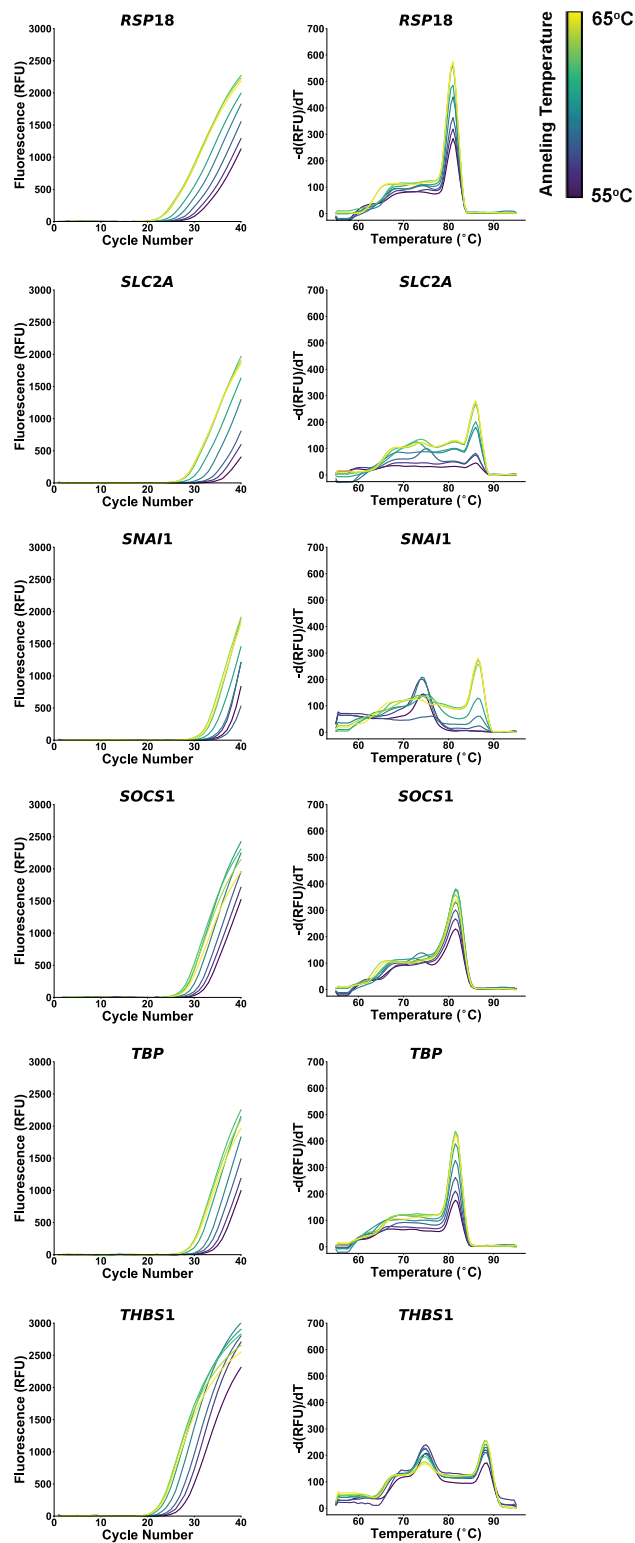
Supplementary Figure 46. Amplification curves and melt curves for primer pairs generated using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



247

248 **Supplementary Figure 47. Amplification curves and melt curves for primer pairs generated**
 249 **using Primer Blast. 8 different temperatures were chosen to create amplification curves (left)**
 250 **and melt temperature curves (right). Source data are provided as a Source Data file.**

251



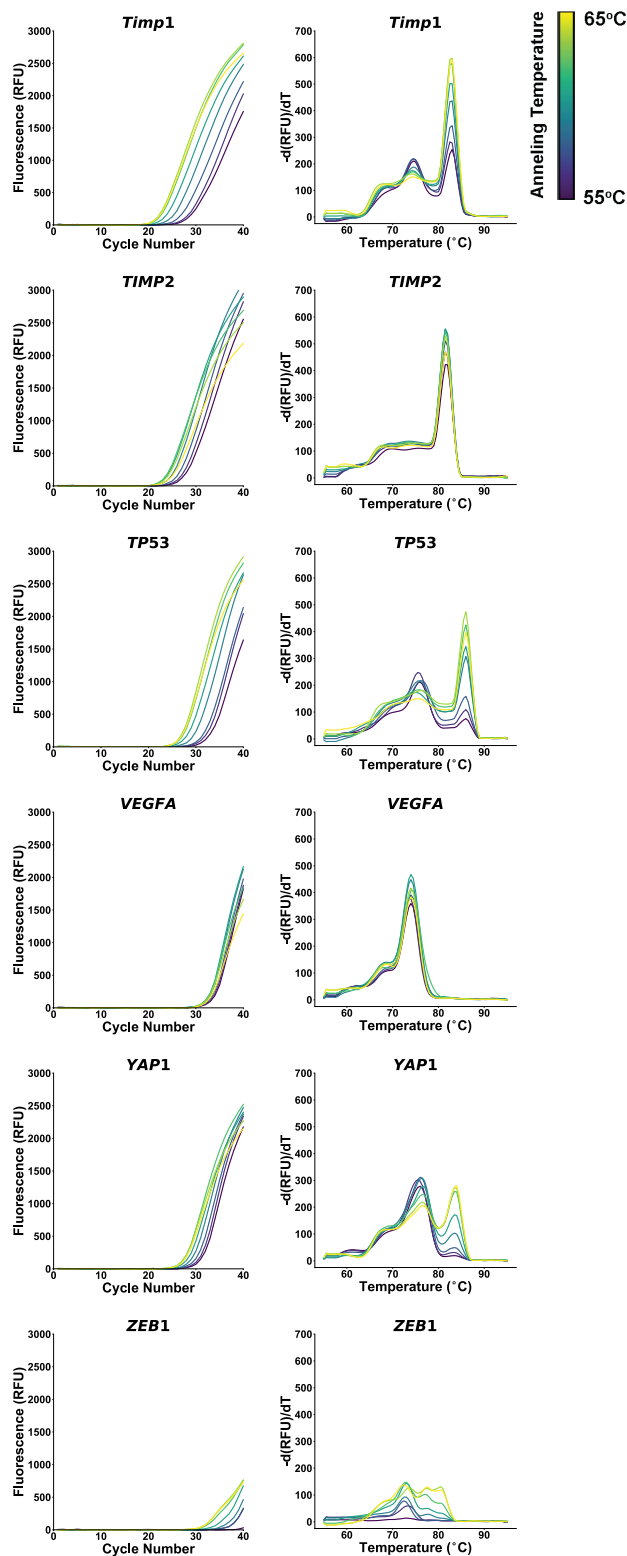
252

253

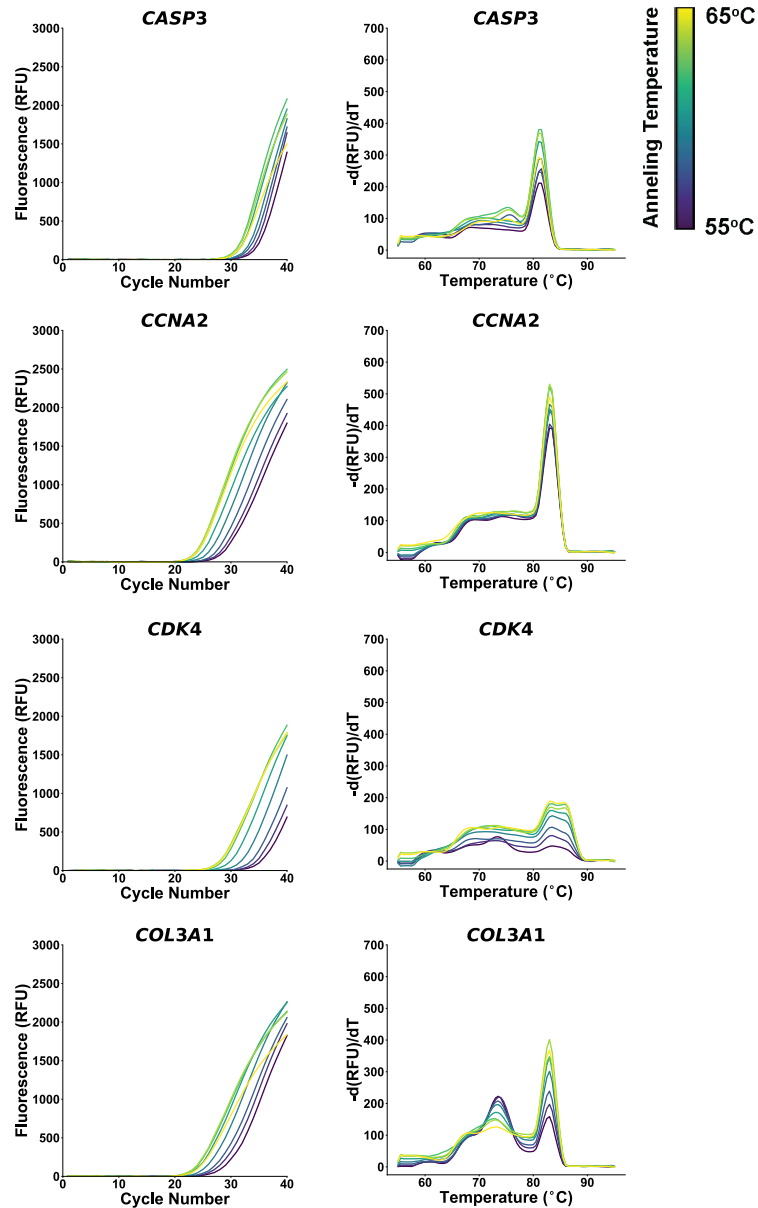
254

255

Supplementary Figure 48. Amplification curves and melt curves for primer pairs generated using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



Supplementary Figure 49. Amplification curves and melt curves for primer pairs generated using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.



Supplementary Figure 50. Amplification curves and melt curves for primer pairs generated using Primer Blast. 8 different temperatures were chosen to create amplification curves (left) and melt temperature curves (right). Source data are provided as a Source Data file.

2

→ ACTGCAACAAGGAATACCTCAG

← GCACTGGTACTTCTTGACATCTG

Cited Together in 50 Paper(s)

69

Combined Citations

Details

242 bp

Amplicon

59.5°C

Avg Tm

91.6

PICKR Score

98.9th percentile

Score Details

Forward Primer Details

Tm: 59.0°C

GC Content: 45.5%

Length: 22 bp

Self-Complementarity: 4.00

Off-Target Matches: 0

Reverse Primer Details

Tm: 60.0°C

GC Content: 47.8%

Length: 23 bp

Self-Complementarity: 4.00

Off-Target Matches: 0

Pair Details

ΔTm: 1.0°C

Pair Complementarity: 4.00

Shared Citations (50)

10147313

10157350

10241944

10543829

11151503

11233391

11264956

11277582

11571564

11720593

4637298

4909706

5041907

5122352

5406608

5410258

5418622

5505986

5543456

5595092

+30 more

Forward Citations (58)

10147313

10157350

10241944

10381260

10543829

11151503

11233391

11264956

11277582

11380517

11569164

11571564

11720593

4637298

4909706

5041907

5122352

5406608

5410258

5418622

+35 more

Reverse Citations (61)

10082504

10105420

10147313

10157350

10241944

10516749

10543829

10884757

11151503

11233391

11264956

11277582

11414211

11571564

11720593

4637298

4909706

5041907

5122352

5406608

+38 more

265
266
267
268
269

Supplementary Figure 51. Screen shot of the primer card from www.pirmerpickr.com. Cards display relevant data to help users consider primer details when choosing which primers to use. The code used in the creation of these data are published under a CC-BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>).