

Degraded RNA from human anterior cruciate ligaments yields valid gene expression profiles

Short title: Valid expression profiles from degraded RNA

Matthew D. Strawn^{1,#a}, Geoffrey D. Fisher², Cody J. Perry^{1,3}, Javad Hashemi⁴, Hossein Mansouri⁵, Matthew P. Ferguson³, Mimi Zumwalt³, George W. Brindley³, James R. Slauterbeck⁶, and Daniel M. Hardy^{1,3*}

¹ Department of Cell Biology & Biochemistry, Texas Tech University Health Sciences Center, Lubbock, Texas, United States of America

² Department of Anesthesiology, Virginia Commonwealth University Health System, Richmond, Virginia, United States of America

³ Department of Orthopaedic Surgery & Rehabilitation, Texas Tech University Health Sciences Center, Lubbock, Texas, United States of America

⁴ Department of Ocean & Mechanical Engineering, Florida Atlantic University, Boca Raton, Florida, United States of America

⁵ Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX

⁶ Department of Orthopaedic Surgery, University of South Alabama, Mobile, Alabama, United States of America

^{#a} Current address: Department of Emergency Medicine, Texas Tech University Health Sciences Center, Amarillo, Texas, United States of America

*Corresponding author
Email: Daniel.Hardy@ttuhsc.edu

ABSTRACT

Background: Correlating gene expression patterns with biomechanical properties of connective tissues provides insight into molecular processes underlying development and maintenance of the tissues' physical characteristics. Cadaveric tissues such as human knees are widely considered suitable for biomechanical studies, but their usefulness for gene expression experiments is potentially limited by unavoidable, nuclease-mediated degradation of RNA.

Methods: We quantified mRNAs encoding 84 extracellular matrix components and five housekeeping proteins in intact and partially degraded RNA from human anterior cruciate ligaments (ACLs), and compared relative amounts of the individual mRNAs by regression analysis.

Results: Human ACL RNA degraded *in vitro* by limited ribonuclease digestion (N=6) resembled degraded RNA isolated from cadaveric tissue. PCR threshold cycle (C_t) values from degraded RNAs ranged variably higher than values obtained from their corresponding non-degraded RNAs, reflecting both the expected loss of target templates in the degraded preparations as well as differences in the extent of degradation. Relative C_t values obtained for mRNA targets in degraded preparations correlated strongly with the corresponding target levels in non-degraded RNA, both for each ACL as well as for pooled results from all six ACLs. Nuclease-mediated degradation produced similar, strongly correlated losses of housekeeping and non-housekeeping gene mRNAs. Expression profiling of RNA degraded *in situ* yielded comparable results, confirming that *in vitro* digestion suitably mimicked degradation by endogenous ribonucleases in frozen and thawed ACL.

Conclusion: PCR-based expression analyses can yield valid mRNA profiles from partially degraded RNA preparations such as those obtained from cadaveric knees and other skeletal motion segments used in biomechanical studies. Legitimate comparisons between variably degraded tissues can be made by normalizing quantitative data to an appropriate housekeeping transcript.

INTRODUCTION

Profiles of expressed mRNAs reflect the physiological states of cells and tissues ^{1,2}. Consequently, comparative expression profiling can identify genes and metabolic pathways associated with onset, progression, and treatability of disease ^{3,4}. Approaches to mRNA profiling range from global analyses such as RNA-Seq or microarray screening of all expressed genes, to more focused, PCR-based profiling of gene subsets. Regardless of the approach, expression profiling methods universally call for use of only high quality, intact RNA to ensure that results accurately reflect the pattern of mRNAs present in living cells or tissues. Appropriate methods for tissue acquisition and handling and for RNA extraction and purification can yield intact, high quality RNA even from difficult tissues such as pancreas and ligament^{5,6}. However, RNA degradation prior to isolation is sometimes unavoidable in precious specimens such as human cadaveric tissues that cannot always be handled and stored under ideal conditions, thus potentially limiting the usefulness of RNA from these tissues for expression analyses.

RNA profiling studies have characterized molecular responses of connective tissues to injury in non-human animals ^{6,7}, and identified sex differences in gene expression that may underlie susceptibility of female athletes to anterior cruciate ligament (ACL) injury ^{8,9}. However, RNA integrity often cannot be preserved completely in biomechanical studies of human cadaveric tissues, which confounds studies seeking, for example, to correlate gene expression with the human ACL's ultrastructure and biomechanical properties (refs. 10, 11, and unpublished observations), or to identify expression patterns associated with degeneration and biomechanical properties of human intervertebral discs (Fisher et al., unpublished). Some studies have shown that meaningful expression data can be obtained using degraded RNA from sources such as rectal tumors ¹² or cultured cells ¹³, but none have assessed validity of profiles generated using degraded RNA from human connective tissues commonly used for biomechanical studies.

Here, we tested the hypothesis that gene expression profiles generated using partially degraded RNA from human ACLs accurately reflect those generated using intact RNA.

METHODS

Tissue acquisition. We recovered human ACL tissues, by IRB-approved protocol (TTUHSC IRB# L06-154), during total knee arthroplasty or ACL reconstruction surgeries (8 male, 5 female; ages 29-73 years, mean 62.3 years), and immediately froze [N₂ (*l*)] the tissues for storage at -80 °C. For human intervertebral disc tissue, we excised the L4-5 disc from the frozen spine of a willed body (non-embalmed female, age 40; spine stored at -20°C). As is commonplace for donated bodies, no information was available regarding the time between donor death and freezing of the spine, or the specific conditions of its storage prior to receipt in the lab.

RNA isolation. We isolated RNA from frozen [N₂ (*l*)], thoroughly pulverized tissues (whole ACL or the outer anterior segment of the intervertebral disc's annulus fibrosus) using either the RNeasy Mini Kit (100-350 mg of tissue; QIAGEN, Valencia, CA) or the guanidinium thiocyanate/acidic phenol extraction⁸ method (1 g tissue) and assessed RNA purity and concentration by spectrophotometry (NanoDrop™, Thermo Fisher Scientific, Wilmington, DE).

RNA degradation. For *in vitro* degradation of ACL RNA, we digested paired aliquots (4 µg, isolated by RNeasy kit) either with RNase ONE™ (0.005 units, 37 °C, 1 h; Promega, Madison, WI) to produce 'partially degraded' RNA, or with no added enzyme to produce a mock-digested, 'intact' control RNA, and assessed extent of degradation by formaldehyde/agarose gel electrophoresis (1 µg RNA/lane)^{8, 14}.

For *in situ* degradation of ACL RNA, we incubated 1 g of thawed, pulverized ACL for 20 min at 23°C to permit autolytic degradation of RNA by endogenous RNases, isolated RNA by the guanidinium thiocyanate/acidic phenol extraction method⁸, then assessed RNA quality with

an Agilent Technologies 2100 Bioanalyzer (Santa Clara, CA). A matching 1 g portion of unthawed, pulverized ACL served as a non-degraded control.

Quantitative PCR. We used first-strand complementary DNA (cDNA) synthesized from 500 ng of RNA with combined oligo-dT and random hexamer priming (RT2 First Strand Kit; SABiosciences, Frederick, MD) as template for real time PCR (Human Extracellular Matrix and Adhesion Molecules RT2 Profiler™ PCR Array kit; SABiosciences), with SYBR Green detection. We calculated threshold cycle (C_t) values at threshold=0.4 from amplification plots (ABI Prism 7000 Sequence Detection System; Applied Biosystems, Foster City, CA), and omitted from further analysis any genes the software identified as “undetermined” because they did not amplify sufficiently during the exponential phase. All five housekeeping genes in the array amplified successfully in all profiles, so downstream correlation analyses (see below) included C_t values from the full complement of these genes.

Data analysis. To determine the effect of RNA degradation on expression profile, we correlated C_t values obtained from degraded RNAs and non-degraded RNAs using hierarchical regression models. We first established a linear relationship by analyzing all data from the six ACL RNA preparations, then produced individual regression equations for ACL RNAs by treating them as random samples from the population regression, and examined the assumption of randomness of the regression coefficients by fitting hierarchical linear regressions with random slopes and random intercepts. To analyze variation in degradation, we averaged the differences in C_t values between intact and degraded RNAs for each gene in each of the six ACLs, then plotted the averages for housekeeping and non-housekeeping genes against each other to test for correlation. Finally, we performed two-tailed t-tests to identify differences in extent of degradation between the housekeeping and non-housekeeping groups.

RESULTS

Limited RNase digestion of ACL RNAs yielded degraded preparations with electrophoretic properties similar to partially degraded RNA isolated from human cadaveric intervertebral disc (Fig. 1). Notwithstanding identical digestion conditions, the six degraded ACL RNA preparations exhibited subtle electrophoretic differences, with RNAs from ACL #98 and #104 appearing slightly less degraded than the other four (Fig. 1). Using hierarchical models, regression analysis of all C_t value pairs (degraded vs. intact) obtained from the six ACL RNAs revealed a strong linear relationship (Pearson correlation coefficient $r=+0.80$) given by the equation:

$$C_{t_degraded} = 5.769 + 0.906 \times C_{t_intact}$$

The line intercept (95% confidence interval 3.254 - 8.284 cycles) reflected an increase in the number of PCR cycles required to reach the threshold signal as a consequence of degradation-induced decreases in the amounts of amplifiable mRNAs for the 89 targets. The line slope near but slightly below 1.0 (95% confidence interval: 0.868 – 0.943) revealed that degradation produced a generally uniform increase in C_t irrespective of an individual RNA's abundance, with a possible trend toward proportionally greater loss of higher abundance mRNAs.

Fig. 1. Formaldehyde-agarose electrophoresis of intact and partially degraded RNA preparations.

We evaluated paired samples of intact and *in vitro*-degraded RNA from six ACLs (designated #105, 108, 125, 98, 99, and 104). Note the loss of ribosomal RNA bands and appearance of smaller-sized RNA fragments upon RNase digestion of the ACL RNAs, as well as the similarity of the electrophoretic patterns to the degraded RNA isolated from cadaveric intervertebral discs.

Hierarchical analysis of individual regression lines for the six RNA preparations (see Fig. 2 for example individual regression plots) revealed strong linear relationships (range of Pearson

coefficients: 0.86-0.98; Table 1), with non-random slopes of the six lines each fixed at 0.906 as for the population overall. In contrast, randomly distributed intercepts ranged from 2.93 – 7.48, reflecting possibly stochastic variation in the extent of RNA degradation between preparations. Mean C_t increase ranged from 0.75 cycles for the 86 targets detected in degraded ACL 104 RNA, to 5.2 cycles for the 80 targets detected in degraded ACL 105 RNA, corresponding to 40% and 97% average losses of amplifiable target, respectively [fraction remaining = $f = 2^{\exp(-\Delta C_t)}$; % loss = $(1 - f) \times 100$]. The randomly distributed y-intercept (b) values correlated inversely with the number of targets quantified (N) and with magnitude of the Pearson coefficient (r), indicating that as average C_t increased because of RNA degradation, more targets were lost and the scatter of the data increased. Nonetheless, even at the highest levels of degradation (ACLs 99, 105, 108, and 125), profiles from individual degraded RNAs correlated strongly with those from the corresponding intact RNAs ($r \geq +0.86$). Likewise, the Coefficients of Determination (R^2) ranging from 0.75-0.96 indicated that 75-96% of the variation in target signal from degraded RNAs reflected the biologically relevant distribution of target signals from the corresponding intact RNA.

Fig. 2. Example linear regression plots of C_t values from intact vs. *in vitro*-degraded RNA for two ACLs (#98 and #125) representing relatively less or more degradation, respectively (compare to Fig. 1). See Table 1 for a summary of results from corresponding analyses of degraded RNA preparations from all six ACLs.

ACL #	Targets Compared (N)	Slope (m)	y-intercept (b)	Pearson Coefficient (r)	Coefficient of Determination (R^2)
Combined	497	0.906	5.77	+0.80	0.64
98	87	0.906	3.28	+0.97	0.94
99	83	0.906	7.19	+0.88	0.77
104	86	0.906	2.93	+0.98	0.96
105	80	0.906	7.48	+0.87	0.76
108	78	0.906	7.01	+0.88	0.77
125	83	0.906	6.72	+0.86	0.75

Table 1. Correlation of mRNA expression profiles from partially degraded ACL RNAs with profiles from their corresponding intact RNAs. Differences in N (number of genes compared) reflect differences in the number of C_t values returned for each ACL, with N=89 representing no “undetermined” targets for an individual ACL owing to degradation of RNA target template to a level below that required for PCR amplification. Note that all six ACL RNAs yielded lines with slope 0.906, the same as for the combined data, but with randomly distributed y-intercept values (b). Note also the higher Coefficients of Determination, ranging up to a very strong positive correlation of 0.96, for all individual lines in comparison to the combined data.

For each of our degraded RNA preparations, mean increases in C_t for housekeeping and non-housekeeping genes did not differ (t-tests; $0.25 < p < 0.40$). Furthermore, regression analysis revealed strong correlations ($r > 0.999$; Fig. 3) between the increases in housekeeping and non-housekeeping C_t among the six degraded RNA preparations, spanning a wide range of mean C_t increases corresponding to levels of RNA degradation from approximately 50% to more than 95%.

Fig. 3. Regression analysis of relative RNA degradation effects on mean C_t values for housekeeping and non-housekeeping genes among the 89 targets evaluated. Note the strong linear relationship spanning a range of C_t increases from approximately one cycle in the least

degraded preparations (corresponding to ~50% mean loss of target mRNAs) to approximately six cycles in the most highly degraded preparations (corresponding to >95% mean loss of target mRNAs). Black data points are results from *in vitro* RNA digestions, and the red data point is from *in situ* RNA degradation in frozen and thawed ACL (see next).

To determine if RNase digestion of RNA *in vitro* accurately modeled the *in situ* degradation of RNA in poorly stored tissues, we thawed ACL tissue to permit action of endogenous RNases prior to RNA isolation, and compared the gene expression profile from the degraded RNA to that obtained with intact RNA. Thawing 20 min before initiating extraction produced overt degradation, with decreased amounts of intact rRNAs and increased amounts of smaller digestion products (Fig. 4A) in comparison to RNA from unthawed ACL. Nevertheless, as observed for *in vitro* digested RNAs, the gene expression profile obtained using RNA from thawed ACL correlated strongly with the profile from unthawed ACL ($R^2 = 0.92$). Furthermore, *in situ* digestion produced a slope of the degraded-vs.-intact regression line (0.915, Fig. 4B) very close to slope of 0.906 from the *in vitro* digestion experiments. Finally, *in situ* digestion produced very similar losses of housekeeping and non-housekeeping gene mRNAs (mean C_t increases of 2.2 and 2.3 respectively, reflecting more than 75% loss of amplifiable targets in the tissue thawed for 20 min), comparable (red data point in Fig. 3) to the relative losses observed in the *in vitro* digestion experiments.

Fig. 4. Comparison of expression profiles from intact and *in situ*-degraded RNA. Panel A shows Bioanalyzer traces of RNA isolated 0 min (upper trace) or 20 min (lower trace) after thawing the ACL tissue at 23°C. The peaks at 44 and 52 min are ribosomal RNAs. For the RNA isolated 20 min post-thaw, note the increased baseline and smaller rRNA peaks, reflecting the autolytic degradation of RNAs to smaller products. Panel B shows the corresponding linear regression plot

of Ct values from the intact vs. *in situ*-degraded RNAs. Note the strong correlation ($R^2 = 0.92$) between the profiles and the very similar slope of the regression line (0.91) to that observed in the *in vitro* digestion experiments (Table 1).

DISCUSSION

Here we found that partial degradation of human ACL RNA decreased the amounts of individual mRNAs amplified in RT-PCR-based expression analyses but did not dramatically alter the overall gene expression profile in comparison to intact RNA. We also found no significant difference between the loss of housekeeping and loss of non-housekeeping mRNAs within degraded RNA preparations. Therefore, by normalizing to levels of housekeeping mRNAs, legitimate comparisons can be made between profiles of mRNAs with varying levels of degradation. Indeed, up to 97% mean loss of amplifiable target still yielded profiles that correlated strongly with profiles obtained from intact RNA, albeit with loss of signals from some low abundance targets. Further digestion of RNA would be required to define the point at which excessive degradation of ACL RNA yields invalid expression profiles. Regardless, our results revealed that valid, PCR-based expression profiles can be generated from RNA that is much more highly degraded than is generally thought to be required for such studies.

ACL RNA digested *in vitro* with RNase ONE™ could differ from RNA degraded by endogenous cellular ribonucleases. RNase ONE™ is a broad specificity enzyme from *E. coli* that cleaves between any two ribonucleotides, whereas human cellular ribonucleases exhibit more restricted nucleobase specificities¹². Nevertheless, differences in enzyme action are unlikely to affect our findings for two reasons. First, the promiscuous activity of RNase ONE™ theoretically approximates the combined specificities of the multiple RNases¹⁵ in human cells that together mediate *in situ* RNA degradation in a poorly handled or stored tissue. Second, hydrolysis of an

mRNA anywhere between the upstream PCR primer sequence and the downstream priming site for reverse transcriptase would render it incapable of serving as a template for RT-PCR. Consequently, any endogenous RNase, including one that preferentially hydrolyzed at certain ribonucleotides (such as the CpX specificity of pancreatic RNase A ¹⁶), would be expected to digest target mRNAs sufficiently to cause loss of RT-PCR signal much as we achieved by digestion with RNase ONE™. Indeed, expression analysis of RNA degraded *in situ* in frozen-and-thawed ACL confirmed the conceptual soundness of the *in vitro* digestion approach in this study.

Our results do not rule out the possibility that certain individual RNAs might be highly sensitive to hydrolysis by specific RNases; however, use of RNA that is presumed to be completely intact also does not necessarily preclude that possibility. Of note, by using a combination of oligo-dT and random hexamers to prime the RT step for cDNA synthesis, we maximized yields of first strand template for downstream amplification by PCR. Thus, even RNA preparations degraded entirely to relatively small fragments could still generate expression signals so long as individual mRNA fragments both spanned their corresponding PCR-amplified target sequences and extended far enough downstream to provide a site for oligo-dT or random hexamer priming. Accordingly, the results of this study do show that useful expression information for most genes can be obtained from human connective tissues acquired in a way that does not assure complete integrity of RNA (e.g., from biopsies or cadaveric specimens).

Although RNA in poorly stored or handled tissues will ultimately degrade to the point that valid expression data cannot be obtained, we conclude it is nevertheless possible to generate meaningful results from partially degraded RNA so long as sufficient mRNA template remains for PCR amplification. This finding is particularly relevant to research correlating gene expression profiles with biomechanical properties of cadaveric tissues, including studies on sex

differences in molecular and biomechanical properties of the human ACL, and on age- and pathology-associated differences in the molecular and biomechanical properties of intervertebral discs.

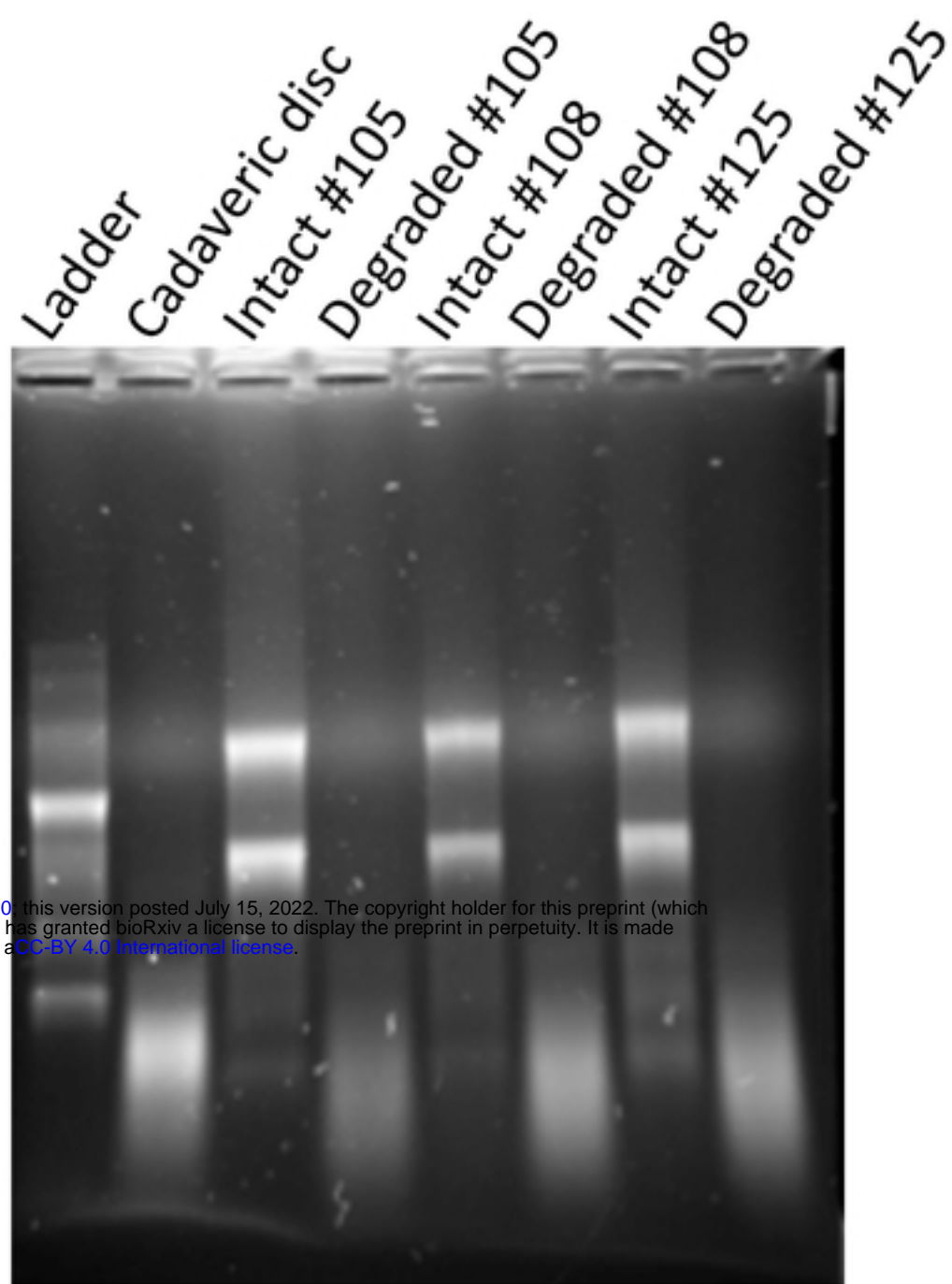
ACKNOWLEDGEMENTS

We thank Elizabeth White for outstanding technical assistance. DMH and JRS received grant number AR049767 from the National Institute for Arthritis and Musculoskeletal and Skin Diseases (<https://www.niams.nih.gov/>), and DMH and JH received funding from the TTU/TTUHSC Clinical/Basic Science Seed Grant Program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

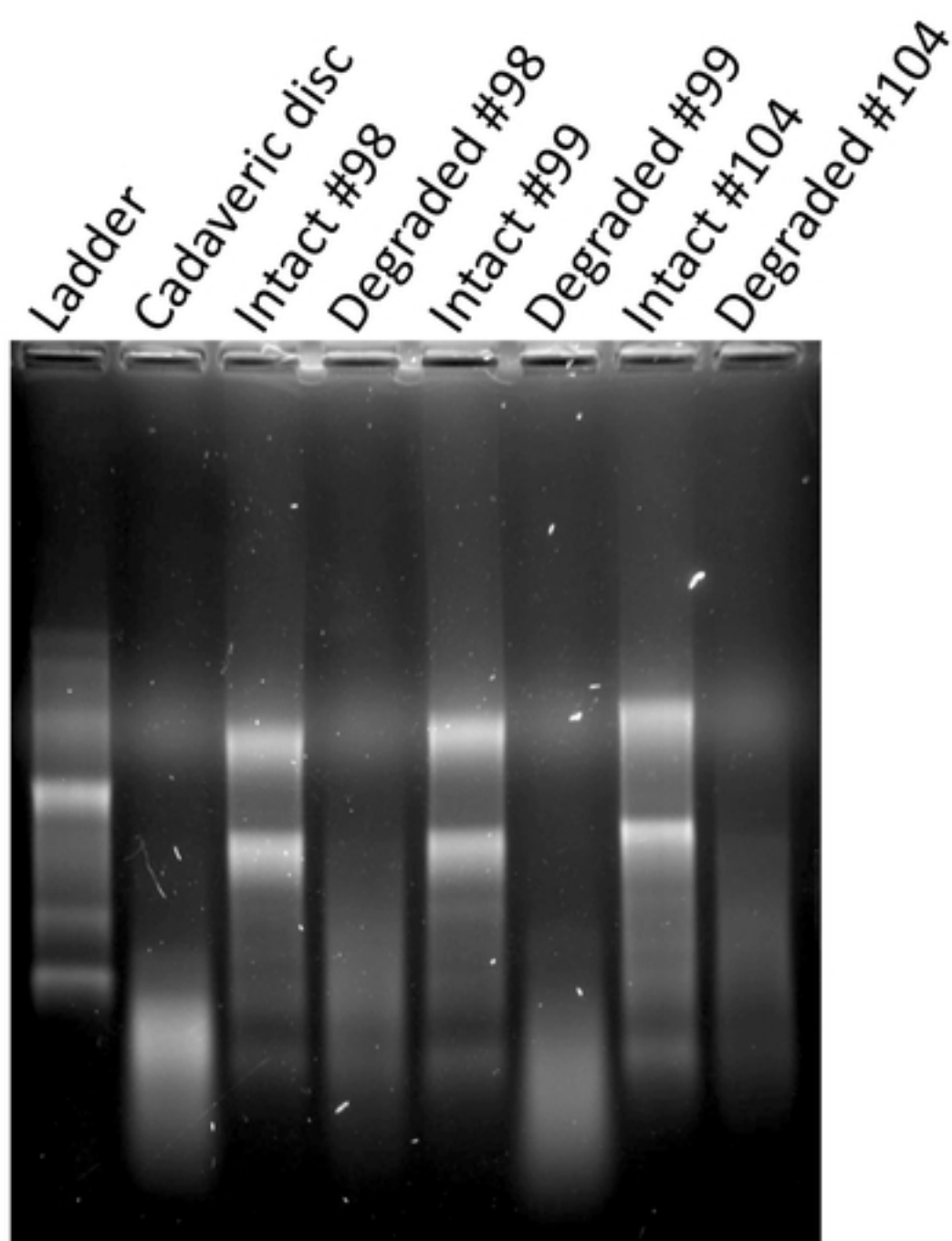
REFERENCES

1. Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JC, et al. The transcriptional program in the response of human fibroblasts to serum. *Science* 1999; 283: 83-87
2. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci USA* 1999; 96: 9212-9217
3. Chang HY, Chi JT, Dudoit S, Bondre C, van de Rijn M, Botstein D, et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc Natl Acad Sci USA* 2002; 99: 12877-12882
4. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415: 530-536
5. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987; 162: 156-159
6. Attia E, Brown H, Henshaw R, George S, Hannafin JA. Patterns of gene expression in a rabbit partial anterior cruciate ligament transection model: the potential role of mechanical forces. *Am J Sports Med* 2010; 38: 348-356

7. Achari Y, Chin JWS, Heard BJ, Rattner JB, Shrive NG, Frank CB, Hart DA. Molecular events surrounding collagen fibril assembly in the early healing rabbit medial collateral ligament-failure to recapitulate normal ligament development. *Connect Tiss Res* 2011; 52: 301-312
8. Foos MJ, Hickox JR, Mansour PG, Slauterbeck JR, Hardy DM. Expression of matrix metalloproteinase and tissue inhibitor of metalloproteinase genes in human anterior cruciate ligament. *J Orthop Res* 2001; 19: 642-649
9. Slauterbeck JR, Hardy DM. Sex hormones and knee ligament injuries in females. *Am J Med Sci* 2001; 322: 196-199
10. Chandrashekar N, Mansouri H, Slauterbeck J, Hashemi J. Sex-based differences in the tensile properties of the human anterior cruciate ligament. *J Biomech* 2006; 39: 2943-2950
11. Hashemi J, Chandrashekar N, Mansouri H, Slauterbeck JR, Hardy DM. The human anterior cruciate ligament: sex differences in ultrastructure and correlation with biomechanical properties. *J Orthop Res* 2008; 26: 945-950
12. Opitz L, Salinas-Riester G, Grade M, Jung K, Jo P, Emons G, Ghadimi BM, et al. Impact of RNA degradation on gene expression profiling. *BMC Med Genomics* 2010; 3: 36
13. Tong D, Schneeberger C, Leodolter S, Zeillinger R. Quantitative determination of gene expression by competitive reverse transcription-polymerase chain reaction in degraded RNA samples. *Anal Biochem* 1997; 251: 173-177
14. Hardy DM, Garbers DL. A sperm membrane protein that binds in a species-specific manner to the egg extracellular matrix is homologous to von Willebrand factor. *J Biol Chem* 1995; 270: 26025-26028
15. Kim WC, Lee CH. The role of mammalian ribonucleases (RNases) in cancer. *Biochim Biophys Acta* 2009; 1796: 99-113
16. Raines RT. Ribonuclease A. *Chem Rev* 1998; 98: 1045-1065

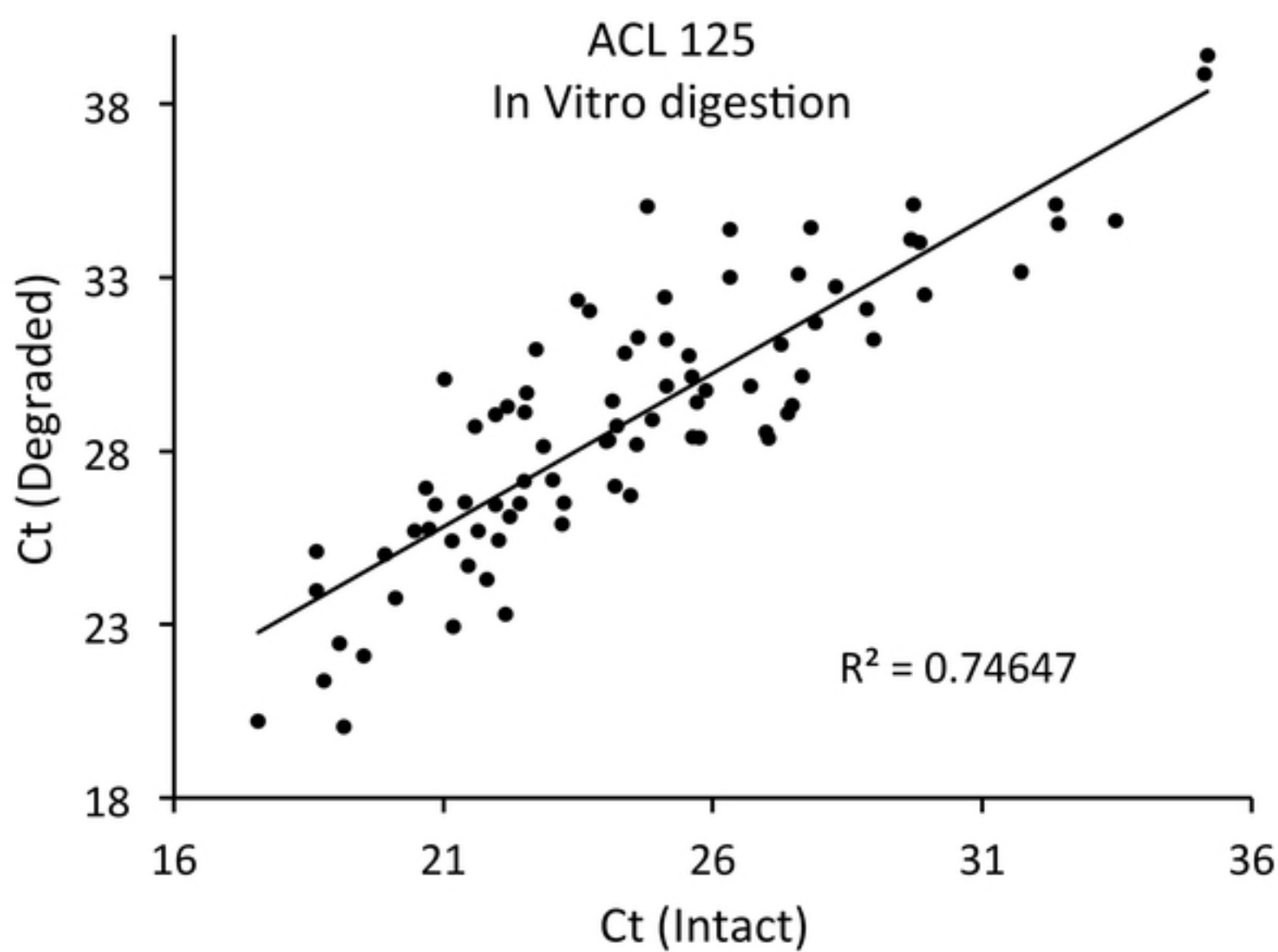
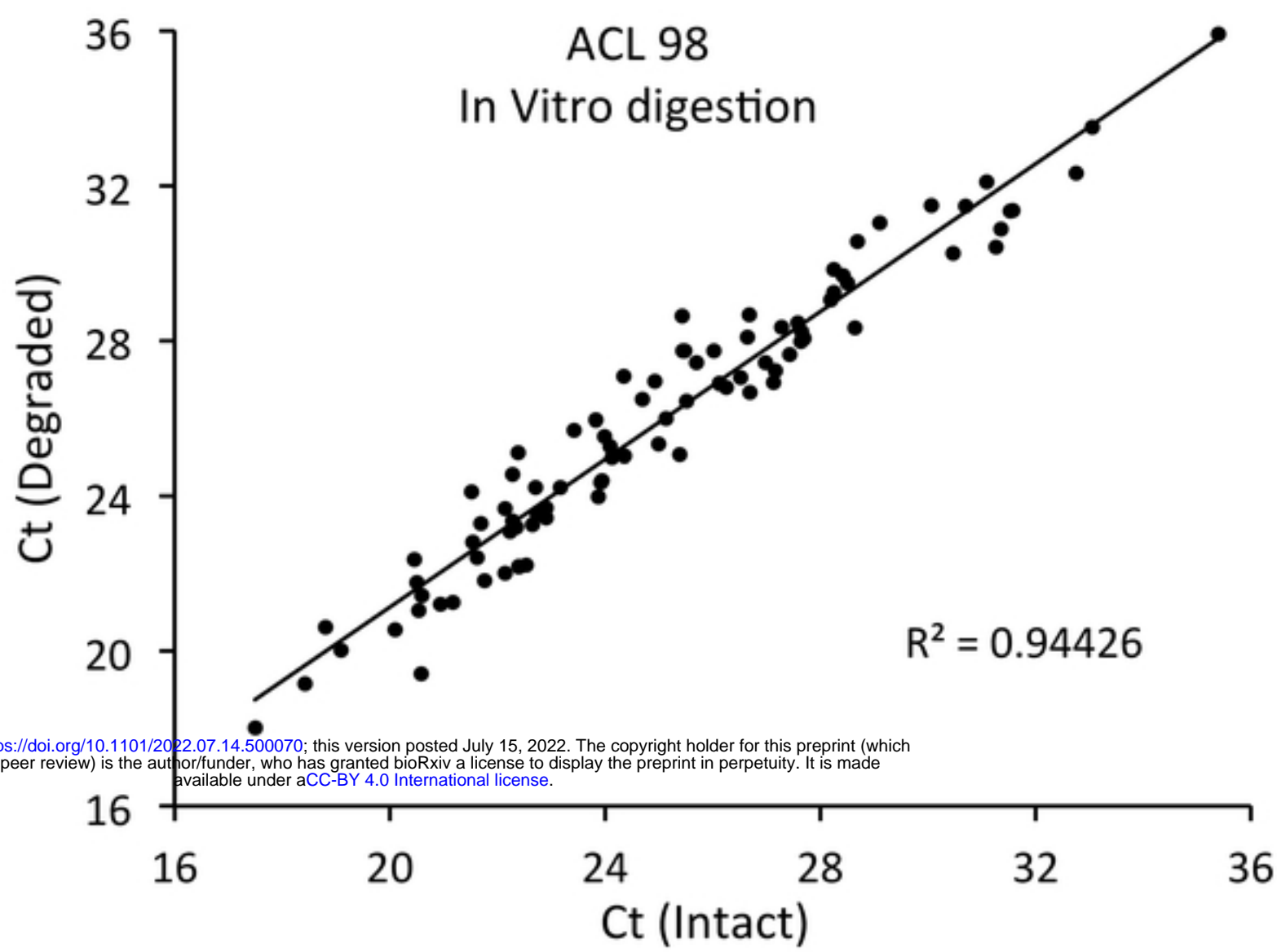


bioRxiv preprint doi: <https://doi.org/10.1101/2022.07.14.500070>; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

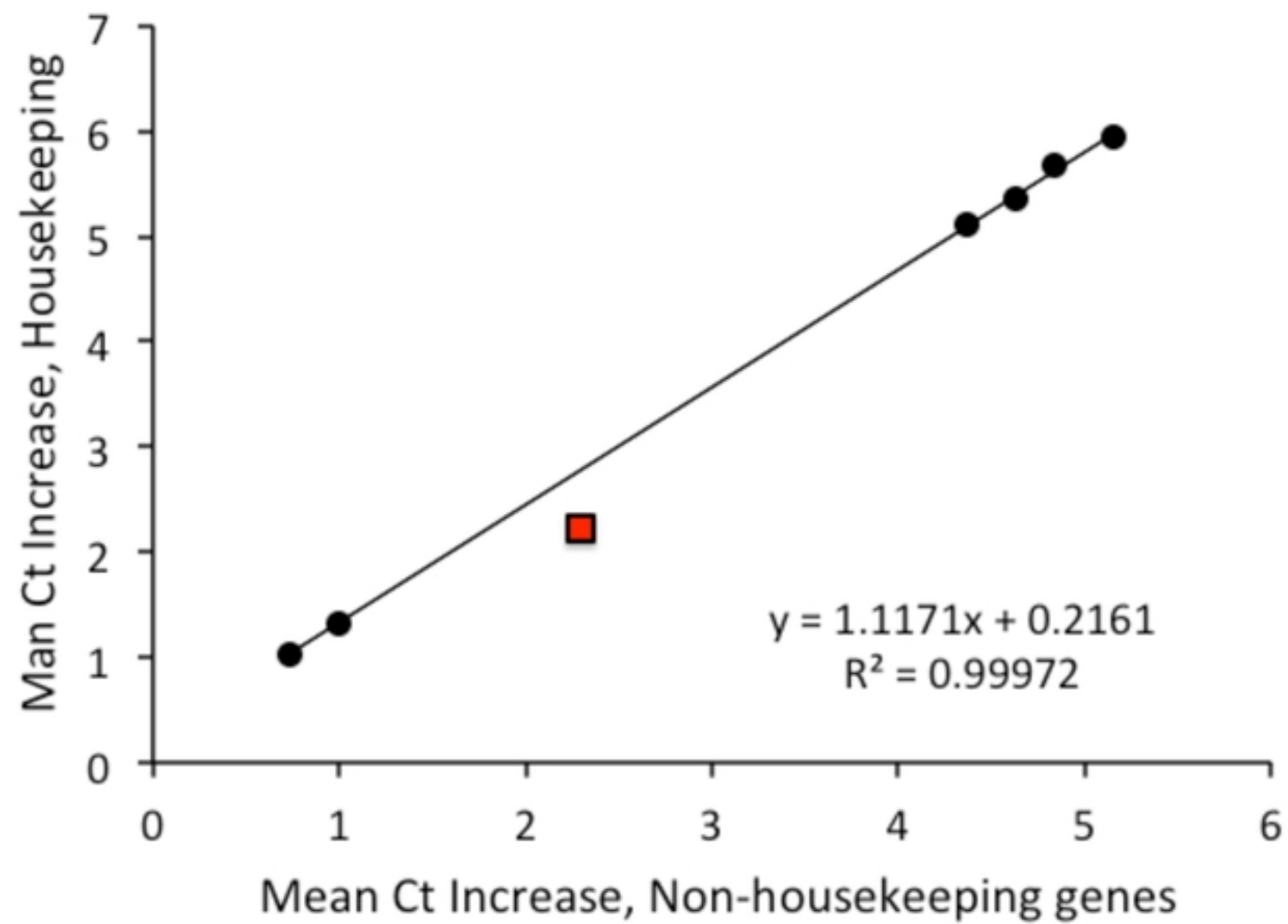


Figure

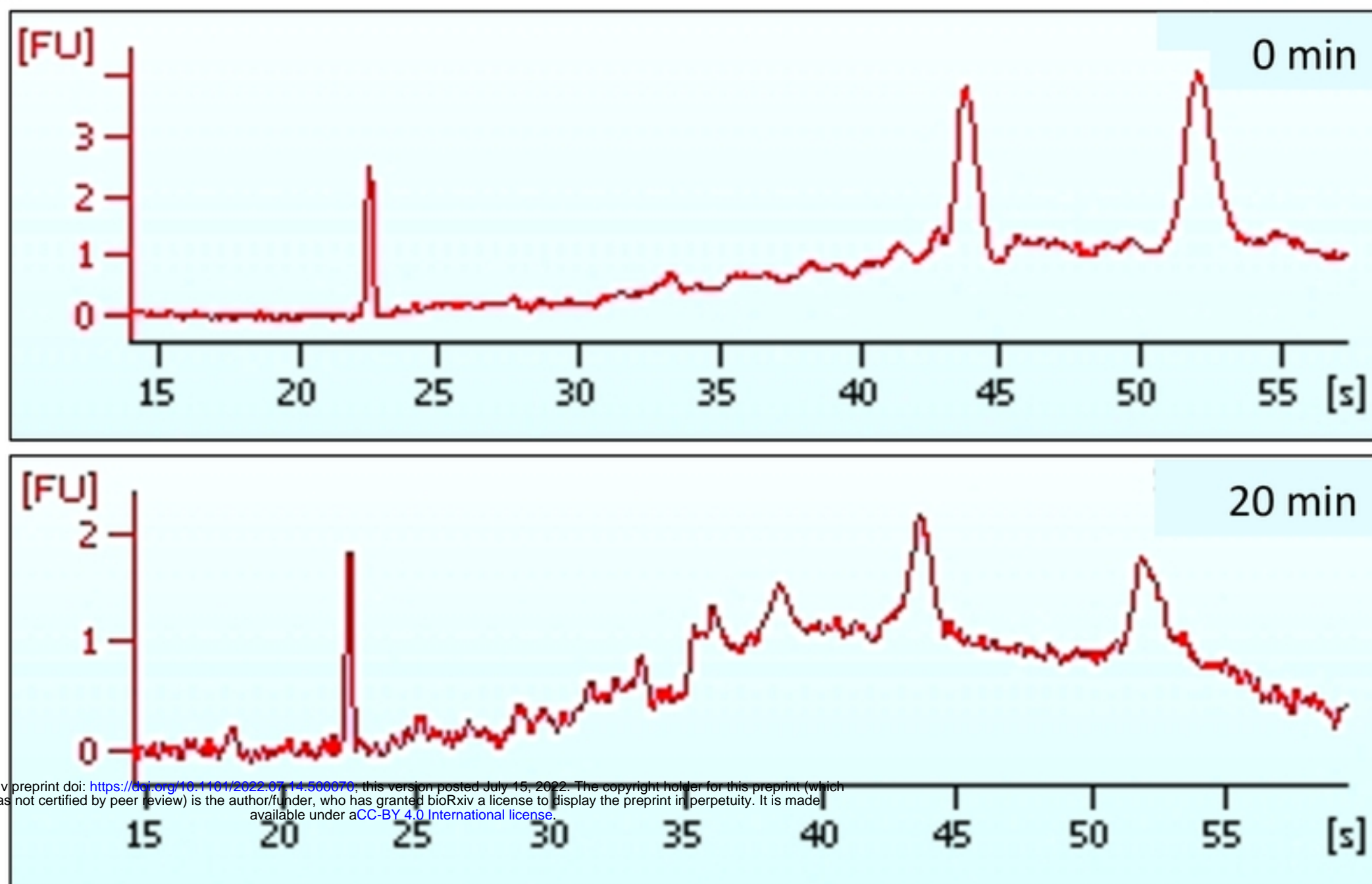
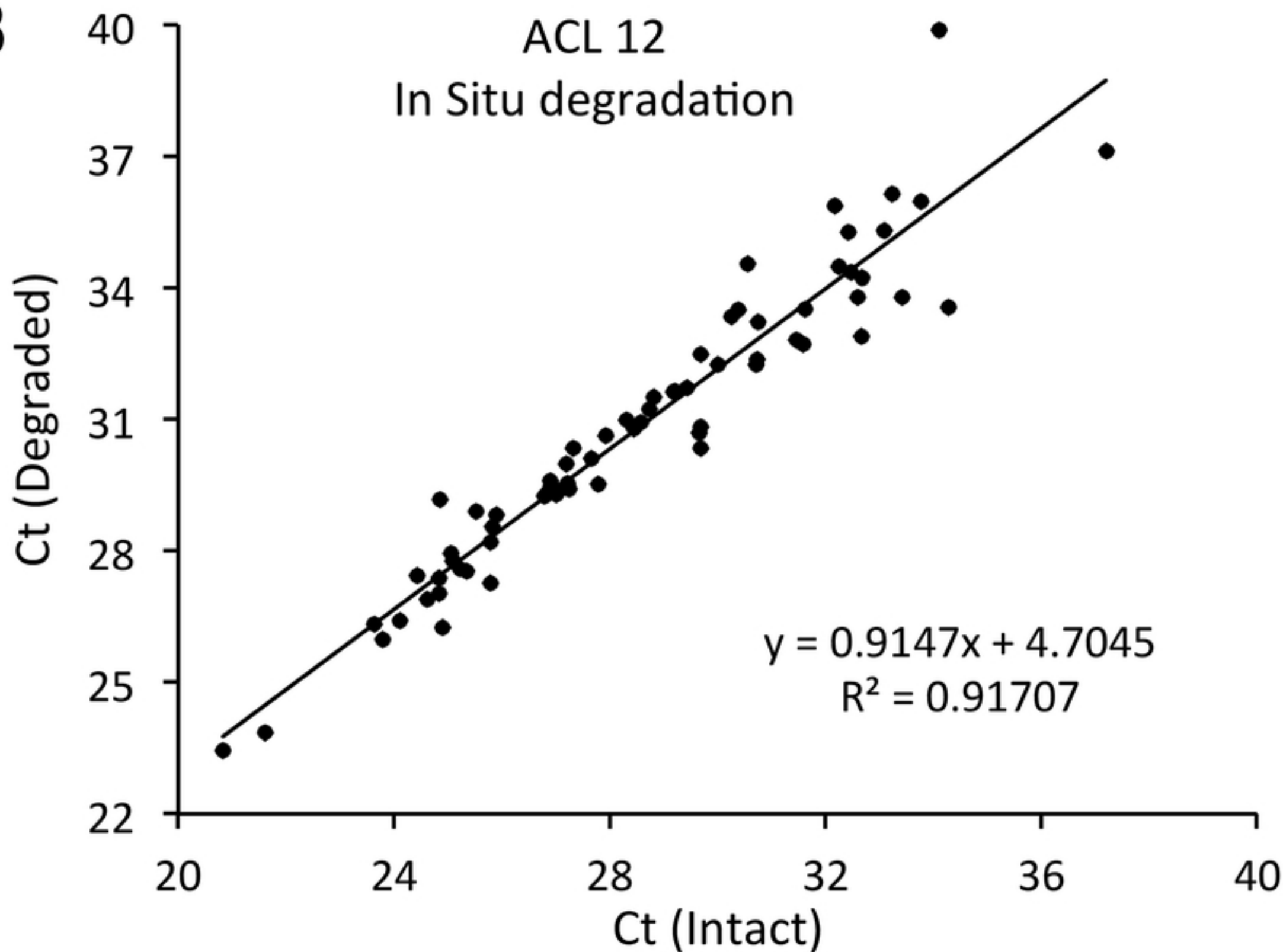
bioRxiv preprint doi: <https://doi.org/10.1101/2022.07.14.500070>; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Figure



Figure

A**B**

Figure