Dissertation

This manuscript ([permalink](#)) was automatically generated from [scherzerthesis/thesis@6518211](#) on February 28, 2022.

Authors

- **Michael**
  ⓘ [XXXX-XXXX-XXXX-XXXX](#) · ⬡ [johndoe](#) · 🐦 [johndoe](#)
  Department of Something, University of Whatever · Funded by Grant XXXXXXXX

Chapter One

Non-small cell lung cancer

Lung cancer is a global menace on human health. The historic consumption of cigarette smoke has caused lung cancer to be the most common and deadly form of cancer in both men women. Fortunately, smoking-rates have decreased in the past 30 years with lung cancer rates also on the decline. However, even if cigarette consumption were eradicated, lung cancer would still significantly affect human health. It is estimated that approximately 25% of all lung cancer cases arise in never smokers. Secondary causes of lung cancer include Radon exposure, high-heat cooking, family history, air pollution. An individual's risk of lung cancer also increases with age due to a decreased ability of the immune system to combat early neoplasms.

Lung cancer can be classified broadly into two major histological types that were named based on how the malignant cells look under a microscope: Non-small Cell Lung Cancer (NSCLC) and Small Cell Lung Cancer (SCLC). NSCLC is the most common type and represents 75% of lung cancer cases and is the most common subtype for patients who have never smoked cigarettes. Within NSCLC, tumors are further characterized based on histological architecture, anatomical location, cell-of-origin, as well as genomic alteration. For example, lung adenocarcinoma is the most common subtype and is characterized by glandular and papillary structures while squamous cell lung carcinoma characterized by keratin-pearls.Moreover, adenocarcinomas arise from alveolar-type-2 (AT2) cells in the distal lung and alveoli and are typified by genomic alterations throughout the MAPK pathway, such as EGFR, KRAS, BRAF, PI3K. Squamous cells carcinoma likely arise from basal cells and typically harbor SOX2 gene amplification.(Trudy cite).

The 5-year survival rate for patients with lung cancer was 25% in 2020 but varies depending on stage of disease at time of diagnosis. Although patient survival has improved over the past few decades due to improved targeted and immune therapies, the still poor prognosis reflects a need to better understand the molecular mechanisms underlying lung tumor progression, maintenance and response to targeted- or immune-therapy (reword).

Currently, the standard of care for LUAD are combination of several conventional chemotherapeutic agents, such as Cisplatin, Carboplatin, Paclitaxel, or Pemetrexed. Fortunately for patients who have been identified to have EGFR, ALK, ROS, TRK, or BRAF (V600E) genetic alternations significantly benefit from pathway-targeted therapies. (cite)

Of the solid malignancies, lung cancer accounts for the greatest number of new cancer diagnoses worldwide with an estimated 2.1 million new cases in 2018 (Siegel, Miller, & Jemal, 2019). It is the most commonly diagnosed cancer in men, and second most commonly diagnosed cancer in women after breast cancer. Unfortunately, this particularly insidious disease is the leading cause of cancer related death in men and often the leading cause of death in women each year, accounting for 1.8 million deaths last year. In the United States, these grim statistics have shown little improvement in the last decade. Kentucky bears the highest lung cancer incidence rate and this outcome is correlated with its top ranking smoking rate among the population. Lung cancer is a disease of aging and commonly affects people over the age of 60. Data provided by the National Cancer Institute and the Surveillance, Epidemiology, and End Results program indicate that patients in which their disease is detected early, while the tumor is confined to the primary location and has not disseminated to the adjacent lymph nodes have a five-year survival rate of approximately 57.4% (Duggan, Anderson, Altekruse, 2 Penberthy, & Sherman, 2016). Once the cancer has spread to adjacent regional lymph nodes, this survival rate drops markedly to 30.8%. Tragically, those diagnosed with late stage and distant metastatic disease have only a 5.2% chance of survival past five years. It is a grim fact that a majority of new patients that will be diagnosed with lung cancer, around 80%, will be informed that they have

regional or distant disease. Metastatic disease is the true killer in lung cancer. ~~Non-small cell lung cancer is also subdivided into several histological subtypes including adenocarcinoma, a disease arising from the distal glandular epithelial lining of the bronchi, squamous cell carcinoma, arising from the more proximal squamous epithelium of the bronchus, and large cell carcinoma (Cooper, Lam, O'Toole, & Minna, 2013). Notably, the mean doubling time of lung adenocarcinoma and squamous cell carcinoma is around 160 days (Arai et al., 1994). This means that from the time of oncogenic transformation that results in a tumor cell able to continuously proliferate and evade immune detection to the time of a detectable tumor mass by chest radiograph can be over a decade. For the vast majority of this time, patients do not experience pain from the growing mass and do not experience other related symptoms specific to the 3 disease. Often the first symptoms experienced by the patient are associated with metastatic dissemination such as pain associated with tumor invasion into the chest wall, liver capsule, or bone structure, or a number of other paraneoplastic syndromes. This delays the critical time to diagnosis for the patient, leading to poor outcomes, but also allows ample time for mutated cells to adapt and modulate a local tumor microenvironment that is conducive to support tumor growth, progression, and metastasis.~~

BRAF (V600E) signaling and cooperating genetic events to LUAD initiation and progression.

Due to the strong association between lung adenocarcinoma formation and mutational activation of the MAPK-pathway, many labs have generated genetically-engineered mouse models (GEMMs) that harbor conditionally-activated alleles of either KRAS (G12D) or BRAF(V600E). Expression on BRAF (V600E) in the Surfactant-Protein C expression cells (alveolar type 2 pneumocytes) elicits benign tumor formation that fail to progress to malignant lung adenocarcinoma. It is hypothesized that the cell cycle arrest is dependent on the tumor suppressors P53, and CDKN2A, as loss of either P53 or CDKN2A allows tumors to progress to malignant and deadly lung adenocarcinoma. Also, activation of WNT signaling and PI3K signaling similarly allow BRAF (V600E) adenomas to progress to adenocarcinoma. These observations are in concordance to the "multiple-hit" hypothesis. This hypothesis states that cancers do not arise from single mutagenic events, but are the consequence of sequential assaults on the genome that activate more than one oncogenic pathway. These multiple genetic hits allow cells to adopt traits to form cancer, which is coined by Bob Weinberg and Douglas Hanahan as the "Hallmarks of Cancer". The traits include suppression of apoptosis, active proliferation, altered metabolism, changes in cell identity, and forming blood-vessels (latest Hanahan and Weinberg). BRAF(V600E) expression alone is sufficient to activate many of the hallmarks of cancer, but additional pathways need to be activated to form deadly metastasis.

Great effort has elucidated the proto-oncogenes that drive cancer formation when mutated, such as *MYC*, *RAS*, *PI3K*, and others. Yet we still do not fully understand the mechanistic details that determine how genes cooperate to drive cancer.

*MAPK **Signaling**

Normal activation of the MAPK pathways begins with extracellular ligands binding to Receptor-Tyrosine Kinases (RTKs). This leads to the association of the G-protein coupled RAS with GTP, which activates RAS. GTP-loaded RAS recruits and activates RAF kinases (ARAF, BRAF, CRAF) which then phosphorylates and activates MEK kinases which then phosphorylates and activates ERK kinases. This ultimately leads to the activation of a suite of cellular processes that are required for proliferation, differentiation, and cell survival. Therefore, activating mutation in this pathway, from RTKs to kinases, serve to keep the pathway in a constitutively active state. This active state drives many of the hallmarks of cancer that is required for tumor formation. It is important to note that oncogenic activation of the MAPK pathway differs from normal ligand-mediated pathway activation. For example, mitogens, such as EGF, only transiently activate the pathway even with constant ligand stimulation. This immediate and strong spike in pathway activation is integrated by various biological processes in the cell that results in transcription of certain target genes. In contrast, oncogenic signaling thru

mutational activation of BRAF(V600E) for example results in constant kinase signaling that is integrated differently than normal pathway stimulation. Thus, an oncogene-specific transcriptional program is activated.

The importance of this pathway is reflected in the strong anti-tumor responses that are seen when small molecule inhibitors are used to target various molecules involved in the MAPK pathway. For example, Dabrafenib and Trametinib, inhibitors of BRAF and MEK kinases, respectfully, are used in patients that harbor BRAF(V600E) melanomas and lung cancers, both offering a clinical benefit (refs). Although there is strong pre-clinical and clinical evidence that BRAF+MEK blockade can lead to tumor regression, not all patients respond to treatment (ref). Therefore, to see a complete response, multiple pathways need to be targeted.

Genetically engineered mouse models (GEMMs)

Genetically-Engineered mouse models (GEMMs) of human cancer have allowed for the reproducible analysis of cancer biology driven by specific mutations that are often found in cancer patients. These models are useful for understanding the molecular mechanisms that drive tumor formation but also serve as a platform for testing tumor responses to targeted- or immune- therapies. Early mouse models utilized either the inherent tumor susceptibility of inbred strains of mice, such as FVB , or carcinogen-induced models such as UV, Urethane, or DMBA treatment. With the ability to genetically manipulate embryonic stem-cells (ES cells) came the development of mice with mutations engineered to spontaneously form neoplasms without the need for carcinogen treatments. Genetic knockouts or over-expression can lead to embryonic lethality depending upon the specific role of the manipulated gene, therefore conditional gene manipulation was developed to allow for temporal control thus overcoming the issues with embryonic lethality or other developmental defects.

The most common conditional system relies on chemically inducible transcription factors, such as tetracycline-dependent regulatory system to relies on the inoculation of tetracycline into the body to bind to a tetracycline-trans-activator to turn on gene expression. More recently, Cre-Lox recombination strategies have been utilized to get more cell type specific control. In this system, genes of interested are constructed to contain LoxP-sites that flank a particular genetic element and upon delivery of Cre-recombinase the DNA sequence in between LoxP sites are excised. With this approach, one can knockout a gene by engineering LoxP sites around critical exons, or one could turn on gene transcription by flanking a strong stop signal such as a polyadenylation sequence upstream of a gene-of-interest sequence such as KRAS G12D.

Genetically engineered mouse models of human cancer are important preclinical models, because they resemble the physiological environment of tumor growth in which tumors arise as progeny from a single initiating cell. These tumors can approximate the genetic alterations, transcriptional landscape, histology, and responsive disposition or lack therefore, seen in human cancers(cite). However, these model often take months to develop tumors and compound genetic alterations take time to develop. Later in this thesis, we will utilize TUBA-SEQ to both quantify tumor burden and cooperation of BRAF(V600E) with other common alterations in lung cancer.

Conditional mouse models of BRAF (V600E) lung cancer

Our lab has previously developed a conditional mouse model of BRAF (V600E) human lung cancers. In the *Braf(CAT)* model, normal BRAF is expressed from a conditional allele prior to Cre-mediated recombination . LoxP sites flank human *BRAF* cDNA encoding normal *BRAF* exons 15-18. Downstream of the LoxP sites is the mutant exon 15 that encodes the murine equivalent of the T1799A mutation that gives rise to the BRAF (V600E) onco-protein. Downstream of the mutant exon there is a P2A element and a CAAX-tagged *TdTomato*. Therefore, after Cre-mediated recombination, the BRAF (V600E) oncoprotein and fluorophore TdTomato is expressed at normal levels.

Talk here about the prevalence of mutations in the egfr-ras-raf-mek-erk pathway

Identification of this pathway as central to luad tumorigenesis

Development of gemms to study the disease (BRAF)

Proliferation arrest by this model

How does p53 loss lead to sustained wnt-signalling? Do dominant negative recapitulate?

RAS-RAF-MEK-ERK signaling

Talk about the details of raf activation and subsequent regulation of important cellular processes

P53- mediated tumor suppression

In biology, there are highly conserved pathways that control a bewildering amount of cellular processes that are both dependent and independent of each other, such as NOTCH- or WNT-signaling. The same could be said about P53. At the time of writing this dissertation, there are approximately 105,000 manuscripts on PUBMED that mention P53. P53 is so well studied because of both its apparent role in tumor suppression, but also because it is the most frequently altered gene in human cancer.

P53 is a transcriptional factor that has a DNA-binding domain that specifically recognizes two decameric half-sites. P53 also contains two N-terminal transactivation domains and a C-terminal oligomerization domain that are critical for P53-target gene activation. Central to the ability of P53 to induce transcriptional activation is tetramerization. Thus, if one or more P53 proteins in the complex is compromised in transactivation or DNA binding, transcriptional activity is compromised. Therefore, if one copy of P53 is mutated, the organism's ability to suppress tumor formation is compromised.

The importance of P53 is best demonstrated in people, and mice, that have just one mutated copy of P53. In humans with the familial- inherited Li-Fraumeni syndrome, in which they have one or more mutations in p53 are almost certain to develop cancer early on and throughout their life. Similarly, mice lacking two function P53 genes are prone to leukemias and lymphomas and have a shortened lifespan due to cancer incidence.

The most well-known cellular functions of P53 include its ability to induce cell-cycle arrest in response to DNA damage. Many stress signals, including oncogene activation has been shown to stimulate a reversible or irreversible cell cycle arrest. However, the strength of P53 induced cell cycle arrest or apoptosis is likely cell type- and cellular stress- specific.

Once a cell encounters a stressor such as DNA damage or oncogene activation, P53 is stabilized and accumulates in the nucleus to activate a suite of target genes that can lead to a number of cellular phenotypes, such as cell-cycle arrest or apoptosis.

P53 was originally thought to be an oncogene as P53 is accumulated in many human tumors which is not common in normal tissues. Moreover, ectopic expression of a P53 cDNA was found to aide in the transformation of primary cells induced by RAS. However, early studies erroneously used mutated P53 instead of wild-type leading to the misclassification of P53 as an oncogene. We known know P53 suppresses cell growth and transformation.

Often times, P53 is mutated in the DNA-binding domain at sites commonly referred to as "hotspots" due to there extraordinary frequency. These hotspot mutations poison the ability of p53 to bind DNA

therefore blocking its function. Interestingly, since one mutated P53 can disrupt the entire P53 tetramer complex, a single mutation can exert dominant-negative effects by inhibiting the normal tumor suppressive functions of P53. Although one mutant P53 allele is enough to compromised transcriptional activity, there is still selective pressure to lose the other wild -type copy. Loss of hemizygosity implies there is still residual tumor suppression in the presence of a wild-type P53 allele.

Early on in P53 studies, there were reports of certain P53 mutations having a "gain-of-function" (GOF) effect. For example, Li-Fraumeni patients with certain missense mutations in P53 would develop tumors earlier than Li-Fraumeni patients with loss-of-function (LOF) mutations (Bougeard 2008). Furthermore, there were experimental cell biology studies that would express mutant-p53 in P53-null cells and demonstrate enhanced tumorigenic potential (arnie levine). Further evidence for GOF mutant-P53 in mice indicated that missense mutant P53 induced different cancer types and enhanced metastasis than LOF P53 (Olive 2004). Furthermore, many groups have shown mutant-p53 can alter signal transduction affecting chemoresistance and altering metabolism.

The cellular mechanisms that GOF P53 are involved in are also well-characterized but context-dependent. For example, in Pancreatic cancers driven by KRAS(G12D) Mutant P3 interacts with CREB to induce FOXA1 transcription which enhances Beta-Catenin signaling to augment liver metastasis (Lozano paper).

P53 in lung cancer

Chapter Two

Chapter Three

Chapter Four

This manuscript is a template (aka "rootstock") for [Manubot](#), a tool for writing scholarly manuscripts. Use this template as a starting point for your manuscript.

The rest of this document is a full list of formatting elements/features supported by Manubot. Compare the input (`.md` files in the `/content` directory) to the output you see below.

Basic formatting

Bold text

Semi-bold text

<div align="center">Centered text</div>

<div align="right">Right-aligned text</div>

Italic text

Combined *italics and **bold***

~~Strikethrough~~

1. Ordered list item
2. Ordered list item
   a. Sub-item
   b. Sub-item
      i. Sub-sub-item
3. Ordered list item
   a. Sub-item

- List item
- List item
- List item

subscript: H_2O is a liquid

superscript: 2^{10} is 1024.

[unicode superscripts](#)⁰¹²³⁴⁵⁶⁷⁸⁹

[unicode subscripts](#)₀₁₂₃₄₅₆₇₈₉

A long paragraph of text. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Putting each sentence on its own line has numerous benefits with regard to [editing](#) and [version control](#).

Line break without starting a new paragraph by putting
two spaces at end of line.

Document organization

Document section headings:

Heading 1

Heading 2

Heading 3

Heading 4

Heading 5

Heading 6

A heading centered on its own printed page

Horizontal rule:

`Heading 1`'s are recommended to be reserved for the title of the manuscript.

`Heading 2`'s are recommended for broad sections such as *Abstract*, *Methods*, *Conclusion*, etc.

`Heading 3`'s and `Heading 4`'s are recommended for sub-sections.

Links

Bare URL link: https://manubot.org

Long link with lots of words and stuff and junk and bleep and blah and stuff and other stuff and more stuff yeah

Link with text

Link with hover text

Link by reference

Citations

Citation by DOI (Himmelstein *et al.*, 2018).

Citation by PubMed Central ID (Beaulieu-Jones and Greene, 2017).

Citation by PubMed ID (Heaven, 2019).

Citation by Wikidata ID (S, 2018).

Citation by ISBN (Suber, 2012).

Citation by URL (Himmelstein *et al.*, 2020).

Citation by alias (Ching *et al.*, 2018).

Multiple citations can be put inside the same set of brackets (Himmelstein *et al.*, 2018; Ching *et al.*, 2018; Suber, 2012). Manubot plugins provide easier, more convenient visualization of and navigation between citations (Himmelstein *et al.*, 2019; Heaven, 2019; Beaulieu-Jones and Greene, 2017; Ching *et al.*, 2018).

Citation tags (i.e. aliases) can be defined in their own paragraphs using Markdown's reference link syntax:

Referencing figures, tables, equations

Figure 1

Quotes and code

> Quoted text

> Quoted block of text
>
> Two roads diverged in a wood, and I—
> I took the one less traveled by,
> And that has made all the difference.

Code `in the middle` of normal text, aka `inline code`.

Code block with Python syntax highlighting:

```python
from manubot.cite.doi import expand_short_doi

def test_expand_short_doi():
    doi = expand_short_doi("10/c3bp")
    # a string too long to fit within page:
    assert doi == "10.25313/2524-2695-2018-3-vliyanie-enhansera-copia-i-
        insulyatora-gypsy-na-sintez-ernk-modifikatsii-hromatina-i-
        svyazyvanie-insulyatornyh-belkov-vtransfetsirovannyh-geneticheskih-
        konstruktsiyah"
```

Code block with no syntax highlighting:

```
Exporting HTML manuscript
Exporting DOCX manuscript
Exporting PDF manuscript
```
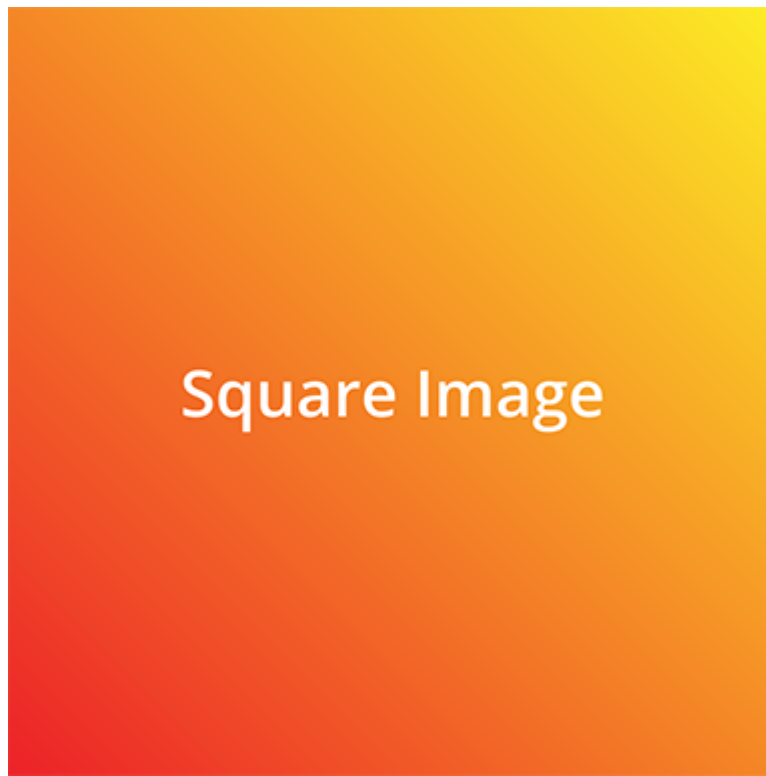
Figures

Figure 1: A square image at actual size and with a bottom caption. Loaded from the latest version of image on GitHub.
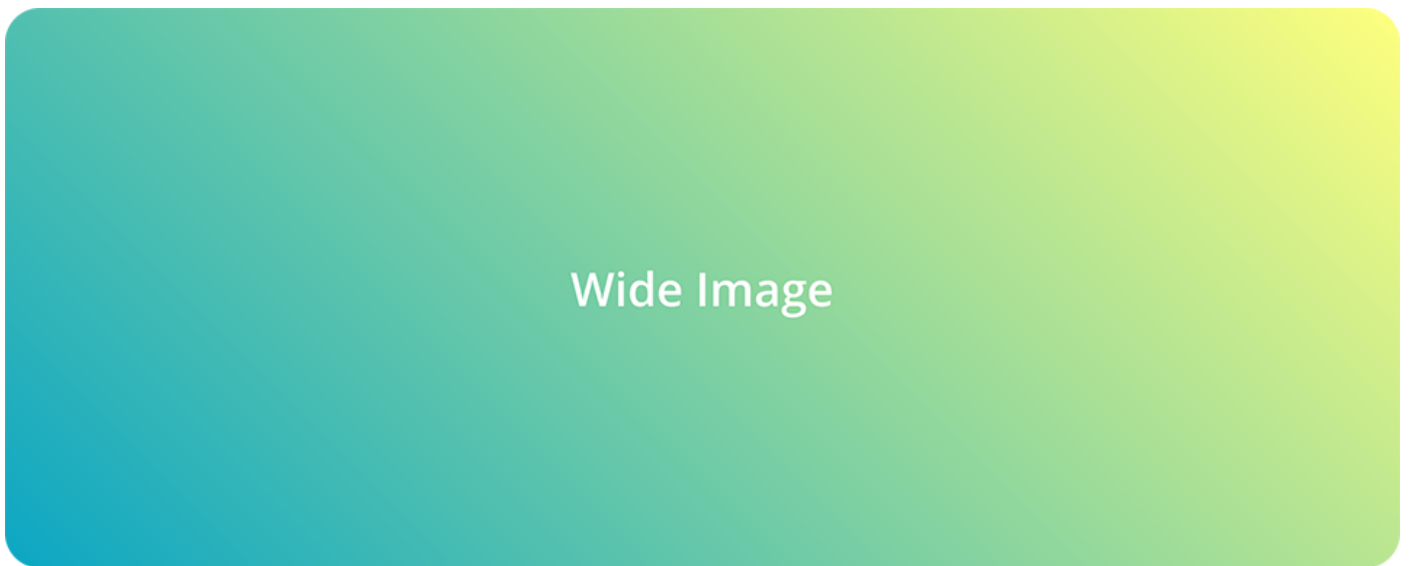

Figure 2: An image too wide to fit within page at full size. Loaded from a specific (hashed) version of the image on GitHub.
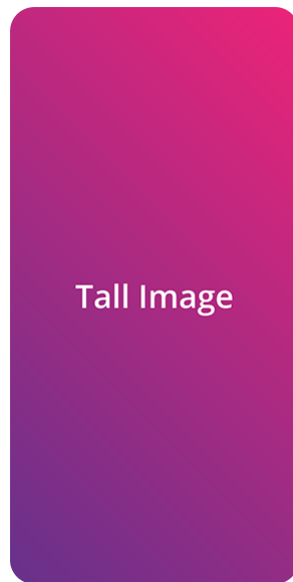
Figure 3: A tall image with a specified height. Loaded from a specific (hashed) version of the image on GitHub.

Figure 4: A vector `.svg` image loaded from GitHub. The parameter `sanitize=true` is necessary to properly load SVGs hosted via GitHub URLs. White background specified to serve as a backdrop for transparent sections of the image.

Tables

Table 1: A table with a top caption and specified relative column widths.

Bowling Scores	Jane	John	Alice	Bob
Game 1	150	187	210	105
Game 2	98	202	197	102
Game 3	123	180	238	134

Table 2: A table too wide to fit within page.

	Digits 1-33	Digits 34-66	Digits 67-99	Ref.
pi	3.14159265358979323 846264338327950	28841971693993751 0582097494459230	78164062862089986 2803482534211706	`piday.org`
e	2.71828182845904523 536028747135266	24977572470936999 5957496696762772	40766303535475945 7138217852516642	`nasa.gov`

Table 3: A table with merged cells using the `attributes` plugin.

	Colors	
Size	**Text Color**	**Background Color**
big	blue	orange
small	black	white

Equations

A LaTeX equation:

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \tag{1}$$

An equation too long to fit within page:

$$x = a + b + c + d + e + f + g + h + i + j + k + l + m + n + o + p + q + r + s + t$$
$$+ u + v + w + x + y + z + 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 \tag{2}$$

Special

⚠ **WARNING** *The following features are only supported and intended for* `.html` *and* `.pdf` *exports. Journals are not likely to support them, and they may not display correctly when converted to other formats such as* `.docx`*.*

[LINK STYLED AS A BUTTON]

Adding arbitrary HTML attributes to an element using Pandoc's attribute syntax:

Manubot Manubot Manubot Manubot Manubot. Manubot Manubot Manubot Manubot. Manubot Manubot Manubot. Manubot Manubot. Manubot.

Adding arbitrary HTML attributes to an element with the Manubot `attributes` plugin (more flexible than Pandoc's method in terms of which elements you can add attributes to):

Manubot Manubot Manubot Manubot Manubot. Manubot Manubot Manubot Manubot. Manubot Manubot Manubot. Manubot Manubot. Manubot.

Available background colors for text, images, code, banners, etc:

white `lightgrey` `grey` `darkgrey` `black` `lightred` `lightyellow` `lightgreen` `lightblue` `lightpurple` `red` `orange` `yellow` `green` `blue` `purple`

Using the Font Awesome icon set:

✔ ? ★ 🔔 ✖ ⋯

🏳 Light Grey Banner
useful for *general information* - manubot.org

🛈 Blue Banner
useful for *important information* - manubot.org

🚫 Light Red Banner
useful for *warnings* - manubot.org

References

Beaulieu-Jones,B.K. and Greene,C.S. (2017) Reproducibility of computational workflows is automated using continuous analysis. *Nat Biotechnol*, **35**, 342–346.

Ching,T. *et al.* (2018) Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface.*, **15**, 20170387.

Heaven,D. (2019) Bitcoin for the biological literature. *Nature*, **566**, 141–142.

Himmelstein,D.S. *et al.* (2019) Open collaborative writing with Manubot. *PLoS Comput Biol*, **15**, e1007128.

Himmelstein,D.S. *et al.* (2020) Open collaborative writing with Manubot Manubot.

Himmelstein,D.S. *et al.* (2018) Sci-Hub provides access to nearly all scholarly literature. *eLife*, **7**.

S,cOAlition. (2018) Plan S: Accelerating the transition to full and immediate Open Access to scientific publications.

Suber,P. (2012) Open access MIT Press, Cambridge, Mass.