

Logbook

class ~~2~~

Science Fair

Sept. 14

On Sept. 14 I:

- decided I wanted to be a part of science fair
- bought the logbook and binder

Sept. 18

On Sept. 18 I:

- got introduced to the science fair at Webber
- met my mentor
- decided on which type of project I would work on (study)

- Narrowed my topic down to a field:

Biochemistry

Sept 22

On Sept 22 I:

- Looked at interesting topics in biochemistry and found:

→ MicroRNAs in Cancer

→ CRISPR for studying the relationship between genes for allergies and the environment

Sept 26

On Sept 26 I:

- Decided officially what my topic would be:

→ microRNAs in Cancer

- Learned useful websites for research

→ NIH.gov

→ pub med

→ Sci.hub.Cite-DOI# from pub med

Oct. 4 On Oct. 4 I:

- Found around five papers about the topic and I printed one out -

Oct. 7

On Oct. 7 I:

Found two more papers that I can use in the future

- Decided on what question I am going to answer (Can microRNA be used to cure

Oct. 10

On Oct. 10 I:

- Looked at other CXSF projects to see what other people did

- Jessica thought my question:

→ Does microRNA have any role in cancer

and how microRNA be used in the

future to cure cancer

was good

then I had a idea

Oct. 12

On Oct. 12 I:

- Printed my first paper

by the NIH on microRNA in cancer

Oct 17

On Oct 17 I:

→ Read and annotated "Micro RNAs in Cancer" by the NIH.

Important Info in "Micro RNAs in Cancer"

① Biogenesis of miRNA

1. RNA Polymerase II (RNA pol II) transcribes a non coding gene (long) into 'pri-miRNA', also long.

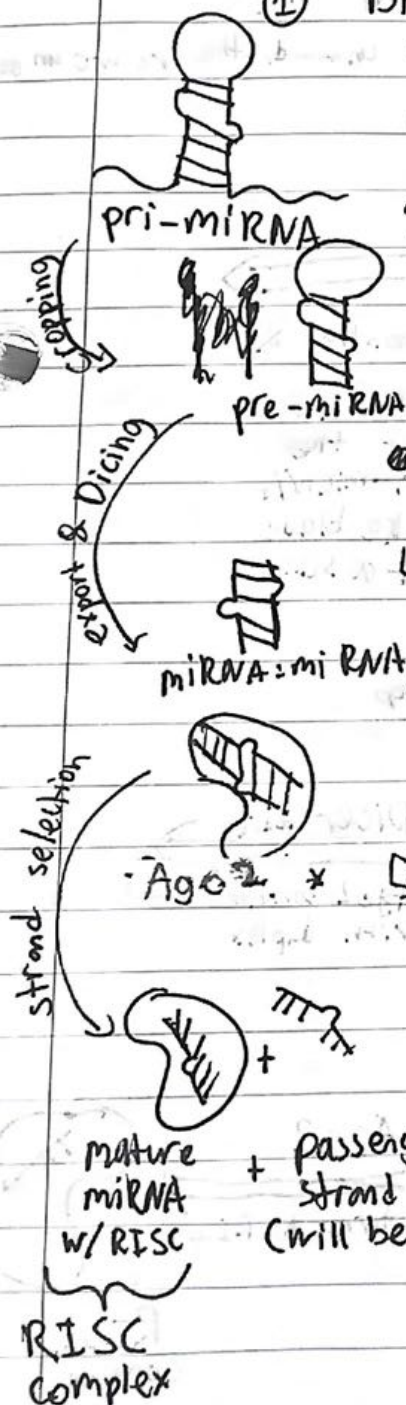
2. The Drosha Complex crops the long pri-miRNA into pre-miRNA (hair pin shape), DGCR8 binds to the pri-miRNA first to call Drosha.

3. Exportin-5 brings pre-miRNA out of the nucleus.

4. The Dicer Complex further processes the pre-miRNA to form double stranded miRNA:miRNA duplex. Ago2 attaches to the duplex and unwinds the duplex.

5. The miRNA:miRNA "breaks" and forms mature miRNA. (thanks to Ago 2)

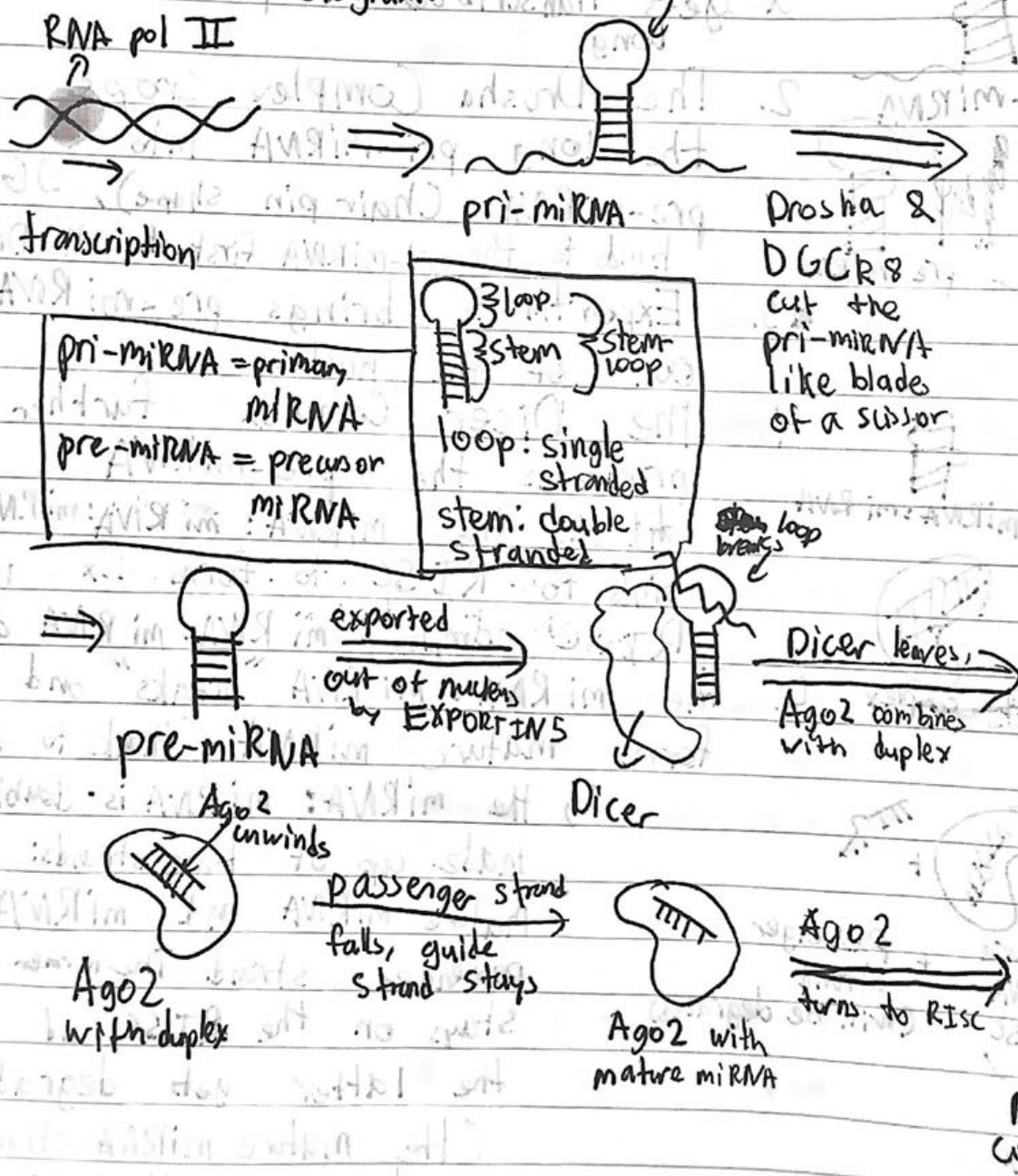
↳ the miRNA:miRNA is double stranded & made up of two strands: mature miRNA and miRNA passenger strand. The former stays on the Ago2 and the latter gets degraded (the mature miRNA strand is also called the guide strand)



6. By now, we have Ago 2 attached to the guide strand. Ago 2, the guide strand and some proteins combine to form RISC

↳ RNA
Induced
Silencing
Complex

* In this diagram I show the early stages of the miRNA (pri-, pre-, duplex) as unwound, they are unwound. A diagram!



② Overview on how miRNA stops mRNA

1. miRNA have flexible base pairing
1 miRNA can bond to tens to hundreds of mRNAs

2. Once bound either:

→ RISC cuts the mRNA which will be degraded by the cell

→ stay there to prevent RNA polymerase from binding

Binds in 3'UTR (untranslated region)

③ Important miRNAs for my purpose

miRNAs can be oncogenes or tumor suppressors

1. **miR-21 (Oncogene)**

→ Frequently found to be over-expressed in cancer

→ targets multiple tumor suppressors

→ increased expression of miR-21

results in increased proliferation

(growth of cells by multiplying)

and/or decreased apoptosis

2. **mi-155 (Oncogene)**

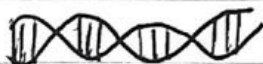
→ Similar to miR21, an oncogene

→ highly expressed in tumors, promotes oncogenesis

→ Is coded for by the gene BIC,

B-cell Integration cluster. (called B-cell ... since discovered in B-cells)

BIC



miR155

miR155

Regular BIC gene with healthy amount of miR155

with virus that causes lymphomas (lymph node cancer)



viral DNA

miR155

miR155

miR155

miR155

too much miR155 leading to cancer

→ miR-155 targets FOXO3B,
part of a family of proteins called
FOXO, which controls cell cycle
regulation (tumor suppressor)

→ Also targets the mismatch DNA
repair (MMR) system. This
corrects mistakes in base pairing.
miR-155 breaks this leading
to higher mutation rates called
MSI (microsatellite instability)

→ high expression of miR-155
causes increased resistance to
chemotherapy.

↳ Other oncogenes: miR-17-92 cluster,
miR-372 / miR-273 and miR-10b

↳ Some studies show miRNAs promote
homeostasis

3. let-7 (tumor suppressor)

→ Frequently reduced expression
in cancers

→ in a study let-7 was given to lung cancer
and that reduced the growth

4. miR-43a

→ made more when p53 binds to
the gene that makes it (p53 is
an enhancer)

→ repressed in pancreatic carcinoma
(cancer in pancreas)

→ controls cell cycle, DNA repair, apoptosis

5. miR-143, miR-145

→ both are repressed in colorectal cancer (end of intestine cancer) and prostate cancer

→ adding these back in reduces size of tumor

6. Some are either oncogenes or tumor suppressors based on context. Ex. miR-29a

→ acts as an oncogene or tumor suppressor based on context

End of Work Done on Oct. 17

pg. 4/14, 4 pages done.

Oct. 18

On Oct. 18 I:

→ Continued annotating the paper

"MicroRNAs in Cancer"

Sizes of: (nucleotides, nt)

→ pre-miRNA: 60-100nt

→ duplex & mature miRNA: 22 (approx) nt

→ pri-miRNA: thousands of nucleotides

Additional notes on

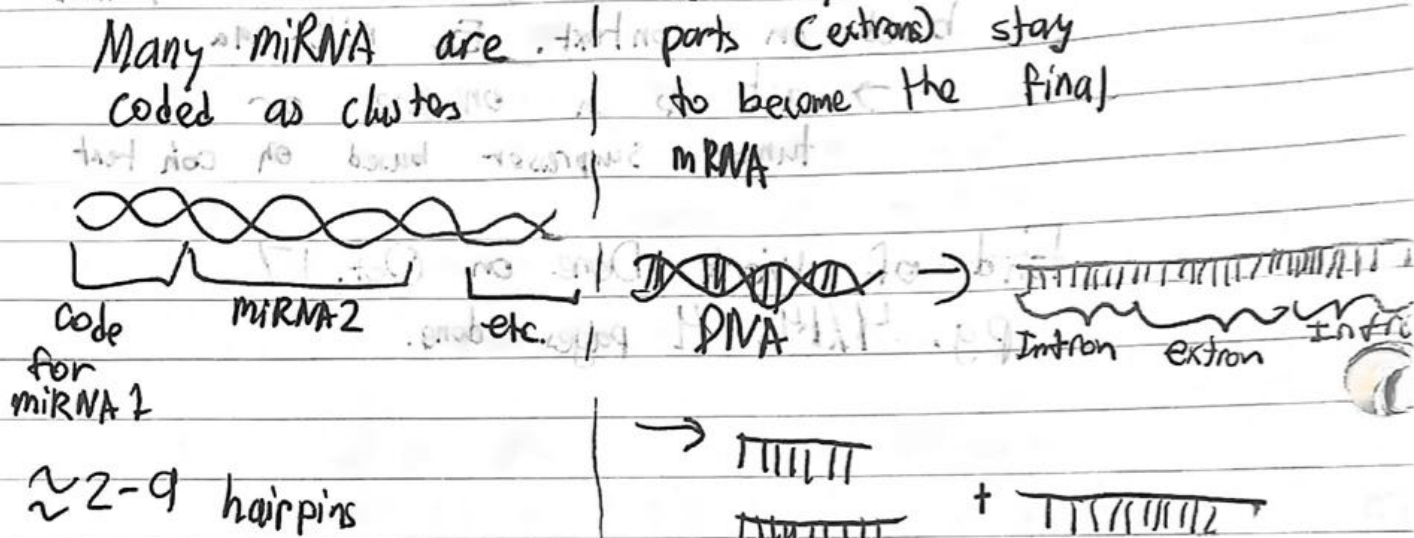
Synthesis of miRNAs

→ Along with Exportin 5, Ran-GTP also exports miRNAs out of the nucleus

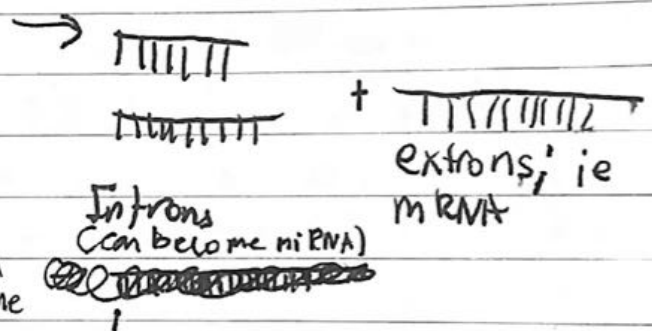
Less is known about how the synthesis of miRNAs is regulated, but the deregulation of these systems cause cancer.

50% of miRNA: non protein coding RNA
 40% of miRNA: derived from introns

When DNA is translated/transcripted the RNA produced as waste parts in it (introns). These are broken away from the RNA, the coding parts (exons) stay to become the final mRNA



Clustered miRNAs can be transcribed at the same time genes'



Many miRNAs are located where mutations in the DNA occur that cause cancer

④ Control and regulation of miRNAs

→ As seen in ③, miRNAs have a good part in cancer

→ according to the paper, "... it is plausible that restoration of normal miRNA expression in tumor cells could have therapeutic potential

RNA Pol II
makes most miRNA,
but RNA Pol III
does some too

RNA Pol II
adds a 7-methyl
guanylate cap at 5'
and poly (A) tail 3'

↓
bunch of
adenines

→ miRNA gene promoters
are similar to mRNA
gene promoters

→ The DNA binding factors
that control miRNA transcription
overlap with those that
control mRNA transcription

→ The important ones
are by c-myc & p53

→ c-myc

a) a proto-oncogene (supports oncogenes
and nullifies tumor suppressors)

b) controls 10-15% of human genes

c) controls cell growth & apoptosis

by activating or repressing
transcription of mRNA

d) also controls miRNA transcription

by binding to E-boxes

e) c-myc controls the activation
of the miR-17-92 cluster which
is often highly expressed in tumors

since c-myc is highly expressed
in tumors

f) this suggests that c-myc
& miR-17-92 cluster both
work together to promote
cancer

↳ over expression of c-myc
and miR-17-92 is known
to give more aggressive
tumors

↳ the miR-17-92 contains 6 miRNAs

the cell cycle regulator E2F1
f) c-myc also promotes
E2F1 transcription slowing
how c-myc regulates
both mRNA & miRNA
to fine tune the

cell cycle progression
g) c-myc also represses these
tumor suppressors, including
miR-15a, -29, -34, & let-7 families

h) the ~~exp~~ When scientists put these
miRNAs in a tumor, it decreased
cell growth, this shows that a
down regulation of these miRNAs
is an important part of c-myc
cancer

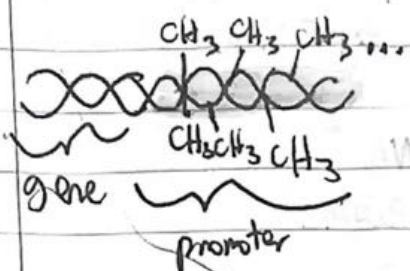
↳ c-myc regulates this through
direct binding to promoters
on miRNA

↳ c-myc increases transcription
of Lin-28B, this
represses the let-7 members

⑤ How We Control miRNAs

→ miRNAs are transcribed mainly
by RNA Pol II, so the
methods used to control mRNA
transcribed by RNA Pol II also
work well with miRNA

→ many tumor suppressors are
under regulated since they
are inaccessible since their
promoter doesn't activate them



a) This happens since the promoter part of the gene is wrong.

It's cytosine's now have too many methyl groups making the promoter not work

↳ known as promoter hyper methylation

b) The miRNAs that get promoter hyper methylated

are miR-127, -9-1 cluster, -193a, -137, -342,

-203, -34 b/c, When these

tumor suppressors get hypermethylated

it can cause cancer

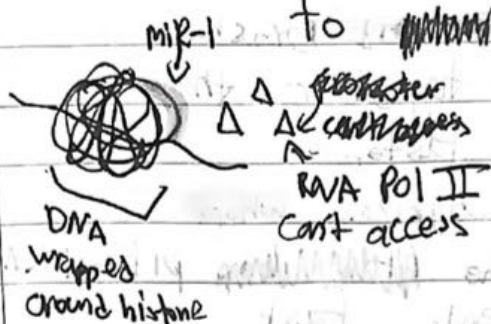
→ To fix this, HDACs are used

(histone deacetylase inhibitors)

These unravel the DNA around

the histones, making it more accessible

to ~~transcription~~ RNA Pol II



↳ This helps activate any miRNAs that have been silenced by histones

↳ ex. miR-1 gets silenced when the ~~transcription~~ RNA Pol II can't reach the gene since the gene is wrapped too tightly around a histone. Treating

it with HDAC will

unravel the gene, the ~~transcription~~ RNA Pol II will access the gene and miR-1 will be made

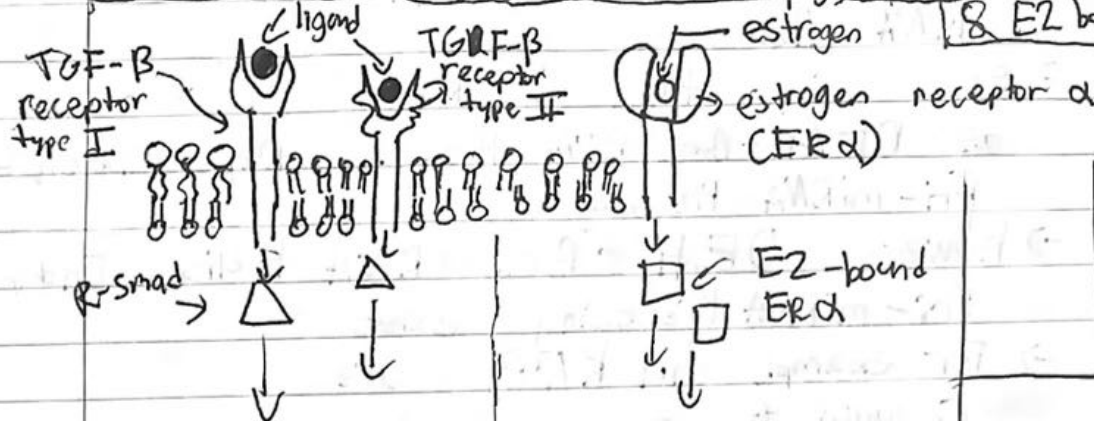


Key Point: disruption of miRNA biogenesis can alter the homeostasis of normal cellular biology

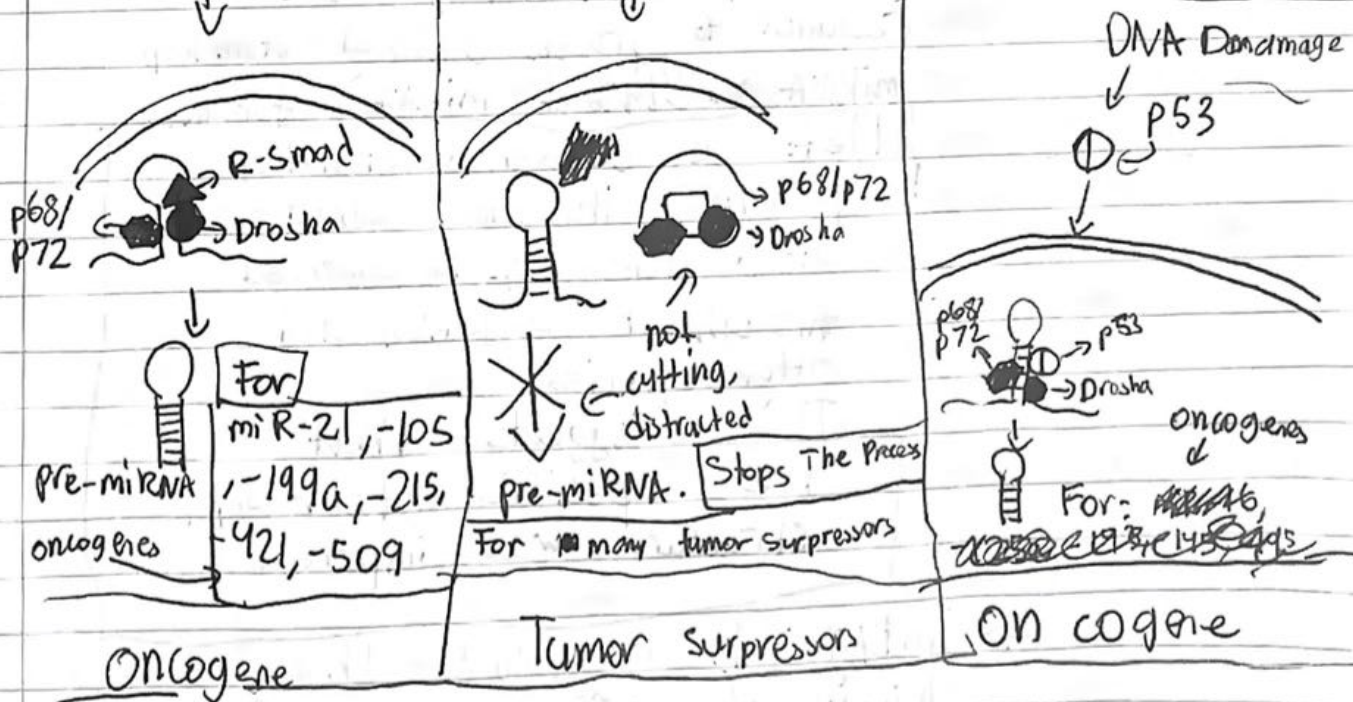
→ three important co-factors are

R-Smad (from TGF- β), p53 (from DNA Damage)

& E2 bound ER α (from estrogen)



p68/p72 () is a cofactor used to direct Drosha and changes the hairpin easier



End of work done on Oct. 18 & Oct. 19 pg. 8/15

Oct. 20

On Oct. 20 I:

→ Continued annotating the paper

↳ The goal was 12/15 pages done

Cont. Annotating:

→ The methods above use p. 68/p. 72

These are known as DEAD-Box

RNA Helicases

→ So, using these is known

as DEAD-Box RNA Helicases - Dependent +
Pri-miRNA Processing

→ however, DEAD-Box RNA Helicase-Independent
Pri-miRNA Processing exists

→ For example hnRMPs are
required to process conserved stem loop
miRNAs (~14% of miRNAs) for most

Def: a conserved stem-loop
is when the same structure
of a stem loop is conserved
throughout evolution and
other species

This suggests that
the specific stem-loop
structure was important

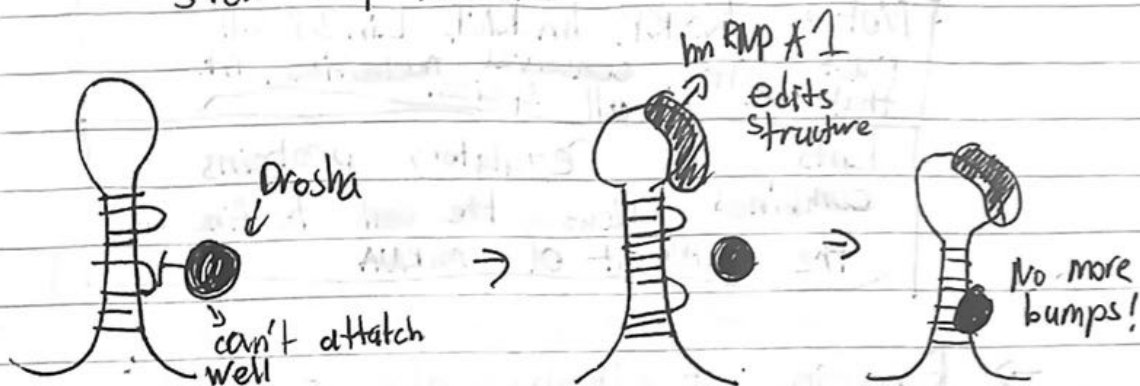
→ hnRMP A1 re-adjusts the
hairpin of miR-18a so Drosha
can cut pri-miRNA-18a easier

→ The stem-loop structure of miR-18a
is conserved, so the role of
hnRMP must be very important for evolution to not

→ 14% of all miRNAs have conserved stem-loop.
this suggests that hnRMPs might have a
very important role for these conserved

change
the

stem-loop miRNAs



→ ~~KSRP~~ KSRP interacts with the guanine-rich regions of the loop of a subset of miRNAs to promote both Drosha and Dicer ^{tumor suppressors}

↳ helps let-7a and miR-206

↳ without KSRP, the functions that let-7a & miR-206 have will be lost, causes increased proliferation (cells multiplying)

↳ KSRP is present in both Drosha complex & Dicer complex

↳ gene knockout of KSRP gene reduced cutting ability of Drosha & Dicer

↳ KSRP also controls the oncogene miR-155. When lipopolysaccharides are present, miR-155 levels increase thanks to KSRP, not more RNA Pol II binding to the promoter

→ Lin 28 negatively regulates Drosha.

↳ negatively regulates Drosha for let-7, unclear how,

either by blocking drosha or by altering the loop's shape so Drosha can't bind

Notes: KSRP, hnRNP, Lin28 all react with conserved nucleotides, nts that are in all species.

Lots of regulatory proteins combined allows the cell to fine tune the amount of miRNA

→ Repression of Drosha also occurs thanks to nuclear factor 45 (NF45) and NF90 proteins

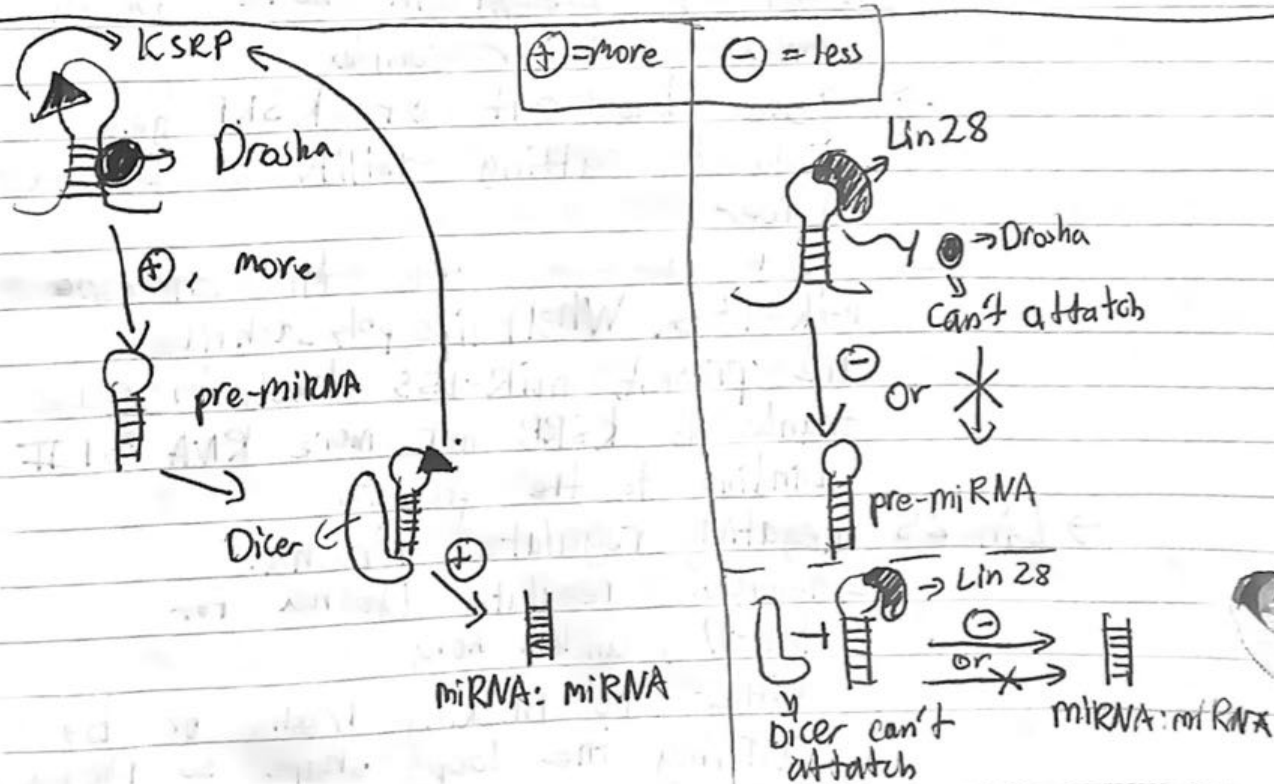
The NF45 / NF90 complex stops Drosha so pre-miRNA and mature miRNA isn't made

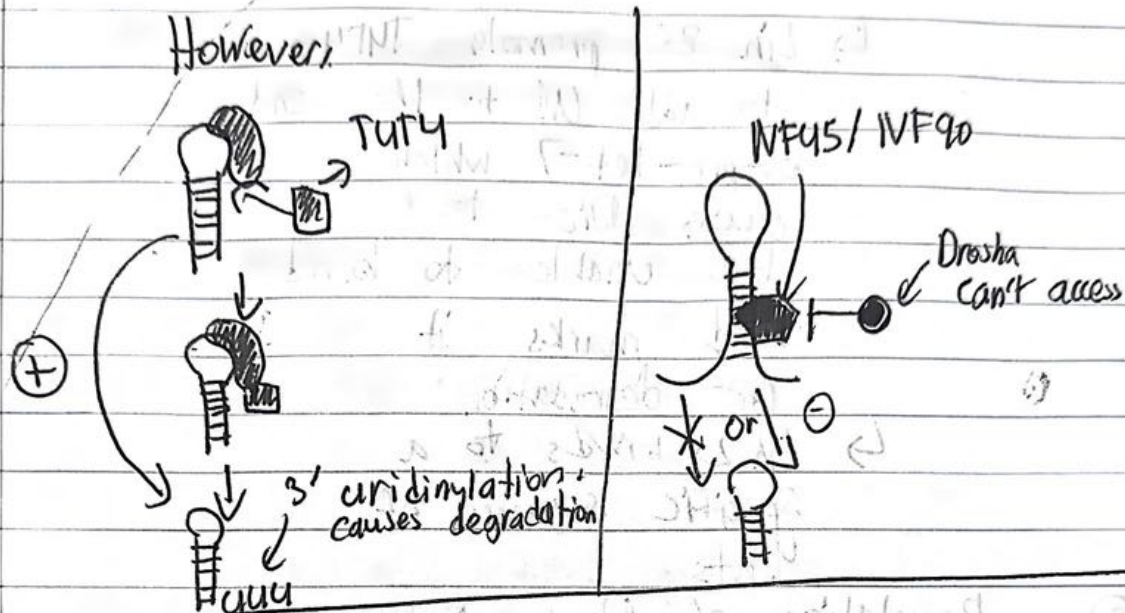
↳ Likes small miRNAs that are double stranded

↳ attaches before Drosha

↳ ~~also~~ Represses let-7

and also 3 other miRNAs





Oct. 21

On Oct. 21 I:

- continued annotating the paper

→ goal: $\frac{13}{15}$

→ Exportin 5 works with GTP bound form of Ran,

→ Dicer is an RNase type III protein

⑦ More on Dicer & RISC

→ Dicer present in nearly all eukaryotic cells (highly conserved)

→ has some helpers, TRBP & PACT

→ both help the stability of Dicer

And its speed

→ helps with Dicer is called the RLC (RISC Loading Complex)

→ Dicer is very good at its job,

doesn't need enhancers, only ⊖ regulators

a) Ex. Lin 28 (from DBP-A-independent

PrP-miRNA processing)

↳ cytoplasm is the primary location of Lin 28-let-7 interaction

- ↳ Lin 28 promotes TUT4 to add U6 to the end of pre-let-7 which causes Dicer to be unable to bind and marks it for degradation
- ↳ Lin 28 binds to a specific sequence of U nts

⑧ Regulation of Drosha & Dicer

- ↳ Increased Drosha's are observed in cancer
 - ↳ 2-7 fold increase but only miR-31 seemed to increase a lot
- DGCR8, when decreased minimally, also doesn't cause large decreases in miRNA
- This happens because the two are in a feedback loop
 - ↳ When less Drosha is present, Drosha cuts the mRNA that makes DGCR8 so that the levels can be balanced
 - ↳ When less DGCR8 is present, it causes less Drosha to bind on pri-miRNA triggering the reduce of Drosha
 - ↳ DGCR8 & Drosha need to be roughly equal for effectiveness

End of work done On
Oct 2nd
pg. $\frac{13}{15}$

Oct 22

On Oct. 22 I:

→ Tried to Finish annotating
the paper
goal: $\frac{15}{15}$

→ altered expression of Dicer occur
in cancer

a) Dicer is increased in prostate
tumors

b) Dicer is decreased in
non-small cell lung cancer

c) less Dicer → less
prognosis (likelihood of

normal life, symptoms becoming
normal, not dying)

d) Dicer levels also change
based on cell type and
stage of tumor

→ ex. precursor of
lung adenocarcinoma shows
high Dicer levels, mature /
advanced invasive tumor
showed less Dicer levels

→ each cell has two copies of the
Dicer gene. One from each parent.

Losing one means you are hemizygous
Lose both: homozygous

a) Being hemizygous causes more
cancer than being homozygous

b) The human cancer genome data support this: more tumors happen from hemizygous ^{not} homozygous (dramatic reduction)

c) Why? When hemizygous cell can still survive but can't make enough miRNAs so becomes cancerous. homozygous means the cell can make no Dicer so just dies. "known as haploinsufficiency")

→ TBRP (Also PACT) is also important for it

dsRNA = double stranded RNA
ssRNA = single stranded RNA

a) Less TBRP & PACT decreases stability of Dicer a lot

b) mutation in TBRP → less miRNA → less Dicer → Cancer

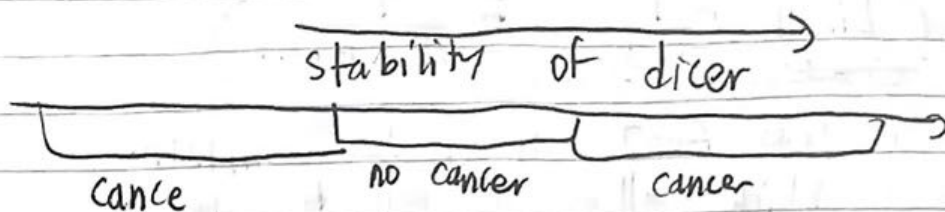
c) mutation here is a truncation mutation, shorter protein (from nonsense gene mutation, i.e. premature stop codon)

d) putting full-length TBRP back in fixes Dicer

e) cellular signaling pathways also affect TBRP levels

↳ ex. MAPK/ERK promotes phosphorylation of TBRP → more stable dicer

→ more growth miRNAs, less tumor
 Suppressor miRNAs → cancer



The MAPK/ERK pathway when ON
 inhibits apoptosis, promote proliferation.
 If it is always ON it can cause cancer

Ago2 regulation

↳ humans contain 8 ago proteins:
 Ago 1→4, Piwi 1→4
 ↳ Piwi are structurally similar to
 Agos but interact with piRNAs

→ Levels of Ago2 contribute to
 miRNA regulation

(Ago2 $\propto$ miRNA levels
 no Ago2, no miRNA
 more Ago2, more miRNA)

→ Ago proteins are to stabilize
 miRNA levels

→ Ago2 found to be increased in

breast cancer
 → EGF (Epidermal Growth Factor, lig and
 to —————) increases Ago2
 stability causing

more Ago2

→ Ago2 targeted for degradation
by Lin 41 (expressed in stem cells
& undifferentiated cells)

$$\left(\text{Lin 41} \propto \frac{1}{\# \text{ Ago2}} \propto \frac{1}{\# \text{ of miRNA}} \right)$$

a) less Ago2 means less miRNA
which allows the cell to stay
stem-like

b) Lin 41 is the target of let-7
 $\# \text{ Lin 41} \propto \frac{1}{\# \text{ of let-7}}$

Similar to:

$$\# \text{ Lin 28} \propto \frac{1}{\text{let-7}}$$

c) In lung cancer $\# \text{ of let-7}$
is less, so Lin-48 is
more so Ago2 is less
(might be a characteristic
of poorly differentiated tumors)

④ Mutations in miRNA processing

- ↳ single nucleotide polymorphisms (SNPs)
- ↳ point mutations

→ These mutations in pri-miRNA
can change its physical structure
and cause it to not be able
to bond to Drosha (mutations in pri- cause less pre-)

→ mutations in pre-miRNA reduce
mature miRNA levels

↳ ex. when one nt
changes from C → T
in pre-miR-16-1
causes less miR-16-1

→ SNPs also alter levels

↳ Ex. a common G→C mutation in pre-miR-146 causes more risk of papillary thyroid cancer. When the C is there, less mature miR-146 gets formed since it alters the 2nd structure (physical structure) of the pri-miR-146 and stops Drosha

↳ a mutation (polymorphism) of miR-125a makes the loop bigger blocking Drosha

→ Additionally, SNP of miRNAs can cause the miRNA not be able to bind to the target

↳ ex. SNP in let-7 makes it unable to bind to the 3'UTR region of K-Ras which increases the risk of non-small cell lung cancer.

→ Therefore, loss or gain of miRNA binding sites in 3'UTR can be an important determinant in cancer

→ Length of 3'UTR also plays a role

a) in cancer 3'UTR is smaller so it is harder for tumor suppressor miRNAs to bind

b) the precise 3' end of some mRNAs could distinguish different tumors

→ The combined effect of reduced expression of miRNAs (tumor suppressors) and shorter 3' UTR might contribute to development of cancer

⑩ Conclusion

→ miRNAs regulate cell differentiation, proliferation and survival

→ alteration of miRNA expression cause cancer

→ regulation of miRNAs is diverse

→ ~~More~~ an understanding on how miRNAs are regulated under normal conditions will allow us to better understand what gets de-regulated in cancer

End of work done on Oct, 22
pg 15/15

Done Annotating and writing notes on: "MiRNAs in Cancer"

Oct. 23

On Oct. 23 I:

- had 3rd Science Fair meeting

Homework:

- Find what do you want to find out

- No book report!

- Look at other

studies to say how they

Structure it

↳ Literature review?

↳ Analysis?

↳ see how much data
you want to incorporate

- Email Ms. Garcia on this?

From Oct. 24 - Nov. 9 I:

→ Decided where I was going to get
my data (GDC)

→ Learned how to use the GDC

→ Decided which specific miRNA
family I am going to look
at (let-7)

→ Decided which cancer region
I am going to look at
(Kidney)

→ Decided on which specific type
of Kidney cancer I am going
to look at:

Kidney Clear Cell Adenocarcinoma
Based on # of cases present
on type in GDC

→ Downloaded Data
1. Categorized by Stage or if
Normal!
it is
Stage 1 Stage 4
Stage 2 Normal
Stage 3

2. I grouped
all these folders
under the folder
"Stage" for simplicity
when searching.

3. Then, I went
into my terminal
and "cd" into
Stage

4. Then, I ran
"grep -r" which means
look in all files recursively
until you find

5. I ran this for all
the let-7 miRNAs
available and tabulated
the results

Let-7 available:

hsa-let-7a-1	→ the -1, -2... means isoform. Isoforms of miRNA code for basically the same thing. Very similar.
hsa-let-7b	
hsa-let-7c	
hsa-let-7d	
hsa-let-7e	
hsa-let-7f-1	
hsa-let-7g	
hsa-let-7i	

nsa means
homo sapiens

Cont. on Nov. 10:

Nov. 10

The reason only 1 is shown

for hsa-let-7a & -7f

is because the miRNAs

are expressed basically

the same in the cells.

So, creating a new graph and

table for them is pointless

→ Extract & Display Results

1. On each sheet there are:
tables.

The columns are:

UID	ID	File name	Read count	RPM

Explanation of each column:

UID: This is the ID of the data file

UID: This is the ID of the data file
UID: This is the ID of the data file
UID: This is the ID of the data file

File name: name of the file

read count: how many times

a specific miRNA sequence was
observed in the data

Note: Read count isn't very

accurate since different

samples have different

sequencing depths (reads)

RPM (Reads per million):

This way is much better

and more accurate since it takes
reads into account:

Calculated as:

$$\left(\frac{\text{read count}}{\text{total miRNA reads}} \right) \cdot 10^6$$

Note: Later, I added another column, file URL. This tells the URL for the data from which we got it. The form of the url is

"https://portal.gdc.cancer.gov/files/"

and then we have the URL.

This was done through the formula button:

$f_2 = \text{CONCAT}(\text{"https://portal.gdc.cancer.gov/files/"}, C_2)$

where C is the column for file URL and n is the row

An example for one is:

$f_2 = \text{CONCAT}(\text{"https://portal.gdc.gov/files/"}, C_4)$

Graph:

Using the "Stage" and RPM columns I created 2 graphs, one a line graph the other a scatter plot

In the graph:

x-axis: Stage

y-axis: RPM

Cross mapped: If any ~~gene~~ miRNA has
have similar sequences
to the target miRNA. If
so, it can make it hard
to get true values

Note: I've also added the means
of each stage in bold

This is the formula:

$$F_2 = \text{AVERAGE}(G_{n_1} : G_{n_2})$$

where G is the column

where the RPM is

and, row of

$n_1 = \text{start of stage}$

$n_2 = \text{row of end of stage}$

Also, In addition to the RPM graphs,
there is also the mean graph
which is a bar graph graphing the
Stage: x axis
RPM: y axis

Nov. 19 On this day I:

- Started by background research
- wrote on miRNA and Cancer

Nov. 23 On this day I:

- wrote on Kindey Clear Cell adenocarcinoma



Nov. 27 On Nov 29 I:

- Edited the document
- added more to the para. on miRNAs

Nov. 28 On this day I:

- Started reading paper

Nov. 30 On this day I:

- Read and made notes on pages 1-4 of "The Roles of the Let-7 Family of miRNAs in the Regulation of Cancer Stemness"

On Let-7 Family

#

Discover

① Let-7 gene was discovered in *C. elegans*,
a species of round worm (nematode)

② It was the second gene discovered
in that species and is highly conserved

③ Its expression determines the fate
of adult cells in the worm

④ highly conserved

Dec 1

On this day I:

→ Read on the paper and
summarized info on Background Info

Dec 3

On this day I:

→ Finished 3/5 of info on
paper in BI

Dec 4

On this day I:

→ Finished the entire BI
5/5

Dec 6

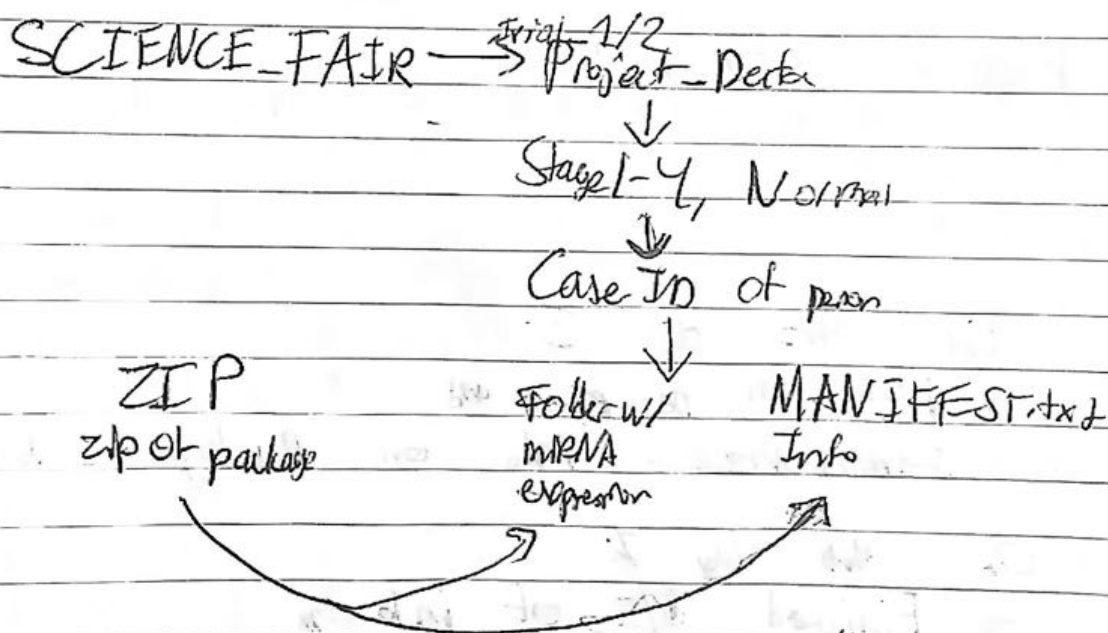
On this day I:

- Started downloading the files for Trial 2 of Project Data as the first trial seemed incorrect with huge outliers.
aka Stage 1-3

Dec 7

On this day I:

- Finished downloading stage 4 & Normal data



Dec 8

I graphed the same way I did trial 1 just that I am including isoforms

→ finished graphing hsa-let-7a-1

→

Dec. 10

Finished graphing hsa-let-7b-e

Dec. 12

On this day I:

- Decided to increase my sample size by downloading all the available files for cancer ~ 500 sample size
- Started downloading files from GDC

Dec. 14

On this day I:

- Finished downloading all GDC files and finished uploading files
- Spreadsheet w/ all data done

Dec. 15

On this day I:

- Started to code Python model demonstrating the ability and high accuracy of a system for let-7b detection.
- p-value: $1.18 \times 10^{-7} = 0.000000118$
- Since a good p-value for scientific analysis

Dec. 17

Continued coding model

Dec. 18

Finished coding model

Dec. 19

Fixed bug for class separation

from making a binary and using
the new sheet, Python-Data.xlsx

Dec 21

Made box plot for Results
added progress to Science Fair Info
so far

Dec 22

- Added more info for testing code
- Exr Screenshot of code, table displaying results, etc.

- Added Info from talk with other PhD, added more papers to study
 - Found paper saying B2-1/b should be down regulated in celiac, showing that there is a disagreement!

Jan 6

On the day I!

- Reviewed Code
- Reviewed SVMs and SMOTE

Jan 12

On the day I!

- Recorded Info on model
Crest pages

Accuracy: 90.470%
 Precision: 88.70%
 Recall: 93.58%
 ROC-AUC: 93.09%
 F1 score: Normal: 0.90; Cancer: 0.91
 # of False Positives: 13 (FP)
 # of False Negatives: 7 (FN)
 # of True Positives: 102 (TP)
 # of True Negatives: 93 (TN)

Definitions:

Accuracy %: Percent of guesses correct

$$\frac{TP + TN}{Total}$$
 (100% of the time model was correct)

Precision %: Percent of time model predicted cancer correctly

$$\frac{TP}{TP + FP}$$
 = $\frac{\text{correctly predicted } \oplus}{\text{total predicted } \oplus}$

Recall %: How many $\oplus$ did the model get out of all $\oplus$?

$$\frac{TP}{TP + FN}$$

F1 score: by getting the harmonic mean of precision and recall, it can display how well the model can identify cancer

$$F1 = 2 \cdot \frac{Precision \times Recall}{Precision + Recall}$$

Roc-AUC: A general score on how well it can classify, based on the effectiveness of its threshold

Jan 11

On this day I:

- Started writing analysis of results
 1. let 7b ~~III~~ ?
 2. let 7b = best biomarker
 3. ~~let 7b~~

- Decided to do a paper

Jan 17

On this day I:

- Finished Analysis (ough)
- Finished Abstract/Summary of Paper → JEI

Jan 19

On this day I:

- Started & Finished Introduction
- Refined Analysis
- Edited w/ Grammarly

Jan 20

On this day I:

- Finished Results of Paper
- Added Concluding para for Discussion
- Added Acknowledgements
- Added 1-6 of References w/ Quilbot
- Added Title

Jan 21

On this day I:

- Finished Rough draft of paper
 - ↳ Discussion & Materials & Methods ✓
- Added References 7-15 or 16
- Started reading (not annotating) paper by Wiley of let-7 in lung cancer.

Jan 31

On this day I:

- Reviewed BI
- Fixed error on paper

Feb 2

On this day I:

- reviewed analysis

Feb 4

OTDI:

- reviewed Discussions & Methodology

Feb 5

OTDI:

- added Bib to BI

ANA → Quant → Ivan Sa Info
11/6/1

Feb 8

OTDI:

- Checked over Results
- did meeting w/ professor

Key! Do Literature Review
for Research Gap

⇒ (Jordan) ⇒

Once done, meet him
also, do comments

Feb
10-12

OTDI:

- Used Grammarly (No AI)
to review my work
gradually

Feb 15

OTDI:

- Reviewed 2013
of his comments

Feb 17

OTDI:

- Started planning
trials

Canva?

Slide Show

dimensions?

Feb 18

OTDI

- submitted everything
to Tim & In. Chicago

Feb 23

OTDI:

Canva

color palette:

red, pink, teal, dark blue,
yellow?

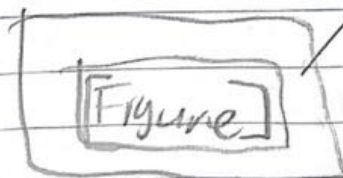
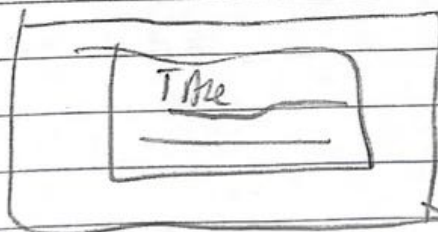
Feb 25:

- OTDI
- Got all graphs / tables
 - looking for icons



~~Mar 2~~ Mar 2 OTDI:

- Extracted & summarized the text



color

Mar 4

OTDI

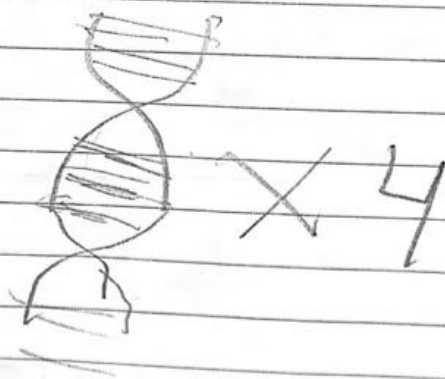
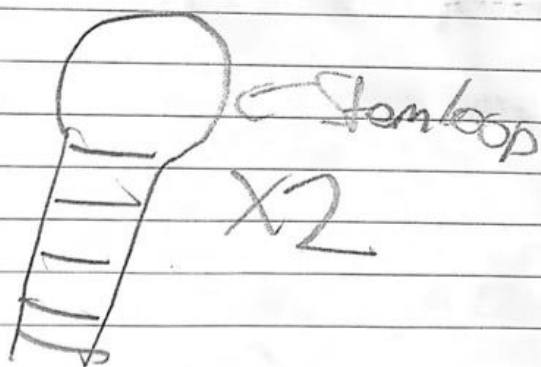
- did 1-8
for Lit review $\rightarrow$ JFI

- Reviewed
all summarized
text

Mar 6

OTDI

- found the teens



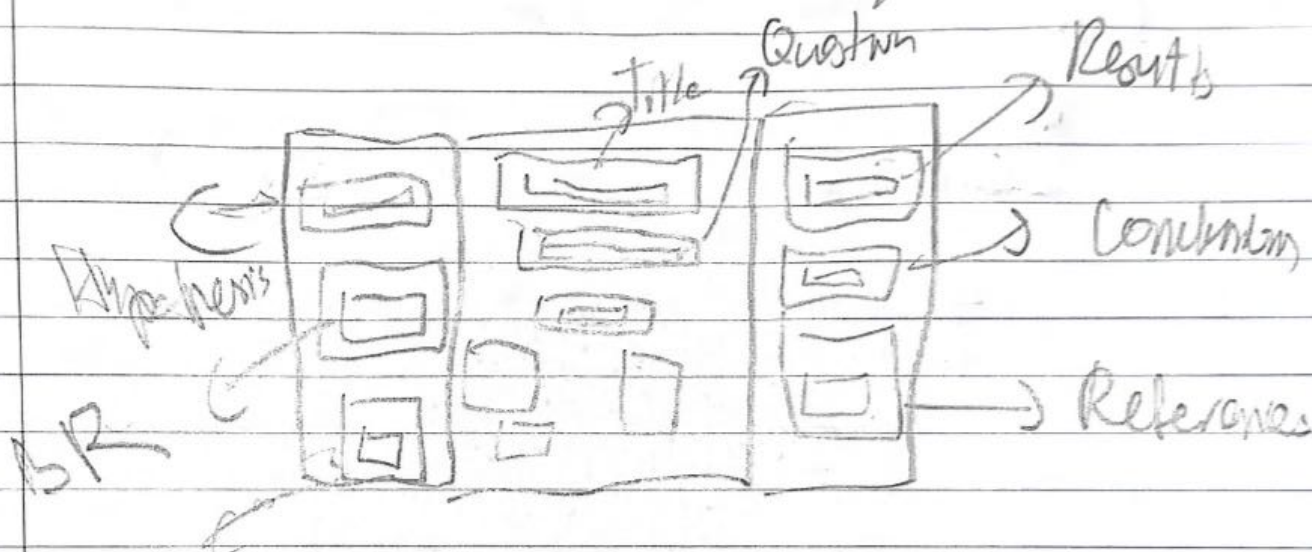
Other
Diagrams

x4

Mar 7

OTDI:

Finished Fully Conva layout



procedure

(flowchart)

Mar 8

OTDI:

- In layout, Switched BR & hypothesis
- Started building trifold
 - Finished middle fold

Mar 9

OTDI Looked in

- Finished entire trifold !!!
Now pdf reporting

Mar 10 OTD I had a GH test:

- Prepared (Reviewed)

Mar 12 OTD I got TODO List:

Science Fair TODO

- Complete CYSF portal
 - Finish declaration form (Finished Mar 13)
 - Finish paper things to portal
 - Methodology, Discussion, BI, etc.
 - Attachments
 - banner
 - 4:3 photo of me
 - 4:3 photo of my project
 - 10mm video
 - ~~to~~ labels on poster

→ Logbook

Mar 15

- OTDI
- Finished Method & BI of CYSF
 - Just IBC → IBC V of the [FINAL] version of the doc
 - (ran through Grammarly)

Mar 16

OTDI

- Did the same as yesterday, but for Discussion & Results

- need to add images & citations for

Mar 19

OTDI locked in (kmda)

- Added citations & images

- added banner, added 4:3

image of representation of project

- participant image: me in a suit

edited it to remove a piece of tape on it with pixlr.

Mar 20

- need to add image citations @ Mar 20

- need to edit poster for

video

- need to scan logbook

Mar. 20