OPERATION GRANNY FILES

The DNA Match-List Decoder

Mission 23 Companion Guide — Methodology Series

Methodology Guide & Workbook

Mission 23: Advanced DNA — Building Genetic Networks & Solving Mysteries

Clearance Level: Open File

operationgrannyfiles.substack.com

QUICK-GRAB SUMMARY

- Most of your 8,000+ matches are noise. The closest 20 to 30 do nearly all the genealogy work.
- **The Leeds Method** (Dana Leeds, 2018) sorts your second-cousin-range matches into four color clusters — one per grandparent.
- **Shared-match clustering** is the foundation of every advanced DNA technique; Leeds is the easiest recipe.
- **Triangulation** is when three or more independent lines share the same segment — that's evidence, not coincidence.
- **DNA Painter** turns segment data into a map of which chromosome region came from which ancestor.

What This Guide Is

This is the working-paper companion to Mission 23 of *Operation Granny Files*. It distills the three core advanced-DNA workflows — Leeds clustering, chromosome mapping, and triangulation — into a recipe you can run beside your laptop at the kitchen table.

It does not replace a textbook. It replaces the moment of *staring at 8,500 matches without knowing where to begin*. That moment is the bottleneck. Once you have a recipe, the rest is patience.

Section 1: Reading a Match List

A DNA match list is a ranked roster of every person in the testing company's database who shares enough DNA with you to call a cousin. The single most important number on the list is **shared centimorgans (cM)** — a unit of genetic distance. The larger the cM number, the closer the relationship.

Every major company sorts the list in descending order by shared cM by default. The matches at the top are your closest relatives. The matches at the bottom are statistical fog. Most of your genealogy energy belongs in the top hundred. Maybe the top thirty.

Centimorgans to Likely Relationship

The table below shows the *typical* range for each relationship. These are guidelines, not laws. Real cousin sharing is messy — a true second cousin might share 200 cM or 350 cM, and either is normal.

Shared cM	Likely Relationship	Confidence	Useful for Research?
3,400+	Identical twin / self	Certain	N/A
2,400 - 3,400	Parent / child / full sibling	Very high	Foundational
1,300 - 2,300	Half-sibling / grandparent / aunt-uncle / niece-nephew	Very high	Excellent

Shared cM	Likely Relationship	Confidence	Useful for Research?
575 - 1,300	First cousin / great-aunt-uncle / half-aunt-uncle	High	Excellent
215 - 650	First cousin once removed / second cousin	High	**Ideal for Leeds Method**
90 - 400	Second cousin / second cousin once removed	Medium-High	**Ideal for Leeds Method**
40 - 150	Third cousin / third cousin once removed	Medium	Good with clustering
15 - 75	Fourth cousin	Lower	Sometimes useful
5 - 15	Distant cousin (often noise)	Low	Rarely useful
Under 5	Statistical noise / IBS	Very low	Skip

IBS stands for *identical by state* — a segment that looks like sharing but is essentially coincidence. Below about 7 cM, most segments are IBS. The major companies have thresholds (Ancestry reports down to about 6 cM; FamilyTreeDNA much lower) but reporting it does not make it useful.

GRANNY PRO TIP

If a match shares less than 15 cM with you, do not even bookmark it for a project unless clustering tells you to. Spend that energy on the matches between 90 and 400 cM. That is where your grandparents are hiding.

Section 2: The Leeds Method — Step by Step

Dana Leeds, a genealogist in Texas, published her clustering recipe in 2018 and it has been the standard entry-point to shared-match analysis ever since. The method is intentionally low-tech: a spreadsheet, four highlighters, and an afternoon. It works on every major testing platform because every platform exposes the one thing you need — the **shared matches** list.

The Recipe

Step 1. Set the range. Open your match list and filter to matches sharing roughly **90 to 400 cM**. On Ancestry, this is the "2nd Cousin - 3rd Cousin" band. The exact cutoffs do not matter — what matters is that you are working with cousins close enough to share a recent ancestor (one of your eight great-grandparents at the most) but distant enough that each match is likely tied to just one of your four grandparent lines.

Step 2. Skip the close family. Remove parents, full siblings, half-siblings, first cousins, aunts, uncles, and any match above 400 cM from this exercise. They share with too many of your other matches across multiple grandparent lines, so they will pollute the clusters. You will use them later — just not here.

Step 3. Build a list. In a spreadsheet (or on paper), list the surviving 20-to-50 matches in descending cM order. Leave four blank columns to the right of each name — these will hold the cluster colors.

Step 4. Cluster from the top. Take the highest match in your list. Open her **shared matches** view on the testing platform. The shared matches are the people who match both you and her. Most testing platforms only show shared matches above a minimum threshold (Ancestry shows 20 cM and up), but that is fine.

Highlight the first match yellow. For every shared match that also appears on your filtered list, mark it yellow too.

Step 5. Find the next cluster. Scroll your list to the highest match that is **not yet yellow**. Open her shared matches. Highlight her green, and mark every one of her shared matches green as well.

Step 6. Repeat until you run out. Continue with blue, then pink. Most testers end up with somewhere between three and five primary clusters. Four is the textbook answer (one per grandparent) but real life is messier.

Step 7. Label the clusters. Once the colors stabilize, open one match in each color group and read her tree (or message her). The shared ancestors you find in her tree are usually one of your grandparents' families. Label the yellow column "Mom's dad," the green column "Mom's mom," and so on.

What the Clusters Mean

- **A clean four-cluster result:** rare and beautiful. Your four grandparents are well-represented in the test pool and you can now slot every new mystery match into a grandparent quadrant on sight.
- **A three-cluster result:** common. One of your grandparents is under-represented in the database, often because that side of the family did not test or comes from a region with low testing rates.
- **A two-cluster result:** suggests endogamy (see Section 6) or that two of your grandparents had a recent shared ancestor.
- **Matches in two clusters at once:** these are golden. They often descend from couples where both spouses share ancestors with you — a clue worth chasing.
- **Matches in no cluster:** distant cousins, noise, or close relatives you forgot to exclude in Step 2.

Automated Clustering Tools

Several platforms now do this work for you. They are wonderful — but only after you have done the manual version once, so you understand what the algorithm decided.

Tool	Platform	Cost	What It Does
Ancestry ProTools Clustering	Ancestry	Subscription add-on	One-click Leeds-style clusters with auto-color assignment
AutoClusters	MyHeritage	Included in subscription	Interactive grid of clusters with shared-match overlay
GEDmatch Tier1	GEDmatch (uploaded	Tier1 subscription	Manual or automated

Tool	Platform	Cost	What It Does
Clustering	kits)		clustering using any uploaded kit
Collins Leeds Method (CLM)	DNAGedcom	Subscription	Bulk clustering plus extended Leeds variants
Genetic Affairs AutoCluster	Stand-alone (when running)	Variable	Independent clustering across multiple platforms

GRANNY PRO TIP

Run the manual Leeds on your top twenty matches first. Then run the automated clustering. Compare. The places where they disagree are usually the most interesting places in your tree.

Section 3: Triangulation — Turning Coincidence into Evidence

Two matches sharing DNA with you, individually, prove almost nothing. You and Match A might descend from your father's mother's family. You and Match B might descend from your father's father's family. Both real cousins, neither related to each other.

Triangulation is when three or more matches share the same DNA segment *and* descend from independent lines that all lead back to the same most-recent-common-ancestor couple (the MRCA). That is no longer coincidence — that is evidence.

The Triangulation Rule

A triangulated segment requires:

1. **Three or more** people sharing the same chromosome segment with each other and with you.
2. **Independent lines of descent** — they are not all from the same branch of one cousin's tree. Three siblings sharing a segment is one line, not three.
3. **A documented MRCA** — a paper-trail ancestor (or ancestral couple) that all three descend from through different children.
4. **A meaningful segment size** — generally 7 cM or larger to reduce the risk of identical-by-state false positives.

When all four hold, you have built genetic evidence that meets the Genealogical Proof Standard's "sufficient quality and quantity" bar — the framework we will walk through in Mission 19.

Where to Do Triangulation

Triangulation requires **segment data** — the chromosome number and start/end positions of the shared DNA. The major platforms split on this:

Platform	Releases Segment Data?	Triangulation-Friendly?
Ancestry	No	No (clustering only)
23andMe	Yes	Yes (via DNA Comparison tool)
MyHeritage	Yes	Yes (Chromosome Browser + Triangulation tool)
FamilyTreeDNA	Yes	Yes (Chromosome Browser + Matrix)
GEDmatch	Yes (uploaded kits)	Yes (Triangulation utilities, multi-kit tools)
Living DNA	Limited	Limited

The practical workflow: take a cluster from your Leeds chart, upload kits to GEDmatch (or use MyHeritage's tools), and look for groups of three or more matches who share the **same chromosome segment** with you. That is your triangulated group.

Section 4: Chromosome Mapping with DNA Painter

DNA Painter (dnapainter.com), built by Jonny Perl, is the dominant free tool for taking the triangulation idea and turning it into a living map of your own chromosomes. The result is sometimes called a **chromosome map** or, less formally, your *painted profile*.

What It Does

You feed DNA Painter the segment data for each of your matches, one match at a time. For each segment, you tell the tool the most-recent-common-ancestor you and that match share — usually a couple, four to seven generations back. DNA Painter then **paints** that segment a color tied to that ancestral couple.

Over time, the 22 autosomal chromosomes (plus the X) develop colored regions. A red region on chromosome 7 is from your great-grandfather Albert and his wife. A blue region on chromosome 14 is from your great-grandmother Agnes's parents. Eventually most of your DNA has a label, and you can read your own chromosomes like a map.

Why It Matters

Once a region is painted, the **next** match whose shared segment falls inside that region is almost certainly descended from the same ancestral couple. You can place that new match on the tree before you have opened a single document. That is a 20-minute task reduced to a 2-minute task.

More importantly: a region painted with three or more independent lines all descending from the same couple **is** triangulation, captured visually and ready for an evidence statement.

Getting Started in 30 Minutes

1. Make a free DNA Painter account at dnapainter.com.

2. Pick your five highest-cM matches who have a known, paper-trail-confirmed ancestor in common with you.
3. Download segment data for each from 23andMe, MyHeritage, FamilyTreeDNA, or GEDmatch.
4. Paint the segments to the shared ancestor.
5. Save the profile. Come back next weekend with five more matches.

GRANNY SAYS

A painted chromosome is your inheritance, drawn in colored pencil. Once you can see which region came from which great-grandparent, every new mystery match is just a question of where her segment lands on your map.

What Ancestry Users Should Know

Ancestry does not release segment data, full stop. This is the single largest limitation in genetic genealogy as a beginner-friendly hobby. Workarounds:

- Find your Ancestry matches who have **also** uploaded to MyHeritage, FamilyTreeDNA, or GEDmatch. You can then triangulate with those second-platform results.
- Test your closest Ancestry relatives on a second platform. A father, mother, sibling, or aunt/uncle on MyHeritage gives you segment data for the lines they share with you.
- Politely ask cooperative matches to upload their raw Ancestry data to GEDmatch (free) — this is increasingly common etiquette in the genealogy community.

Section 5: Building a Genetic Network Diagram

Once you have several clusters and a beginning of a painted profile, you may want a single visual that puts everything in one place. This is the **genetic network diagram** — a hand-drawn or software-drawn picture of your matches connected by shared DNA and shared ancestors.

What Goes on the Diagram

- **You**, at the center.
- **Your closest documented relatives** (parents, siblings, aunts, uncles) on the inside ring.
- **Each cluster** as a colored region: the yellow group on one side, green on another.
- **The MRCA (most-recent-common-ancestor) couple** for each cluster, written as the label for that region.
- **Bridge matches** — matches that fall in two clusters — drawn straddling the border between them, with a note about what shared ancestral couple links the two lines.

You do not need software for this. A blank sheet of paper and four colored pens are enough. The point is to externalize the relationships so you can see them at a glance instead of holding the whole map in your head.

Software Options

If you want a digital version, three options are common:

- **Gephi** — free, technical, powerful. Steep learning curve.

- **NodeXL** — Excel-based network analysis. Friendlier for spreadsheet users.
- **DNAGedcom Client** — produces network graphs directly from your match data on supported platforms.

For most beginner work, paper is faster and shows you the same information.

Section 6: When DNA Confirms vs. Contradicts the Paper Trail

DNA is **evidence**. So is a marriage record, so is a deed, so is a baptismal entry, so is the family Bible. The Genealogical Proof Standard (Mission 19) treats them as the same kind of input — all evidence — but it asks you to weigh them with their respective strengths and weaknesses in mind.

What DNA Does Well

- Confirms or denies that two living people share a recent common ancestor.
- Tests parentage hypotheses with high confidence at close ranges.
- Identifies which side of a tree a mystery cousin belongs to (via clustering).
- Detects long-hidden relationships (adoptions, misattributed parentage, sealed records).

What DNA Does Less Well

- Naming a specific ancestor more than four or five generations back is hard without paper records to anchor the segment.
- Distinguishing between two siblings as the source of a segment is sometimes impossible.
- Endogamy inflates cM and breaks standard relationship predictions.
- Adoption, donor conception, and non-paternity events appear as "anomalies" — the DNA does not lie, but the paper trail might.

The Honest Reconciliation Workflow

When DNA evidence and paper evidence disagree, you have one of three situations:

1. **The paper is wrong.** A typed name, a misremembered marriage, a transcription error, or a deliberately falsified record. DNA usually wins here.
2. **The DNA interpretation is wrong.** The segment was too small, the relationship estimate was off because of endogamy, or the wrong MRCA was assigned. Re-examine before rejecting the paper.
3. **A non-paternity event (or its analogues) occurred.** Adoption, an affair, a re-marriage where children kept the new surname, a child of a stepfather raised as his own. This is genuinely common — the literature suggests 1 to 4 percent of births per generation. The paper trail is correct as social history; the DNA is correct as biology. Both are valid.

GRANNY PRO TIP

Never publish a conclusion that "Aunt Hattie's father was not actually Edward" without first making absolutely sure — and then thinking long and hard about who is alive to read your finding and what they would feel about it. DNA reveals things families spent generations not revealing. Use that power gently.

Section 7: Endogamy — When the Math Breaks

Some communities marry within themselves for many generations. Their descendants share unusually high amounts of DNA simply because their ancestors are related to each other through multiple lines, not just one. This is **endogamy**, and it breaks standard cousin-relationship estimators.

Communities Where Endogamy Matters

- Ashkenazi Jewish populations
- Acadian and Cajun communities
- Mennonite and Amish populations
- Several Caribbean islands (small founding populations)
- Pioneer-era Latter-day Saint families in the Mountain West
- Many Indigenous American nations
- Small island populations worldwide (Iceland, the Faroes, parts of Polynesia)
- Long-isolated rural communities

What Endogamy Does to Your Match List

- A "predicted second cousin" at 350 cM might actually be a fourth cousin twice over.
- The 4th-to-8th-cousin range becomes enormous — thousands of matches that are technically distant cousins through many lines.
- Clustering still works, but each cluster may contain matches who are actually related through *several* of your grandparents, not one.
- Triangulation requires much larger segment thresholds (often 10 cM or 15 cM minimum) to be reliable.

Workarounds

- Use larger cM cutoffs throughout — start your Leeds at 200 cM rather than 90 cM.
- Rely more heavily on paper records to anchor clusters.
- Test multiple relatives — a sibling, parent, or aunt/uncle gives you side-of-family information that endogamous clusters obscure.
- Be very cautious with relationship-predictor tools — they were trained on outbred populations and will mislead you systematically.

A future OGF mission will tackle endogamy in full. For now, simply knowing it exists in your tree is the first step.

Section 8: The Match-List Decoder Worksheet

Use this worksheet at your kitchen table the first time you run the Leeds Method.

Match Cluster Tracking Sheet

Match Name / ID	Shared cM	Yellow	Green	Blue	Pink	Notes

Match Name / ID	Shared cM	Yellow	Green	Blue	Pink	Notes

Cluster Labels (fill in after clustering)

- Yellow cluster: _____
- Green cluster: _____
- Blue cluster: _____
- Pink cluster: _____

Section 9: Glossary

Centimorgan (cM). A unit of genetic distance. The number of cM you share with a match is the standard predictor of how closely related you are. More cM means closer relationship.

Endogamy. Marriage within a closed community over many generations, which inflates shared DNA between any two members.

Identical by descent (IBD). A DNA segment that you and a match share because you inherited it from a common ancestor. This is what you want.

Identical by state (IBS). A DNA segment that looks like sharing but is actually coincidence — the same sequence arose independently. This is statistical noise; usually below 7 cM.

Leeds Method. Dana Leeds's 2018 recipe for clustering DNA matches into colored groups, one per grandparent, using only shared-match data.

Most Recent Common Ancestor (MRCA). The most recent person — or ancestral couple — that you and a match both descend from.

Segment. A specific stretch of DNA on a specific chromosome, defined by a chromosome number, a start position, an end position, and a length in cM.

Shared matches. People who appear in both your match list and another person's match list. The foundation of clustering.

Triangulation. When three or more matches share the same segment, descend from the same most-recent-common-ancestor through independent lines, and the segment is large enough to be confident (typically 7+ cM).

Unknown parentage research. Genetic-genealogy techniques used by adoptees, donor-conceived adults, and others who do not know one or both biological parents.

What Are The Odds (WATO). A DNA Painter tool that lets researchers test specific parental hypotheses against the shared-cM evidence.

Section 10: Quick-Reference Mission Checklist

Before you close this guide, walk through this checklist with your match list open.

- [ ] I have filtered to matches in the 90 to 400 cM range.
- [ ] I have removed close family (parents, siblings, first cousins) from the cluster pool.
- [ ] I have at least 20 matches to work with.
- [ ] I have built or downloaded a tracking spreadsheet.
- [ ] I have assigned colors to the highest unassigned match and its shared matches, working down the list.
- [ ] I have labeled each color cluster with the grandparent line it represents.
- [ ] I have noted any matches that appear in two clusters and flagged them for later investigation.
- [ ] I have decided whether chromosome mapping (DNA Painter) is my next move or whether I am stopping at clustering for now.
- [ ] I have a file or notebook where these clusters live, so I do not have to re-do this work next time.

GRANNY SAYS

The match list is not the enemy. It is just the largest map you have ever been handed, and nobody told you where to put the pin. Pick a corner. Color it in. Come back tomorrow. The picture will keep developing.

Notes

Use this space for personal notes and discoveries.