



Development and Semi-Automated Analysis of an *in vitro* Dissemination Model
for Myeloma Cells Interacting with Mesenchymal Stromal Cells

Entwicklung und semi-automatisierte Analyse eines *in vitro* Modells
für die Disseminierung von Myelomzellen in Interaktion mit mesenchymalen Stromazellen

Doctoral Thesis for a Doctoral Degree

at the

GRADUATE SCHOOL OF LIFE SCIENCES,
JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG,
SECTION BIOMEDICINE

submitted by

Martin Kuric

from

Bad Neustadt a.d. Saale

Würzburg, 2024

Submitted on:
Office stamp

Members of the *Promotionskomitee*:

Chairperson: TODO: CHAIRPERSON
Primary Supervisor: Prof Dr. rer. nat. Regina Ebert
Supervisor (Second): Prof Dr. med. Franziska Jundt
Supervisor (Third): Prof Dr. rer. nat. Torsten Blunk

Date of Public Defense:

Date of Receipt of Certificates:

This work was conducted at the Department of Musculoskeletal Tissue Regeneration (Bernhard-Heine-Centre for Locomotive Research), University of Würzburg from 08.04.2018 to 31.03.2024 under the supervision of Prof Dr. rer. nat. Regina Ebert.

Summary

 Lorem Ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum

Zusammenfassung

 Lorem Ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum bla

Acknowledgements

 Lorem Ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor
 invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam
 et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem
 ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy
 eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero
 eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata
 sanctus est Lorem ipsum

Contents

Summary	i
Acknowledgements	ii
Introduction	1
Human Mesenchymal Stem/Stromal Cells	2
Multiple Myeloma	3
Myeloma-hMSC Interactions	3
Myeloma Bone Disease	3
Dissemination of Myeloma Cells	4
Code-Automation as a Standard in Modern Biosciences	5
How Code Quality Improves Scientific Reproducibility	6
Python as a Programming Language	8
Data Science with Python	12
Aims	14
Chapter 1: Modelling Myeloma Dissemination <i>in vitro</i>	15
Abstract	15
Introduction	15
Methods	15
Results	15
Discussion	15
Research Article: Cancer Research Communications	16
Chapter 2: Semi-Automation of Data Analysis	61
Abstract	61
Introduction	61
Statement of Need	64
Example	64
Overview	66
Discussion	67
Summarising Discussion	69
Time Lapse	69
Myeloma	69
Semi-Automated Analysis Improves Agility During Establishing new <i>in vitro</i> Methods	69
Appendix	76
Methods: Supplementary Figures and Methods	77
Author Contributions to Research Projects	113
Author Contributions to Figures and Tables	116
Affidavit	119
Curriculum Vitae	120

Introduction

To provide a comprehensive background for the following chapters that focus on the interaction of human mesenchymal stromal cells (hMSCs) with multiple myeloma (MM) cells, this

Human Mesenchymal Stem/Stromal Cells

Explaining what a mesenchymal stromal cell (MSC) is, is not such an easy task as one might expect. MSCs are derived from multiple MSCs different sources, serve a wide array of functions and are always isolated as a heterogenous group of cells. This makes it particularly challenging to find a consensus on their exact definition, nomenclature, exact function and *in vivo* differentiation potential. Therefore, the most effective approach to describe hMSCs is to present their historical context.

hMSCs first gained popularity as a stem cell. Stem cells lay the foundation of multicellular organisms. Embryonic stem cells orchestrate the growth and patterning during embryonic development, while adult stem cells are responsible for regeneration during adulthood. The classical definition of a stem cell is that of a relatively undifferentiated cell that divides asymmetrically, producing another stem cell and a differentiated cell (Cooper, 2000; Shenghui et al., 2009). Because of their significance in biology and regenerative medicine, stem cells have become a prominent subject in modern research. Especially human mesenchymal stromal cells (hMSCs) have proven to be a promising candidate in this context (Ullah et al., 2015).

Mesenchyme first appears in embryonic development during gastrulation. There, cells that are committed to a mesodermal fate, lose their cell junctions and exit the epithelial layer in order to migrate freely. This process is called epithelial-mesenchymal transition (Tam & Beddington, 1987; Nowotschin & Hadjantonakis, 2010). Hence, the term mesenchyme describes non-epithelial embryonic tissue differentiating into mesodermal lineages such as bone, muscles and blood. Interestingly, it was shown nearly twenty years earlier that cells within adult bone marrow seemed to have mesenchymal properties as they were able to differentiate into bone tissue (A. J. Friedenstein et al., 1966; A. Friedenstein & Kuralesova, 1971; Bianco, 2014). This was the origin of the “mesengenic process”-hypothesis: This concept states that mesenchymal stem cells serve as progenitors for multiple mesodermal tissues (bone, cartilage, muscle, marrow stroma, tendon, fat, dermis and connective tissue) during both adulthood and embryonic development (A. Caplan, 1991; A. I. Caplan, 1994). The mesenchymal nature of these cells (termed bone marrow stromal cells: BMSCs) was confirmed later when they were shown to differentiate into adipocytic (fat) and chondrocytic (cartilage) lineages (Pittenger et al., 1999). Since then, the term “mesenchymal stem cell” (MSC) has grown popular as an adult multipotent precursor to a couple of mesodermal tissues. hMSCs derived from bone marrow (hMSCs) were shown to differentiate into osteocytes, chondrocytes, adipocytes and cardiomyocytes (Gronthos et al., 1994; Muruganandan et al., 2009; Xu et al., 2004) Most impressively, these cells also exhibited ectodermal and endodermal differentiation potential, as they produced neuronal cells, pancreatic cells and hepatocytes (Barzilay et al., 2009; Wilkins et al., 2009; Gabr et al., 2013; Stock et al., 2014).

Furthermore, cultures with MSC properties can be established from “virtually every post-natal organs and tissues”, and not just bone marrow (da Silva Meirelles et al., 2006). However, it has to be noted that hMSCs can differ greatly in their transcription profile and *in vivo* differentiation potential depending on which tissue they originated from (Jansen et al., 2010; Sacchetti et al., 2016).

Since “hMSCs” are a heterogenous group of cells, they were defined by their *in vitro* characteristics. A minimal set of criteria are the following (Dominici et al., 2006): First, hMSCs must be plastic adherent. Second, they must express or lack a set of specific surface antigens (positive for CD73, CD90, CD105; negative for CD45, CD34, CD11b, CD19). Third, hMSCs must differentiate to osteoblasts, adipocytes and chondroblasts *in vitro*. Together, hMSCs exhibit diverse differentiation potentials and can be isolated from multiple sources of the body. This offers great opportunity for regenerative medicine, if the particular hMSC-subtype is properly characterized.

Multiple Myeloma

Multiple myeloma arises from clonal expansion of malignant plasma cells in the bone marrow (BM). At diagnosis, myeloma cells have disseminated to multiple sites in the skeleton and, in some cases, to virtually any tissue (Rajkumar & Kumar, 2020; Bladé et al., 2022).

Myeloma-hMSC Interactions

Since plasma cells can not survive outside the bone marrow, MM cells also require survival signals for growth and disease progression. These signals are produced by the bone marrow microenvironment, including ECM, MSCs and ACs (Kibler et al., 1998; García-Ortiz et al., 2021).

Myeloma Bone Disease

Bone is a two-phase system in which the mineral phase provides the stiffness and the collagen fibers provide the ductility and ability to absorb energy (Viguet-Carrin et al., 2006). On a molecular level, bone tissue is composed of extracellular matrix (ECM) proteins that are calcified by hydroxyapatite crystals. This ECM consists mostly of collagen type I, but also components with major regulatory activity, such as fibronectin and proteoglycans that are essential for healthy bone physiology (Alcorta-Sevillano et al., 2020). Bone tissue is actively remodeled by bone-forming osteoblasts and bone-degrading osteoclasts. Osteoblasts are derived from mes-

enchymal stromal cells (MSCs) that reside in the bone marrow (A. J. Friedenstein et al., 1966; Pittenger et al., 1999). MSCs also give rise to adipocytes (ACs) to form Bone Marrow Adipose Tissue (BMAT), which can account for up to 70% of bone marrow volume (Fazeli et al., 2013).

MM indirectly degrades bone tissue by stimulating osteoclasts and inhibiting osteoblast differentiation, which leads to MM-related bone disease (MBD) (Glavey et al., 2017). MBD is present in 80% of patients at diagnosis and is characterized by osteolytic lesions, osteopenia and pathological fractures (Terpos et al., 2018).

Dissemination of Myeloma Cells

dissemination is still widely unclear - multistep process - invasion, intravasation, intravascular arrest, extravasation, colonization - overcome adhesion, retention, and dependency on the BM microenvironment - loss of adhesion factors such as CD138

Code-Automation as a Standard in Modern Biosciences

Beschreibe die Situation. - Big Data in Biosciences - what is big data, examples - Define citable challenges: - reproducibility crisis - lack of tools

In recent years, the biosciences have evolved dramatically, with a marked increase in the volume and complexity of data generated (Yang et al., 2017; Ekmekci et al., 2016). This transformation necessitates robust software tools, many of which require coding skills to use effectively. Here we summarize standard tools used by biosciences today and show their reliance on coding. The author argues that the role of a modern independent researcher is now intertwined with coding skills similar to a role of “precision medicine bioninformatician” (Gómez-López et al., 2019).

Statistical analysis in biosciences has traditionally been reliant on user-friendly tools like Excel and GraphPad Prism. While Excel by itself is recognized as limited for complex data analysis (Tanavalee et al., 2016; Incerti et al., 2019), GraphPad Prism offers more advanced statistical models .

However, increasingly demands more sophisticated approaches as data sets grow in size and complexity.

R and Python scripts offer more efficient and versatile solutions, enabling complex analyses with a few lines of code (R Core Team, 2018; Vallat, 2018).

Recognizing this trend, Microsoft has integrated a Python interpreter into Excel to computations more accessible within a widely used platform (?).

Next-generation sequencing, such as bulk RNAseq, has become affordable, allowing for larger sample sets during a single PhD project. This technology offers advanced tools that are most efficiently used through scripting in R or Python. In the absence of a dedicated statistician, researchers are compelled to learn coding.

In gene ontology, tools such as Metascape facilitate the integration of vast datasets and outputs multiple useful data visualizations. Metascape also provides multiple excel sheets, containing all results, sometimes in a nested format, which provides even further information that's adaptable for specific hypotheses, but given the sheer amount of data, is impractical to analyze manually.

since Metascape returns large Excel sheets with complex nested information, a researcher without coding skills requires manual work to adapt the results to specific research hypotheses.

its true potential is unlocked only when researchers can manipulate and analyze these data through scripting.

Modern gene ontology tools like Metascape offer powerful graphical user interfaces. However, their effectiveness is only possible through standardizing multiple large datasets.

The output from Metascape, large Excel sheets with complex nested information, is more efficiently analyzed through scripting, which is often necessary to adapt metascape results to specific research hypotheses.

Image analysis is another area where coding skills are essential. ImageJ/FIJI, a standard tool in the field, requires scripting for batch processing of multiple images and automating multiple processing steps into a pipeline. While macros can be recorded, understanding the underlying code is necessary for troubleshooting and adapting the macro to new datasets.

In the field of protein structural biology, Pymol is a standard tool that also has a Python command interface.

Similarly, artificial intelligence (AI), a game-changer in biomedicine, primarily uses Python due to its extensive libraries for machine learning and scientific computing. Python is also a standard for integrative biomedicine simulations.

Finally, databases and repositories are essential for storing, retrieving, and sharing data. Researchers need to understand common file formats to adhere to standards that ensure re-usability and interoperability. Scripting helps automate the process of formatting data for submission to these databases.

In conclusion, the integration of coding in bioscience research is not just a trend but a necessity. As the field continues to evolve, the demarcation between biologists and computational scientists blurs, underscoring the importance of coding skills for the next generation of researchers. The ability to code is fast becoming an indispensable asset, as integral to bioscience as traditional laboratory skills.

How Code Quality Improves Scientific Reproducibility

A main reason to write software is to define re-usable instructions for task automation (Narzt et al., 1998). However, the complexity of the code makes it prone to errors and can prevent usage by persons other than the author himself. This is a problem for the general scientific community, as the software is often the only way to reproduce the results of a study (Sandve et al., 2013). Hence, modern journals aim to enforce standards to software development, including software written and used by biological researchers (Smith et al., 2018). Here, we provide a brief overview of the standards utilized by `plotastic` that to ensure its reliability and reproducibility by the scientific community (Peng, 2011).

Modern software development is a long-term commitment of maintaining and improving

code after initial release (Boswell & Foucher, 2011). Hence, it is good practice to write the software such that it is scalable, maintainable and usable. Scalability or, to be precise, structural scalability means that the software can easily be expanded with new features without major modifications to its architecture (Bondi, 2000). This is achieved by writing the software in a modular fashion, where each module is responsible for a single function. Maintainability means that the software can easily be fixed from bugs and adapted to new requirements (Kazman et al., 2020). This is achieved by writing the code in a clear and readable manner, and by writing tests that ensure that the code works as expected (Boswell & Foucher, 2011). Usability is hard to define (Brooke, 1996), yet one can consider a software as usable if the commands have intuitive names and if the software’s manual, termed “documentation”, is up-to-date and easy to understand for new users with minimal coding experience. A software package that has not received an update for a long time (approx. one year) could be considered abandoned. Abandoned software is unlikely to be fully functional, since it relies on other software (dependencies) that has changed in functionality or introduce bugs that were not expected by the developers of all dependencies. Together, software that’s scalable, maintainable and usable requires continuous changes to its codebase. There are best practices that standardize the continuous change of the codebase, including version control, continuous integration (often referred to as CI), and software testing.

Version control is a system that records changes to the codebase line by line, allowing the documentation of the history of the codebase, including who made which changes and when. This is required to isolate new and experimental features into newer versions and away from the stable version that’s known to work. The most popular version control system is Git, which is considered the industry standard for software development (Chacon & Straub, 2024). Git can use GitHub.com as a platform to store and host codebases in the form of software repositories. GitHub’s most famous feature is called “pull request”. A pull request is a request from anyone registered on GitHub to include their changes to the codebase (as in “please pull this into your main code”). One could see pull requests as the identifying feature of the open source community, since it exposes the codebase to potentially thousands of independent developers, reaching a workforce that is impossible to achieve with closed source models used by paid software companies.

Continuous integration (CI) is a software development practice in which developers integrate code changes into a shared repository several times a day (Duvall et al., 2007). Each integration triggers the test suite, aiming to detect errors as soon as possible. The test suite includes building the software, setting up an environment for the software to run and then executing the programmed tests, ensuring that the software runs as a whole. Continuous integration is often used together with software branches. Branches are independent copies of the codebase that are meant to be merged back into the original code once the changes are finished. Since branches

accumulate multiple changes over time, this can lead to minor incompatibilities between the branches of all developers (integration conflicts), which is something that CI helps to prevent.

Continuous integration especially relies on a thorough software testing suite. Software testing is the practice of writing code that checks if the codebase works as expected (Myers et al., 2011). The main type of software testing is unit testing, which tests the smallest units of the codebase (functions and classes) in isolation (Listing 1).

Listing 1: Example of an arbitrary python function and its respective unit test function. The first function simply returns the number 5. The second function tests if the first function indeed returns the number 5. The test function is named with the prefix “test_” and is placed in a file that ends with the suffix “_test.py”. The test function is executed by the testing framework pytest. Note that code after “#” is considered a comment and won’t be executed.

```
1 # Define a function called "give_me_five" that returns the number 5
2 def give_me_five():
3     return 5
4 # Define a test function asserting that "give_me_five" returns 5
5 def test_give_me_five():
6     assert give_me_five() == 5
```

The quality of the software testing suite is measured by the code coverage, the precision of the tests, and the number of test-cases that are checked. The code coverage is the percentage of the codebase that is called by the testing functions, which should be as close to 100% as possible, although it does not measure how well the code is tested. The precision of the test is not a measurable quantity, but it represents if the tests truly checks if the code works as expected. The number of test-cases is the number of different scenarios that are checked by the testing functions, for example testing every possible option or combinations of options for functions that have multiple options. The most popular software testing framework for python is `pytest`, which is utilized by `plotastic` (Krekel et al., 2004).

Together, the standards of software development, including version control, continuous integration, and software testing, ensure that the software is scalable, maintainable, and usable. This is especially important for software that is used by the scientific community, as it ensures that the software is working as expected at defined versions years after publishing scientific results.

Python as a Programming Language

Here, we provide a general overview of the python programming language, explaining terms like “*type*”, “*method*”, etc., in order to prepare readers without prior programming experience for the following chapters. We also describe the design principles of python to lay out the key concepts that differentiate python compared to other programming languages. A more detailed tutorial on python that’s specialized for bioscientists is found in Ekmekci et al. 2016

Listing 2: Example of readable python code. This one-line code returns the words (string) 'Hello, World!' when executed. The command is straightforward and easy to understand.

```
1 print("Hello, World!")
2 // Output: Hello, World!
```

Languages such as python are considered “*high-level*”, which means that it is designed to be easy to read and write, but also independent of hardware by hiding (“*abstracting*”) underlying details (*The Python Language Reference*, n.d.). A key principle of python is the emphasis on implementing a syntax that is concise and close to human language (Listing 2, Listing 3).

Listing 3: Example of less readable code written in the low-level programming language C. This code is doing exactly the same as the python code in Listing 2. The command is harder to understand because more steps are needed to access the same functionality, including the definition of a function

```
1 #include <stdio.h>
2 int main() {
3     printf("Hello, World!");
4     return 0;
5 }
6 // Output: Hello, World!
```

Furthermore, python is an *interpreted* language, which means that the code is executed line by line. This makes coding easier because the programmer can see the results of the code immediately after writing it, and error messages point to the exact line where the error occurred. This is in contrast to *compiled* languages, where the code has to be compiled into machine code before it can be executed. The advantage of compiled languages is that the code runs faster, because the machine code is optimized for the hardware.

Python automates tasks that would otherwise require an advanced understanding of computer hardware, like the need for manual allocation of memory space. This is achieved by using a technique called “*garbage collection*”, which automatically frees memory space that is no longer needed by the program. This is a feature that is not present in low-level programming languages like C or C++, that were designed to maximize control over hardware.

Another hallmark of python is its *dynamic typing system*. In python the type is inferred automatically during code execution (Listing 4). This is in contrast to *statically* typed languages like C, where the type of a variable has to be declared explicitly and cannot be changed during code execution (Listing 5) (*The Python Language Reference*, n.d.).

Dynamic typing makes python a very beginner-friendly language, since one does not have to keep track of the type of each variable. However, this also makes python a slower language, because the interpreter has to check the type of each variable during code execution. Also, developing code with dynamic typing systems is prone to introducing bugs (“*type errors*”), because it allows unexperienced developers to convert variables from one type to another without noticing, leading to unexpected behavior. Hence, larger python projects require disciplined

Listing 4: Example of dynamic typing in python. The variable “a” is assigned the value 5, which is of type integer. The variable “a” is then assigned the value “Hello, World!”, which is of type string. Python allows dynamic re-assignment of variables with different types. Note that code after “#” is considered a comment and won’t be executed.

```
1 a = 5 # Type integer
2 a = 5.0 # Type float
3 a = 'Hello, World!' # Type string
4 a = True # Type boolean
5 a = False # Type boolean
6 a = [1, 2, 3] # Type list of integers
7 a = { 'name': 'Regina' } # Type dictionary
```

Listing 5: Example of static typing in C. The variable “a” is declared as an integer (int), and can only store integers. The variable “a” is then assigned the value 5, which is an integer. The variable “a” is then assigned the value ‘Hello, World!’, which is a string. This results in a compilation error, because the variable “a” can only store integers. Note that code after “//” is considered a comment and won’t be executed.

```
1 int a; // Declare type as integer (int)
2 a = 5;
3 a = 'Hello, World!'; // Compilation error!
```

adherence to programming conventions. One such convention is *type hinting*, which is a way to explicitly note the type of a variable. Type hinting does not have an effect on the code, but it makes the code more readable and understandable for other developers, and allows for development environments to detect type errors before execution (Listing 6) (van Rossum et al., 2014).

Listing 6: Example of type hints used in python. Explicitly stating the type of the variable is optional and does not change the behavior of the code as shown in Listing 4.

```
1 a: int = 5
2 a: str = 'Hello, World!'
```

Python supports both functional and object-oriented programming paradigms. In functional programming, the code is written in a way that the program is a sequence of function calls, where each function call returns a value that is used in the next function call (Listing 7). This approach is useful when multiple actions have to be performed on the same data and the structure of the data is relatively simple, for example a string of a gene sequence.

When the data itself gains in complexity, for example when storing not just the gene sequence, but also the promotor sequence, an object-oriented approach is more suitable (Listing 8). Object-oriented programming is a programming paradigm that uses objects and classes. An object is a collection of both data and functions, and a class is a blueprint for creating objects. The data of an object is stored as attributes. Functions that are associated with an object are called methods.

Listing 7: Example of functional programming in Python. The code defines a function called “find_restriction_site” that finds the position of a restriction site in a gene. The function “cut” uses the function “find_restriction_site” to cut the gene at the restriction site.

```

1 def find_restriction_site(gene: str):
2     return gene.find('GCGC')
3
4 def cut(gene: str):
5     position = find_restriction_site(gene)
6     return gene[position:]
7
8 gene1 = 'TGAGCTGAGCTGATGCGCTATATTAGGCG'
9 gene1_cut = cut(gene1)
10 print(gene1_cut)
11 # Output: GCGCTATATTAGGCG

```

Listing 8: Example of object oriented programming in python. The class is called “Gene” and has four methods, “__init__”, “find_promotor”, “find_restriction_site” and “cut”. The function “__init__” is called when creating (“initializing”) an object, which fills the object with user-defined data. The parameter “self” is used to reference the object itself internally. “find_promotor” is a method that finds the position of the promotor in the gene and is called during object initialization.

```

1 class Gene:
2     def __init__(self, sequence: str):
3         self.sequence: str = sequence # Save sequence as attribute
4         self.promotor: str = self.find_promotor()
5     def find_promotor(self):
6         return self.sequence.find('TATA')
7     def find_restriction_site(self):
8         return self.sequence.find('GCGC')
9     def cut(self):
10        position = self.find_restriction_site()
11        return self.sequence[position:]
12
13 gene1 = Gene(sequence='TGAGCTGAGCTGATGCGCTATATTAGGCG') # Create object
14 gene1_cut = gene1.cut() # Call the method cut
15 print(gene1_cut) # Show result
16 # Output: GCGCTATATTAGGCG

```

A major benefit of using an object oriented versus a functional approach is that the data itself programmable, enabling the programmer to define the behavior of the data itself through methods. This is achieved by using the keyword “self” to reference the object itself inside the class. For example, one could extend the class “Gene” with a method that finds the promotor of the gene and stores it as an attribute (Listing 8).

When designing software, both functional and object oriented programming can be used together, where object oriented programming is often used to design the program’s overall architecture, and functional programming is used to implement the algorithms of the program’s features. This allows for scalability of the software, as every single class is extended through the

addition of new methods. Furthermore, classes can be expanded in their functionalities through inheritance (Listing 9).

Listing 9: Example of inheritance in python. The class “mRNA” inherits from the class “Gene”. The class “mRNA” has two methods, “`__init__`” and “`find_stopcodon`”. The method “`find_stopcodon`” finds the position of stop codons.

```
1 # Define a class called mRNA inheriting from the class Gene
2 class mRNA(Gene):
3     def __init__(self, sequence: str):
4         super().__init__(sequence) # Get attributes from parent class
5         self.sequence.replace('T', 'U') # Replace thymine with uracil
6     def find_stopcodons(self):
7         return self.sequence.find('UGA')
8
9 mRNA1 = mRNA(sequence='TGAGCTGAGCTGATGCGCTATTTAGGCC') # Create object
10 print(mRNA1.find_stopcodons()) # Call the method translate
11 # Output: [0, 5, 10]
```

Inheritance is a feature of object-oriented programming that allows a class to access every attribute and method of a parent class. For example, one could extend the class “Gene” with a class “mRNA”, by writing a class “mRNA” that inherits from the class “Gene”.

Together, python is not just beginner-friendly, but also well respected for its ease in development, which is why it is widely used in professional settings for web development, data analysis, machine learning, biosciences and more (Ekmekci et al., 2016).

Data Science with Python

the ease of use has made python a very popular language (Rayhan & Gross, 2023)

Like any other programming language, python alone does not provide specialized tools like those used for data analysis (*The Python Language Reference*, n.d.). However, python was designed to be extended by packages developed by its users. A python package consists of multiple python modules, where each module is a text-file with a .py ending containing python code. Famous examples of such packages are `pytorch` and `tensorflow`, that are used to build models of artificial intelligence, including ChatGPT (Paszke et al., 2019; Abadi et al., 2016; Radford et al., 2019). Here, we outlay the most important packages used for `plotastic`.

Interactive Python - Jupyter

Python overcame the issues of interpreted language by utilizing Code written in C numpy:

- Acceleration, - SIMD instructions

Tabular operations - pandas

Data visualization - matplotlib - seaborn

Inferential Statistics - pingouin

AI: - pytorch and tensorflow - example: VGG19 is just a few lines of code (??) asdfdf

Aims

This project defines these aims:

- Characterise the interaction between myeloma cells and mesenchymal stromal cells
- Aim 2
- Aim 3

Chapter 1: Modelling Myeloma Dissemination *in vitro*

Abstract

lorem ipsum

Introduction

lorem ipsum dolor sit amet

Methods

lorem ipsum

Results

lorem ipsum

Discussion

lorem ipsum

1 Keep it Together: Modelling Myeloma Dissemination *in vitro* with hMSC-
2 Interacting Subpopulations of INA-6 Cells and their Aggregation/Detachment
3 Dynamics

4

5 Martin Kuric¹, Susanne Beck², Doris Schneider¹, Wyonna Rindt³, Marietheres Evers⁴, Jutta
6 Meißner-Weigl¹, Sabine Zeck¹, Melanie Krug¹, Marietta Herrmann⁵, Tanja Nicole Hartmann⁶,
7 Ellen Leich⁴, Maximilian Rudert⁷, Denitsa Docheva¹, Anja Seckinger⁸, Dirk Hose⁸, Franziska
8 Jundt³, Regina Ebert¹

9

10 ¹University of Würzburg, Department of Musculoskeletal Tissue Regeneration, Würzburg,
11 Germany

12 ²University Hospital Heidelberg, Institute of Pathology, Heidelberg, Germany

13 ³University Hospital Würzburg, Department of Internal Medicine II, Würzburg, Germany

14 ⁴University of Würzburg, Institute of Pathology, Comprehensive Cancer Center Mainfranken,
15 Würzburg, Germany

16 ⁵University Hospital Würzburg, IZKF Research Group Tissue Regeneration in
17 Musculoskeletal Diseases, Würzburg, Germany

18 ⁶University of Freiburg, Department of Internal Medicine I, Faculty of Medicine and Medical
19 Center, Freiburg, Germany

20 ⁷University of Würzburg, Orthopedic Department, Clinic König-Ludwig-Haus, Würzburg,
21 Germany

22 ⁸Vrije Universiteit Brussel, Department of Hematology and Immunology, Jette, Belgium

23

24 **Running Title**

25 Keep it Together: Modelling Myeloma Dissemination *in vitro*

26

27 **Keywords**

28 Multiple Myeloma, Human Mesenchymal Stromal Cells (hMSC), Cell Adhesion,
29 Dissemination, Survival

30

31 **Additional Information**

32 Corresponding Authors:

33 Martin Kuric, Department of Musculoskeletal Tissue Regeneration, Julius-Maximilians-
34 Universität Würzburg, Friedrich-Bergius-Ring 15, 97076 Würzburg, Germany, Phone +49
35 931 803 1583, martin.kur4@gmail.com

36 Prof. Dr. Regina Ebert, Department of Musculoskeletal Tissue Regeneration, Julius-
37 Maximilians-Universität Würzburg, Friedrich-Bergius-Ring 15, 97076 Würzburg, Germany,
38 Phone +49 931 803 1597, regina.ebert@uni-wuerzburg.de

39

40 **Conflict of Interest:**

41 The authors declare no potential conflicts of interest.

42

43 Abstract

44 Multiple myeloma involves early dissemination of malignant plasma cells across the bone
45 marrow; however, the initial steps of dissemination remain unclear. Human bone marrow-
46 derived mesenchymal stromal cells (hMSCs) stimulate myeloma cell expansion (e.g., IL-6)
47 and simultaneously retain myeloma cells via chemokines (e.g., CXCL12) and adhesion
48 factors. Hence, we hypothesized that the imbalance between cell division and retention
49 drives dissemination.

50 We present an *in vitro* model using primary hMSCs co-cultured with INA-6 myeloma cells.
51 Time-lapse microscopy revealed proliferation and attachment/detachment dynamics.
52 Separation techniques (V-well adhesion assay and well plate sandwich centrifugation) were
53 established to isolate MSC-interacting myeloma subpopulations that were characterized by
54 RNAseq, cell viability and apoptosis. Results were correlated with gene expression data
55 (n=837) and survival of myeloma patients (n=536).

56 On dispersed hMSCs, INA-6 saturate hMSC-surface before proliferating into large homotypic
57 aggregates, from which single cells detached completely. On confluent hMSCs, aggregates
58 were replaced by strong heterotypic hMSC-INA-6 interactions, which modulated apoptosis
59 time-dependently. Only INA-6 daughter cells (nMA-INA6) detached from hMSCs by cell
60 division but sustained adherence to hMSC-adhering mother cells (MA-INA6).

61 Isolated nMA-INA6 indicated hMSC-autonomy through superior viability after IL-6 withdrawal
62 and upregulation of proliferation-related genes. MA-INA6 upregulated adhesion and retention
63 factors (CXCL12), that, intriguingly, were highly expressed in myeloma samples from patients
64 with longer overall and progression-free survival, but their expression decreased in relapsed
65 myeloma samples.

66 Altogether, *in vitro* dissemination of INA-6 is driven by detaching daughter cells after a cycle
67 of hMSC-(re)attachment and proliferation, involving adhesion factors that represent a bone
68 marrow-retentive phenotype with potential clinical relevance.

69 **Statement of Significance**

70 Novel methods describe *in vitro* dissemination of myeloma cells as detachment of daughter
71 cells after cell division. Myeloma adhesion genes were identified that counteract *in vitro*
72 detachment with potential clinical relevance.

73

74 Introduction

75 Multiple myeloma arises from clonal expansion of malignant plasma cells in the bone marrow
76 (BM). At diagnosis, myeloma cells have disseminated to multiple sites in the skeleton and, in
77 some cases, to “virtually any tissue” (1, 2). However, the mechanism through which myeloma
78 cells initially disseminate remains unclear.

79 Dissemination is a multistep process involving invasion, intravasation, intravascular arrest,
80 extravasation, and colonization (3). To initiate dissemination, myeloma cells overcome
81 adhesion, retention, and dependency on the BM microenvironment, which could involve the
82 loss of adhesion factors such as CD138 (4, 5). BM retention is mediated by multiple factors:
83 First, chemokines (CXCL12 and CXCL8) produced by mesenchymal stromal cells (MSCs),
84 which attract plasma cells and prime their cytoskeleton and integrins for adhesion (6, 7).
85 Second, myeloma cells must overcome the anchorage and physical boundaries of the
86 extracellular matrix (ECM), consisting of e.g. fibronectin, collagens, and proteoglycans such
87 as decorin (8-11). Simultaneously, ECM provides signals inducing myeloma cell cycle arrest
88 or progression the cell cycle (8, 10). ECM is also prone to degradation, which is common in
89 several osteotropic cancers, and is the cause of osteolytic bone disease. This is driven by a
90 ‘vicious cycle’ that maximizes bone destruction by extracting growth factors (EGF and TGF-
91 β) that are stored in calcified tissues (12). Third, direct contact with MSCs physically anchors
92 myeloma cells to the BM (3, 13). Fourth, to disseminate to distant sites, myeloma cells
93 require, at least partially, independence from essential growth and survival signals provided
94 by MSCs in the form of soluble factors or cell adhesion signaling (5, 14, 15). For example,
95 the VLA4 (Myeloma)-VCAM1 (MSC)-interface activates NF-κB in both myeloma and MSCs,
96 inducing IL-6 expression in MSCs. The independence from MSCs is then acquired through
97 autocrine survival signaling (16, 17). In short, anchorage of myeloma cells to MSCs or ECM
98 is a ‘double-edged sword’: adhesion counteracts dissemination, but also presents signaling
99 cues for growth, survival, and drug resistance (18).

100 To address this ambiguity, we developed an *in vitro* co-culture system modeling diverse
101 adhesion modalities to study dissemination, growth, and survival of myeloma cells and

102 hMSCs. Co-cultures of hMSCs and the myeloma cell line INA-6 replicated tight interactions
103 and aggregate growth, akin to "micrometastases" in Ghobrial's metastasis concept (19). We
104 characterized the growth conformations of hMSCs and INA-6 as homotypic aggregation vs.
105 heterotypic hMSC adherence and their effects on myeloma cell survival. We tracked INA-6
106 detachments from aggregates and hMSCs, thereby identifying a potential "disseminated"
107 subpopulation lacking strong adhesion. We developed innovative techniques (V-well
108 adhesion assay and well plate sandwich centrifugation) to separate weakly and strongly
109 adherent subpopulations for the subsequent analysis of differential gene expression and cell
110 survival. Notably, our strategy resolves the differences in gene expression and growth
111 behavior between cells of one cell population in "direct" contact with MSCs. In contrast,
112 previous methods differentiated between "direct" and "indirect" cell-cell contact using
113 transwell inserts (20). To evaluate whether genes mediating adhesion and growth
114 characteristics of INA-6 were associated with patient survival, we analyzed publicly available
115 datasets (21, 22).

116 Materials and Methods

117 See supplemental for a complete method list and description.

118 Ethics Statement

119 Primary human MSCs were collected with informed consent from all patients and the
120 procedure was approved by the local Ethics Committee of the University of Würzburg
121 (186/18).

122 Cultivation and Co-Culturing of primary hMSCs and INA-6

123 Cell isolation, cultivation, and medium composition are provided in the supplemental
124 materials and methods section. Primary human MSCs (hMSC) were obtained from 34 non-
125 myeloma patients undergoing elective hip arthroplasty (Tab. S1: 21 male and 13 female,
126 mean age 68.9 ± 10.6). The INA-6 cell line (DSMZ Cat# ACC-862, RRID:CVCL_5209, [link](#))
127 was initially isolated from a pleural effusion sample obtained from an 80-year-old male with
128 multiple myeloma (23, 24). hMSCs were not tested for mycoplasma, whereas stocks of INA-6

129 were tested in this study (Tab. S1) using the Venor GeM OneStep kit (Minerva Biolabs,
130 Berlin, Germany).

131 For each co-culture, hMSCs were seeded 24 h before INA-6 addition to generate the MSC-
132 conditioned medium (CM). INA-6 cells were washed with PBS, resuspended in MSC
133 medium, and added to hMSCs so that the co-culture comprised 33% (v/v) of CM gathered
134 directly from the respective hMSC donor. The co-cultures were not substituted for IL-6 ([14](#)).

135 **Cell Viability and Apoptosis Assay**

136 Cell viability and apoptosis rates were measured using CellTiter-Glo Luminescent Cell
137 Viability Assay and Caspase-Glo 3/7 Assay, respectively (Promega GmbH, Mannheim,
138 Germany).

139 **Automated Fluorescence Microscopy**

140 Microscopic images were acquired using an Axio Observer 7 (Zeiss) with a COLIBRI LED
141 light source and motorized stage top using 5x and 10x magnification. The tiled images had
142 an automatic 8-10% overlap and were not stitched.

143 **Live Cell Imaging**

144 hMSCs (stained with PKH26) were placed into an ibidi Stage Top Incubation System and
145 equilibrated to 80% humidity and 5% CO₂. INA-6 (2×10^3 cells/cm²) were added directly
146 before the start of acquisition. Brightfield and fluorescence images of up to 13 mm² of the co-
147 culture area were acquired every 15 min for 63 h. Each event of interest was manually
148 analyzed and categorized into defined event parameters.

149 **V-well Adhesion Assay**

150 INA-6 cells were arrested during mitosis by two treatments with thymidine, followed by
151 nocodazole. Arrested INA-6 were released and added to 96 V-well plates (10^4 cells/cm²) on
152 top of confluent hMSCs and adhered for 1-3 h. The co-culture was stained with calcein-AM
153 (Thermo Fisher Scientific, Darmstadt, Germany) before non-adherent INA-6 were pelleted
154 into the tip of the V-well (2000 rpm, 5-10 min). MSC-adhering INA-6 cells were manually
155 detached by rapid pipetting. The pellet brightness was measured microscopically and the
156 pellet was isolated by pipetting.

157 **Cell Cycle Profiling by Image Cytometry**

158 Isolated INA-6 cells were fixed in 70% ice-cold ethanol, washed, resuspended in PBS,
159 distributed in 96-well plates, and stained with Hoechst-33342. The plates were scanned at 5x
160 magnification. A pre-trained convolutional neural network (Intellessis, Zeiss) was fine-tuned to
161 segment the scans into single nuclei and exclude fragmented nuclei. Nuclei were filtered to
162 exclude extremes of size roundness. The G0/G1 frequency was determined by Gaussian
163 curve fitting.

164 **Well Plate Sandwich Centrifugation (WPSC)**

165 hMSCs were grown to confluence in 96-well plates coated with collagen I (rat tail; Corning,
166 NY, USA). INA-6 were added and the cells were allowed to adhere for 24 h. A second plate
167 ("catching plate") was attached upside down to the top of the co-culture plate. That "well
168 plate sandwich" was turned around and the content of the co-culture plate was centrifuged
169 into the catching plate three times (40 seconds at 110 g) while gently adding 30 µL of
170 medium in between centrifugation steps. Non-MSC-adhering INA-6 cells were collected from
171 the catching plate, whereas MSC-adhering INA-6 cells were isolated by digesting the co-
172 culture with Accutase. For RNA sequencing (RNAseq), all samples were purified using anti-
173 CD45 magnetic-assisted cell sorting (Miltenyi Biotec B.V. & Co. KG, Bergisch Gladbach).

174 **RNA Isolation**

175 RNA was isolated using the NucleoSpin RNA II Purification Kit (Macherey-Nagel) according
176 to the manufacturer's instructions. RNA was isolated from INA-6 cells co-cultured with a
177 unique hMSC donor (n=5 for qPCR n=11).

178 **RNA sequencing, Differential Expression, and Functional Enrichment Analysis**

179 RNAseq was performed at the Core Unit Systems Medicine, University of Würzburg. mRNA
180 was enriched with polyA beads. Fastq files were aligned to the GRCh38 reference genome
181 using STAR (RRID:SCR_004463, [link](#)) and raw read counts were generated using HTseq
182 ([25-27](#)). Differential gene expression was analyzed using edgeR in R (version 3.6.3).
183 Functional enrichment analysis was performed using Metascape ([28](#)).

184 **RT-qPCR**

185 RNA (1 µg) was reverse transcribed using SuperScript IV reverse transcriptase (Thermo
186 Fisher Scientific, Darmstadt, Germany). qPCR was performed using 10 µL GoTaq qPCR
187 Master Mix (Promega), 1:10 diluted cDNA, and 5 pmol of primers obtained from Biomers.net
188 or Qiagen (Tab. S3).

189 **Statistics**

190 Inferential statistics were performed using Python (IPython, RRID:SCR_001658, [link](#)) (3.10)
191 packages pingouin (0.5.1) and statsmodels (0.14.0) ([29](#), [30](#)). The figures were plotted using
192 plotastic (0.0.1) ([31](#)). Normality (for $n \geq 4$) and sphericity were ensured using Mauchly's and
193 Shapiro-Wilk tests, respectively. Data points were log10 transformed to convert the scale
194 from multiplicative to additive or to fulfill sphericity requirements. p-value = 0.05 > * > 0.01 >
195 ** > 10^{-3} > *** 10^{-4} > ****. P-values were either adjusted (p-adj) or not adjusted (p-unc) for
196 family wise error rate. Power calculations were not performed to determine the sample size.

197 **Patient Cohort, Analysis of Survival and Expression**

198 Survival and gene expression data were obtained as previously described ([21](#), [22](#)) and are
199 available at the European Nucleotide Archive (ENA) under accession numbers PRJEB36223
200 and PRJEB37100. The expression level was categorized into "high" and "low" using maxstat
201 (Maximally selected Rank Statistics) thresholds ([32](#)).

202 **Data Availability Statement**

203 A detailed description of the methods is provided in the Supplementary Material section. Raw
204 tabular data and examples of analyses and videos are available in the github repository ([33](#)).
205 Raw RNAseq data are available from the NCBI Gene Expression Omnibus (GEO)
206 (RRID:SCR_005012, link) (GSExxxx). Microscopy data are available at the Image Data
207 Resource (IDR) (RRID:SCR_017421, link) (idrxxxx).

208 **Results**

209 **INA-6 Cells Saturate hMSC-Interaction to Proliferate into Aggregates**

210 hMSCs are isolated as a heterogeneous cell population. To analyze whether INA-6 cells
211 could adhere to every hMSC, we saturated hMSCs with INA-6. A seeding ratio of 1:4

212 (hMSC:INA-6) resulted in the occupation of $93\% \pm 6\%$ of single hMSCs by one or more INA-
213 6 cells within 24 h after INA-6 addition, escalating to 98% after 48 h (Fig. 1A, B). Therefore,
214 most hMSCs provide an interacting surface for INA-6 cells.

215 INA-6 exhibits homotypic aggregation when cultured alone, a phenomenon observed in
216 some freshly isolated myeloma samples (up to 100 cells after 6 h) (34, 35). Adding hMSCs at
217 a 1:1 ratio led to smaller aggregates after 24 h (size 1-5 cells), all of which were distributed
218 over $52 \pm 2\%$ of all hMSCs (Fig. 1A, B). Intriguingly, INA-6 aggregation was notably absent
219 when grown on confluent hMSCs, and occurred only when heterotypic interactions were
220 limited to 0.2 hMSCs per INA-6 cell (Fig. 1C). We concluded that INA-6 cells prioritize
221 heterotypic over homotypic interactions.

222 To monitor the formation of such aggregates, we conducted live-cell imaging of hMSC/INA-6
223 co-cultures for 63 h. We observed that INA-6 cells adhered long after cytokinesis,
224 constituting $55 \pm 12\%$ of all homotypic interactions between 13 and 26 h, increasing to >75%
225 for the remainder of the co-culture (Fig. 1D). Therefore, homotypic INA-6 aggregates were
226 mostly formed by cell division.

227 **Apoptosis of INA-6 Depends on Ratio Between Heterotypic and Homotypic Interaction**

228 Although direct interaction with hMSCs has been shown to enhance myeloma cell survival
229 through NF- κ B signaling (15), the impact of aggregation on myeloma cell viability during
230 hMSC interaction remains unclear. To address this, we measured the cell viability (ATP) and
231 apoptosis rates of INA-6 cells growing as homotypic aggregates compared to those in
232 heterotypic interactions with hMSCs by modulating hMSC density (Fig. 1E). To equalize the
233 background signaling caused by soluble MSC-derived factors, all cultures were incubated in
234 hMSC-conditioned medium and the results were normalized to INA-6 cells cultured without
235 direct hMSC contact (Fig. 1E, left).

236 INA-6 viability (ATP) was not affected by the direct adhesion of hMSCs at any density.
237 However, apoptosis rates decreased over time [$F(2, 6) = 23.29$, $p\text{-unc} = 1.49e-03$, Two-factor
238 RM-ANOVA], interacting significantly with MSC density [$F(4, 12) = 6.98$, $p\text{-unc} = 3.83e-3$].
239 For example, 24 h of adhesion to confluent MSCs increased apoptosis rates by 1.46 ± 0.37

240 fold, while culturing INA-6 cells on dispersed hMSCs (ratio 1:1) did not change the apoptosis
241 rate (1.01 ± 0.26).

242 We presumed that sensitive apoptotic cells might have been lost when harvesting INA-6 cells
243 from hMSCs. Hence, we measured survival parameters in the co-culture and in hMSC and
244 INA-6 cells cultured separately (Fig. 1E, right). We defined MSC interaction effects when the
245 survival measured in the co-culture differed from the sum of the signals measured from INA-6
246 and hMSCs alone. RM-ANOVA confirmed that adherence to confluent MSCs increased
247 apoptosis rates of INA-6 cells 24 h after adhesion and decreased after 72 h [interaction
248 between MSC density and time: $F(2, 4) = 26.86$, $p\text{-unc} = 4.80\text{e-}03$, Two-factor RM-ANOVA],
249 whereas INA-6 cells were unaffected when grown on dispersed hMSCs.

250 In summary, the growth conformation of INA-6 cells, measured as the ratio between
251 homotypic aggregation and heterotypic MSC interactions, affected apoptosis rates of INA-6
252 cells.

253 **Single INA-6 Cells Detach Spontaneously from Aggregates of Critical Size**

254 Using time-lapse microscopy, we observed that $26 \pm 8\%$ of INA-6 aggregates growing on
255 single hMSCs spontaneously shed INA-6 cells (Fig. 2A, B; Vid. S1). Notably, all detached
256 cells exhibited similar directional movements, suggesting entrainment in convective streams
257 generated by temperature gradients within the incubation chamber. INA-6 predominantly
258 detached from other INA-6 cells or aggregates (Fig. 2C), indicating weaker adhesive forces
259 in homotypic interactions than in heterotypic interactions. The detachment frequency
260 increased after 52 h, when most aggregates that shed INA-6 cells were categorized as large
261 (> 30 cells) (Fig. 2D). Since ~ 1020 INA6 cells already fully covered a single hMSC, we
262 suggest that myeloma cell detachment depended not only on hMSC saturation, but also
263 required a minimum aggregate size. Interestingly, INA-6 detached mostly as single cells,
264 independent of aggregate size categories [$F(2, 6)=4.68$, $p\text{-unc}=.059$, Two-factor RM-
265 ANOVA] (Fig. 2D), showing that aggregates remained mostly stable despite losing cells.

266 Cell Division Generates a Daughter Cell Detached from hMSC

267 We suspected that cell division drives detachment because we observed that MSC-adhering
268 INA-6 cells could generate daughter cells that “roll over” the mother cell (Fig. 3A; Vid. S2).
269 We recorded and categorized the movement of INA-6 daughter cells in confluent hMSCs
270 after cell division. Half of all INA-6 divisions yielded two daughter cells that remained
271 stationary, indicating hMSC adherence (Fig. 3B, C; Vid. S3). The other half of division
272 events generated one hMSC-adhering (MA) cell and one non-hMSC-adhering (nMA) cell,
273 which rolled around the MA cell for a median time of 2.5 h post division ($Q_1=1.00$ h, $Q_3=6.25$
274 h) until it stopped and re-adhered to the hMSC monolayer (Fig. 3D; Vid. S2, S4). Thus, cell
275 division establishes a time window in which one daughter cell can detach.

276 To validate that cell division reduced adhesion, we measured both the size and cell cycle
277 profile of the nMA and MA populations using an enhanced V-well assay (method described in
278 Fig. 3E, S1, S2). For comparison, we fully synchronized and arrested INA-6 cells at mitosis
279 and released their cell cycle immediately before addition to the hMSC monolayer, rendering
280 them more likely to divide while adhering. Mitotic arrest significantly increased the number of
281 nMA cells and decreased the number of MA cells (Fig. 3F). Furthermore, the nMA population
282 contained significantly more cells cycling in the G₀/G₁ phase than the MA population, both in
283 synchronously and asynchronously cycling INA-6 (Fig. 3G, S3B). The number of nMA INA-6
284 cells increased due to a higher cell division frequency. Taken together, we showed that INA-
285 6 detach from aggregates by generating one temporarily detached daughter cell after cell
286 division, a process that potentially contributes to the initiation of dissemination.

**287 Well Plate Sandwich Centrifugation (WPSC) Separates hMSC-Interacting INA-6
288 Subpopulations**

289 To separate nMA and MA cells for further analysis, we developed the method “Well Plate
290 Sandwich Centrifugation” (WPSC), outlined in Fig. 4A. To equalize the background signaling
291 caused by MSC-derived factors and to focus on differences within directly MSC-interacting
292 INA-6 subpopulations, all cultures were incubated in hMSC-conditioned medium (CM) from
293 the respective donors and compared with INA-6 incubation in CM without hMSCs.

294 Microscopic tracking of nMA and MA INA-6 cell numbers during each WPSC separation step
295 revealed successful separation after the third centrifugation step, whereas CM-treated INA-6
296 cells required only one centrifugation step (Fig. 4B). Thus, WPSC generated cell numbers
297 that were suitable for subsequent analyses.

298 **RNAseq of non-MSC-Adhering and MSC-Adhering Subpopulations**

299 To characterize the subpopulations separated by WPSC, we conducted RNAseq, revealing
300 1291 differentially expressed genes between nMA vs. CM, 484 between MA vs. CM, and 195
301 between MA vs. nMA. We validated RNAseq and found that the differential expression of 18
302 genes correlated with those measured with qPCR for each pairwise comparison (Fig. 4C-E,
303 S5): nMA vs. CM [$\rho(16) = .803, p = 6.09e-5$], MA vs. CM [$\rho(16) = .827, p = 2.30e-5$], and MA
304 vs. nMA cells [$\rho(16) = .746, p = 3.74e-4$] (Spearman's rank correlation). One of the 18 genes
305 (*MUC1*) measured by qPCR showed a mean expression opposite to that obtained by
306 RNAseq (nMA vs. CM), although the difference was insignificant (Fig. 4C). For nMA vs. CM,
307 the difference in expression measured by qPCR was significant for only two of the 11 genes
308 (*DKK1* and *OPG*), whereas the other genes (*BCL6*, *BMP4*, *BTG2*, *IL10RB*, *IL24*, *NOTCH2*,
309 *TNFRSF1A*, *TRAF5*) only confirmed the tendency measured by RNAseq (Fig. 4C-E). For MA
310 vs. CM, qPCR validated the significant upregulation of seven genes (*TGM2*, *DCN*, *LOX*,
311 *MMP14*, *MMP2*, *CXCL12*, *CXCL8*), whereas the downregulation of *BMP4* was insignificant.

312 **Non-MSC-Adhering INA-6 and MSC-Adhering INA-6 Have Distinct Expression Patterns
313 of Proliferation or Adhesion, Respectively**

314 To functionally characterize the unique transcriptional patterns in nMA-INA6 and MA-INA6,
315 we generated lists of genes that were differentially expressed vs. the other two
316 subpopulations [termed nMA vs. (MA & CM) and MA vs. (nMA & CM)]. Functional enrichment
317 analysis was performed, and the enriched terms were displayed as ontology clusters (Fig.
318 5A). nMA-INA6 upregulated genes enriched with loosely connected term clusters associated
319 with proliferation (e.g., "positive regulation of cell cycle"). MA-INA6 upregulated genes
320 enriched with tightly connected term clusters related to cell adhesion and the production of
321 ECM factors (e.g. "cell-substrate adhesion"). Similar ontology terms were enriched in the

322 gene lists obtained from pairwise comparisons (nMA vs. CM, MA vs. CM, and MA vs. nMA)
323 (Fig. 5B). In particular, nMA vs. CM (but not MA vs. CM) upregulated genes that were
324 enriched with “G1/S transition”, showing that WPSC isolated nMA daughter cells after cell
325 division.

326 To check for similarities between lists of differentially expressed genes from hMSC-
327 interacting subpopulations, we performed enrichment analysis on gene lists from the
328 overlaps (“ \cap ”) between all pairwise comparisons (Fig. 5B, S6), and showed the extent of
329 these overlaps in circos plots (Fig. 5C). The overlap between MA vs. CM and nMA vs. CM
330 showed neither enrichment with proliferation- nor adhesion-related terms but with apoptosis-
331 related terms. A direct comparison of MSC-interacting subpopulations (MA vs. nMA) showed
332 a major overlap with MA vs. CM (Fig. 5C, middle). This overlap was enriched with terms
333 related to adhesion but not proliferation. Hence, MA-INA6 and nMA-INA6 mostly differed in
334 their expression of adhesion genes.

335 To assess whether nMA-INA6 and MA-INA6 were regulated by separate transcription factors,
336 we examined the enrichment of curated regulatory networks from the TRRUST database
337 (Fig. 5B bottom). All the lists were enriched for p53 regulation. E2F1 regulation was observed
338 only in genes upregulated in nMA vs. CM and downregulated in MA vs. nMA. Genelists
339 involving MA-INA6 were enriched in regulation by subunits of NF- κ B (NFKB1/p105 and
340 RELA/p65) and factors of immediate early response (SRF, JUN). Correspondingly, NF- κ B
341 and JUN are known to regulate the expression of adhesion factors in multiple myeloma and
342 B-cell lymphoma, respectively (36, 37).

343 Taken together, MSC-interacting subpopulations showed unique regulatory patterns,
344 focusing on either proliferation or adhesion.

345 **nMA-INA6 and MA-INA6 Show Increased Apoptosis Signaling Mediated by ER-Stress,
346 p53 and Death Domain Receptors**

347 As previously stated, apoptosis rates increased in INA-6 cells grown on confluent hMSCs
348 compared to CM-INA6 cells after 24 h of co-culture (Fig. 1D). Since this setup was similar to
349 that used to separate hMSC-interacting subpopulations using WPSC, we looked for

350 enrichment of apoptosis-related terms (Fig. 5B). “Regulation of cellular response to stress”
351 and “intrinsic apoptotic signaling pathway (in response to ER-stress)” are terms that were
352 enriched in nMA vs. CM, MA vs. CM and their overlap. We also found specific stressors for
353 either nMA-INA6 (“intrinsic apoptotic signaling pathway by p53 class mediator”) or MA-INA6
354 (“extrinsic apoptotic signaling pathway via death domain receptor”). Therefore, apoptosis
355 may be driven by ER stress in both nMA-INA6 and MA-INA6, but also by individual pathways
356 such as p53 and death domain receptors, respectively.

357 **nMA-INA6 and MA-INA6 Regulate Genes Associated with Bone Loss**

358 Myeloma cells cause bone loss by degradation and dysregulation of bone turnover via DKK1
359 and OPG ([38-40](#)). RNAseq of hMSC-interacting subpopulations showed enrichment with
360 functional terms “skeletal system development” and “ossification” (Fig. 5A, S6), as well as
361 the regulation of *MMP2*, *MMP14*, *DKK1*, and *OPG*. Validation by qPCR (Fig. 4C, D) showed
362 that MA-INA6 significantly upregulated both *MMP14* and *MMP2* compared with either nMA-
363 INA6 or CM-INA6. The expression of *DKK1*, however, was upregulated significantly in nMA-
364 INA6 (and not significantly upregulated in MA-INA6), while *OPG* was significantly
365 downregulated only in nMA-INA6.

366 Together, hMSC-interacting subpopulations might contribute to bone loss through different
367 mechanisms: MA-INA6 expression of matrix metalloproteinases and nMA cells via paracrine
368 signaling.

369 **MA-INA6 Upregulate Collagen and Chemokines Associated with Bone Marrow
370 Retention**

371 Retention of myeloma cells within the bone marrow is mediated by adhesion to the ECM
372 (e.g., collagen VI) and the secretion of chemokines (CXCL8 and CXCL12) ([7](#), [11](#)). This
373 directly counteracts dissemination, which is a hallmark of MA-INA6. RNAseq of hMSC-
374 interacting subpopulations showed that genes upregulated in MA-INA6 were enriched with
375 collagen biosynthesis and modifying enzymes, as well as chemotaxis and chemotaxis-
376 related terms (Fig. 5B). Using qPCR, we validated the upregulation of collagen crosslinkers
377 (*LOX* and *TGM2*), collagen-binding *DCN* and chemokines (CXCL8 and CXCL12) in MA-INA6

378 compared with both nMA-INA6 and CM-INA6 (Fig. 4D). Therefore, MA-INA6 can provide
379 both an adhesive surface and soluble signals for the retention of malignant plasma cells in
380 the bone marrow.

381 **nMA-INA6 Show Highest Viability During IL-6 Withdrawal**

382 Although RNAseq did not reveal IL-6 induction in any WPSC-isolated subpopulation, nMA-
383 INA6 upregulated IGF-1 135%-fold [RNAseq, nMA vs. (MA & CM)], which was shown to
384 stimulate growth in CD45+ and IL-6 dependent myeloma cell lines such as INA-6, implying
385 increased autonomy for nMA-INA6 (41).

386 To test the autonomy of hMSC-interacting INA-6 subpopulations, we isolated them using
387 WPSC after 24 h and 48 h of co-culture, sub-cultured them for 48 h under IL-6 withdrawal,
388 and measured both viability and apoptosis (Fig. 5D). Among the subpopulations, nMA-INA6
389 was the most viable. Compared to MA-INA6, nMA-INA6 increased cell viability by 8 or 4 fold
390 when co-cultured for 24 or 48 h, respectively [Hedges g of $\text{Log}_{10}(\text{Fold Change})$ = 2.31 or
391 0.82]. However, the difference was no longer significant after 48 h of co-culture, probably
392 because nMA-INA6 adhered to the hMSC layer (turning into MA-INA6) during prolonged co-
393 culture, which could also explain why the viability of MA-INA6 cell subcultures increased with
394 prolonged co-culture. Nevertheless, nMA-INA6 did not achieve the same viability as that of
395 INA-6 cells cultured with IL-6. Despite the differences in viability, subcultures of hMSC-
396 interacting subpopulations did not show any differences in caspase 3/7 activity when co-
397 cultured for 48 h (Fig. 5D, right).

398 Overall, among the hMSC-interacting subpopulations, nMA-INA6 had the highest chance of
399 surviving IL-6 withdrawal.

400 **Genes Upregulated by MA-INA6 are Associated with an Improved Disease Prognosis**

401 To relate the adhesion of MA-INA6 observed *in vitro* to the progression of multiple myeloma,
402 we assessed patient survival [$n = 535$, Seckinger et al. 2018 (21, 22) depending on the
403 expression level of 101 genes, which were upregulated in MA vs. (nMA & CM) and are part
404 of the ontology terms “Extracellular matrix organization,” “ECM proteoglycans,” “cell-
405 substrate adhesion” and “negative regulation of cell-substrate adhesion” (Fig. 6A, Tab. S2).

406 As a reference, we generated a list of 173 cell cycle-related genes that were upregulated by
407 nMA-INA6 vs. (MA-INA6 & CM-INA6).
408 As expected, longer patient survival was associated with low expression of the majority of
409 cell cycle genes [71 or 68 genes for progression-free survival (PFS) or overall survival (OS)].
410 Only a few cell cycle genes (two for PFS and seven for OS) were associated with survival
411 when highly expressed. Intriguingly, adhesion genes showed an inverse pattern: a large
412 group of adhesion genes (24 for PFS and 26 for OS) was significantly associated with
413 improved survival when highly expressed, whereas only a few genes (two for PFS and four
414 for OS) improved survival when expressed at low levels (Tab. 1). We concluded that the
415 myeloma-dependent expression of adhesion factors determined in our *in vitro* study
416 correlates with improved patient survival.

417 **Expression of Adhesion- or Retention-related genes (CXCL12, DCN and TGM2) is
418 Decreased During Progression of Multiple Myeloma**

419 To examine how the disease stage affects the adhesion and bone marrow retention of
420 myeloma cells *in vitro*, we analyzed the expression of CXCL12 in healthy plasma cell
421 (BMPC) cohorts of patients at different disease stages and in myeloma cell lines (HMCL)
422 [described in Seckinger et al. 2018 (29)] (Fig. 6C). We also included DCN and TGM2 since
423 both are suggested to inhibit metastasis in different cancers by promoting cell-matrix
424 interactions (8, 42). In accordance with independent reports (9, 43), high expression of
425 CXCL12 and DCN by myeloma cells was associated with improved overall survival (adj. p =
426 .009 and .008, respectively) (Fig. 6B).

427 CXCL12 is expressed by BMPCs (median = 219 normalized counts), but its expression
428 levels are significantly lower from MGUS to relapsed multiple myeloma (MMR) (median = 9
429 normalized counts in MMR and absent expression in most HMCL). DCN (but not TGM2) was
430 weakly expressed in BMPCs ($Q_1=0.7$, $Q_3=3.7$, normalized counts), whereas TGM2 was
431 weakly expressed only in patients with monoclonal gammopathy of undetermined
432 significance (MGUS) ($Q_1=0.4$, $Q_3=4.1$ normalized counts). The median and upper quartiles of
433 both DCN- and TGM2 decreased continuously after each stage, ending at $Q_3=0.9$ and

434 $Q_3=0.6$, respectively, in MMR. 49 of the 101 adhesion genes (Fig. 6A) followed a similar
435 pattern of continuous downregulation in the advanced stages of multiple myeloma (Fig. S7
436 and S8), of which 19 genes were associated with longer PFS when they were highly
437 expressed. The other 52 (out of 101) adhesion genes that were not downregulated across
438 disease progression (or were expressed at a level too low to make that categorization)
439 contained only five genes that were associated with longer PFS at high expression (Tab. 1,
440 Tab. S2).

441 Together, the expression of adhesion or bone marrow retention-related markers (*CXCL12*,
442 *DCN*, and *TGM2*) is reduced or lost at advanced stages of multiple myeloma, which could
443 enhance dissemination and reduce retention in the BM microenvironment.

444 Discussion

445 In this study, we developed an *in vitro* model to investigate the attachment/detachment
446 dynamics of INA-6 cells to/from hMSCs and established methods to isolate the attached and
447 detached intermediates nMA-INA6 and MA-INA6. Second, we characterized a cycle of
448 (re)attachment, division, and detachment, linking cell division to the switch that causes
449 myeloma cells to detach from hMSC adhesion (Fig. 7). Thirdly, we identified clinically
450 relevant genes associated with patient survival, in which better or worse survival was based
451 on the adherence status of INA-6 to hMSCs.

452 INA-6 cells emerged as a robust choice for studying myeloma dissemination *in vitro*, showing
453 rapid and strong adherence, as well as aggregation exceeding MSC saturation. The IL-6
454 dependency of INA-6 enhanced the resemblance of myeloma cell lines to patient samples,
455 with INA-6 ranking 13th among 66 cell lines (44). Despite variations in bone marrow MSCs
456 between multiple myeloma (MM) and healthy states, we anticipated the robustness of our
457 results, given the persistent strong adherence and growth signaling from MSCs to INA-6
458 during co-cultures (45).

459 We acknowledge that INA-6 cells alone cannot fully represent the complexity of myeloma
460 aggregation and detachment dynamics. However, the diverse adhesive properties of
461 myeloma cell lines pose a challenge. We reasoned that attempting to capture this complexity

462 within a single publication would not be possible. Our focus on INA-6 interactions with
463 hMSCs allowed for a detailed exploration of the observed phenomena, such as the unique
464 aggregation capabilities that facilitate the easy detection of detaching cells *in vitro*. The
465 validity of our data was demonstrated by matching the *in vitro* findings with the gene
466 expression and survival data of the patients (e.g. CXCL12, DCN, and TGM2 expression,
467 n=873), ensuring biological consistency and generalizability regardless of the cell line used.
468 The protocols presented in this study offer a cost-efficient and convenient solution, making
469 them potentially valuable for a broader study of cell interactions. We encourage optimizations
470 to meet the varied adhesive properties of the samples, such as decreasing the number of
471 washing steps if the adhesive strength is low. We caution against strategies that average
472 over multiple cell lines without prior understanding their diverse attachment/detachment
473 dynamics, such as homotypic aggregation. Such detailed insights may prove instrumental
474 when considering the diversity of myeloma patient samples across different disease stages
475 ([34](#), [35](#)).

476 The intermediates, nMA-INA6 and MA-INA6, were distinct but shared similarities in response
477 to cell stress, intrinsic apoptosis, and regulation by p53. Unique regulatory patterns were
478 related to central transcription factors: E2F1 for nMA-INA6; and NF- κ B, SRF, and JUN for
479 MA-INA6. This distinction may have been established through antagonism between p53 and
480 the NF- κ B subunit RELA/p65 ([46](#), [47](#)). Similar regulatory patterns were found in transwell
481 experiments with RPMI1-8226 myeloma cells, where direct contact with the MSC cell line
482 HS5 led to NF- κ B signaling and soluble factors to E2F signaling ([20](#)).

483 The first subpopulation, nMA-INA6, represented proliferative and disseminative cells; nMA-
484 INA6 drove detachment through cell division, which was regulated by E2F, p53, and likely
485 their crosstalk ([48](#)). They upregulate cell cycle progression genes associated with worse
486 prognosis, because proliferation is a general risk factor for an aggressive disease course
487 ([49](#)). Additionally, nMA-INA6 survived IL-6 withdrawal better than CM-INA6 and MA-INA6,
488 implying their ability to proliferate independently of the bone marrow ([1](#)). Indeed, xenografted
489 INA-6 cells developed autocrine IL-6 signaling but remained IL-6-dependent after

490 explantation (23). The increased autonomy of nMA-INA6 cells can be explained by the
491 upregulation of IGF-1, being the major growth factor for myeloma cell lines (41). Other
492 reports characterized disseminating cells differently: Unlike nMA-INA6, circulating myeloma
493 tumor cells were reported to be non-proliferative and bone marrow retentive (50). In contrast
494 to circulating myeloma tumor cells, nMA-INA6 were isolated shortly after detachment and
495 therefore these cells are not representative of further steps of dissemination, such as
496 intravasation, circulation or intravascular arrest (3). Furthermore, Brandl et al. described
497 proliferative and disseminative myeloma cells as separate entities, depending on the surface
498 expression of CD138 or JAM-C (4, 51). Although CD138 was not differentially regulated in
499 nMA-INA6 or MA-INA6, both subpopulations upregulated JAM-C, indicating disease
500 progression (51).

501 Furthermore, nMA-INA6 showed that cell division directly contributed to dissemination. This
502 was because INA-6 daughter cells emerged from the mother cell with distance to the hMSC
503 plane in the 2D setup. A similar mechanism was described in an intravasation model in which
504 tumor cells disrupt the vessel endothelium through cell division and detach into blood
505 circulation (52). Overall, cell division offers key mechanistic insights into dissemination and
506 metastasis.

507 The other subpopulation, MA-INA6, represented cells retained in the bone marrow; MA-INA6
508 strongly adhered to MSCs, showed NF- κ B signaling, and upregulated several retention,
509 adhesion, and ECM factors. The production of ECM-associated factors has recently been
510 described in MM.1S and RPMI-8226 myeloma cells (53). Another report did not identify the
511 upregulation of such factors after direct contact with the MSC cell line HS5; hence, primary
512 hMSCs may be crucial for studying myeloma-MSC interactions (20). Moreover, MA-INA6
513 upregulated adhesion genes associated with prolonged patient survival and showed
514 decreased expression in relapsed myeloma. As myeloma progression implies the
515 independence of myeloma cells from the bone marrow (1, 44), we interpreted these adhesion
516 genes as mediators of bone marrow retention, decreasing the risk for dissemination and
517 thereby potentially prolonging patient survival. However, the overall impact of cell adhesion

518 and ECM on patient survival remains unclear. Several adhesion factors have been proposed
519 as potential therapeutic targets (51, 54). Recent studies have described the prognostic value
520 of multiple ECM genes, such as those driven by NOTCH (53). Another study focused on
521 ECM gene families, of which only six of the 26 genes overlapped with our gene set (Tab. S2)
522 (55). The expression of only one gene (*COL4A1*) showed a different association with overall
523 survival than that in our cohort. The lack of overlap and differences can be explained by
524 dissimilar definitions of gene sets (homology vs. gene ontology), methodological
525 discrepancies, and cohort composition.

526 In summary, our *in vitro* model provides a starting point for understanding the initiation of
527 dissemination and its implications for patient survival, providing innovative methods,
528 mechanistic insights into attachment/detachment, and a set of clinically relevant genes that
529 play a role in bone marrow retention. These results and methods might prove useful when
530 facing the heterogeneity of disseminative behaviors among myeloma cell lines and primary
531 materials.

532 Acknowledgements

533 This work was supported by the Deutsche Forschungsgemeinschaft (DFG) SPP microBONE
534 grants EB 447/10-1 (491715122), JA 504/17-1, HO 4462/1-1 (401358321), JU 426/10-1
535 (491715122), and JU 426/11-1 (496963451). MH was funded by the IZKF project D-361, and
536 EL was funded by Deutsche Krebshilfe (70112693) and Wilhelm Sander-Stiftung
537 (2014.903.1). We thank the Core Unit Systemmedizin (SysMed), c/o Institut für Molekulare
538 Infektionsbiologie (IMIB), University of Würzburg for performing the RNAseq analyses, as
539 well as the Elite Netzwerk Bayern and the Graduate School of Life Sciences of the University
540 of Würzburg.

541

542 References

- 543 1. Bladé J, Beksać M, Caers J, Jurczyszyn A, von Lilienfeld-Toal M, Moreau P, et al.
544 Extramedullary disease in multiple myeloma: a systematic literature review. *Blood Cancer
545 Journal.* 2022;12(3):1-10.
- 546 2. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos M-V, et al.
547 International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma.
548 *The Lancet Oncology.* 2014;15(12):e538-48.
- 549 3. Zeissig MN, Zannettino ACW, Vandyke K. Tumour Dissemination in Multiple Myeloma
550 Disease Progression and Relapse: A Potential Therapeutic Target in High-Risk Myeloma.
551 *Cancers (Basel).* 2020;12(12).
- 552 4. Akhmetzyanova I, McCarron MJ, Parekh S, Chesi M, Bergsagel PL, Fooksman DR.
553 Dynamic CD138 surface expression regulates switch between myeloma growth and
554 dissemination. *Leukemia.* 2020;34(1):245-56.
- 555 5. Garcia-Ortiz A, Rodriguez-Garcia Y, Encinas J, Maroto-Martin E, Castellano E,
556 Teixido J, et al. The Role of Tumor Microenvironment in Multiple Myeloma Development and
557 Progression. *Cancers (Basel).* 2021;13(2).
- 558 6. Aggarwal R, Ghobrial IM, Roodman GD. Chemokines in multiple myeloma.
559 *Experimental hematology.* 2006;34(10):1289-95.
- 560 7. Alsayed Y, Ngo H, Runnels J, Leleu X, Singha UK, Pitsillides CM, et al. Mechanisms
561 of regulation of CXCR4/SDF-1 (CXCL12)-dependent migration and homing in multiple
562 myeloma. *Blood.* 2007;109(7):2708-17.
- 563 8. Hu X, Villodre ES, Larson R, Rahal OM, Wang X, Gong Y, et al. Decorin-mediated
564 suppression of tumorigenesis, invasion, and metastasis in inflammatory breast cancer.
565 *Communications Biology.* 2021;4(1):72.
- 566 9. Huang S-Y, Lin H-H, Yao M, Tang J-L, Wu S-J, Hou H-A, et al. Higher Decorin Levels
567 in Bone Marrow Plasma Are Associated with Superior Treatment Response to Novel Agent-
568 Based Induction in Patients with Newly Diagnosed Myeloma - A Retrospective Study. *PloS
569 One.* 2015;10(9):e0137552.

- 570 10. Katz B-Z. Adhesion molecules—The lifelines of multiple myeloma cells. Seminars in
571 Cancer Biology. 2010;20(3):186-95.
- 572 11. Kibler C, Schermutzki F, Waller HD, Timpl R, Müller CA, Klein G. Adhesive
573 interactions of human multiple myeloma cell lines with different extracellular matrix
574 molecules. Cell Adhesion and Communication. 1998;5(4):307-23.
- 575 12. Glavey SV, Naba A, Manier S, Clauser K, Tahri S, Park J, et al. Proteomic
576 characterization of human multiple myeloma bone marrow extracellular matrix. Leukemia.
577 2017;31(11):2426-34.
- 578 13. Sanz-Rodríguez F, Ruiz-Velasco N, Pascual-Salcedo D, Teixidó J. Characterization
579 of VLA-4-dependent myeloma cell adhesion to fibronectin and VCAM-1: VLA-4-dependent
580 Myeloma Cell Adhesion. British Journal of Haematology. 1999;107(4):825-34.
- 581 14. Chatterjee M, Hönenmann D, Lentzsch S, Bommert K, Sers C, Herrmann P, et al. In
582 the presence of bone marrow stromal cells human multiple myeloma cells become
583 independent of the IL-6/gp130/STAT3 pathway. Blood. 2002;100(9):3311-8.
- 584 15. Hideshima T, Mitsiades C, Tonon G, Richardson PG, Anderson KC. Understanding
585 multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets.
586 Nature Reviews Cancer. 2007;7(8):585-98.
- 587 16. Frassanito MA, Cusmai A, Iodice G, Dammacco F. Autocrine interleukin-6 production
588 and highly malignant multiple myeloma: relation with resistance to drug-induced apoptosis.
589 Blood. 2001;97(2):483-9.
- 590 17. Urashima M, Chauhan D, Uchiyama H, Freeman G, Anderson K. CD40 ligand
591 triggered interleukin-6 secretion in multiple myeloma. Blood. 1995;85(7):1903-12.
- 592 18. Solimando AG, Malerba E, Leone P, Prete M, Terragna C, Cavo M, et al. Drug
593 resistance in multiple myeloma: Soldiers and weapons in the bone marrow niche. Frontiers in
594 Oncology. 2022;12:973836.
- 595 19. Ghobrial IM. Myeloma as a model for the process of metastasis: implications for
596 therapy. Blood. 2012;120(1):20-30.

- 597 20. Dziadowicz SA, Wang L, Akhter H, Aesoph D, Sharma T, Adjeroh DA, et al. Bone
598 Marrow Stroma-Induced Transcriptome and Regulome Signatures of Multiple Myeloma.
599 Cancers. 2022;14(4):927.
- 600 21. Seckinger A, Delgado JA, Moser S, Moreno L, Neuber B, Grab A, et al. Target
601 Expression, Generation, Preclinical Activity, and Pharmacokinetics of the BCMA-T Cell
602 Bispecific Antibody EM801 for Multiple Myeloma Treatment. Cancer Cell. 2017;31(3):396-
603 410.
- 604 22. Seckinger A, Hillengass J, Emde M, Beck S, Kimmich C, Dittrich T, et al. CD38 as
605 Immunotherapeutic Target in Light Chain Amyloidosis and Multiple Myeloma-Association
606 With Molecular Entities, Risk, Survival, and Mechanisms of Upfront Resistance. Frontiers in
607 Immunology. 2018;9:1676.
- 608 23. Burger R, Guenther A, Bakker F, Schmalzing M, Bernand S, Baum W, et al. Gp130
609 and ras mediated signaling in human plasma cell line INA-6: a cytokine-regulated tumor
610 model for plasmacytoma. The Hematology Journal: The Official Journal of the European
611 Haematology Association. 2001;2(1):42-53.
- 612 24. Gramatzki M, Burger R, Trautman U, Marschalek R, Lorenz H, Hansen-Hagge TE, et
613 al. Two new interleukin-6 dependent plasma cell lines carrying a chromosomal abnormality
614 involving the IL-6 gene locus. 1994;84 Suppl. 1:173a-a.
- 615 25. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-
616 throughput sequencing data. Bioinformatics (Oxford, England). 2015;31(2):166-9.
- 617 26. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast
618 universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
- 619 27. Zerbino DR, Achuthan P, Akanni W, Amode M R, Barrell D, Bhai J, et al. Ensembl
620 2018. Nucleic Acids Research. 2018;46(D1):D754-D61.
- 621 28. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al.
622 Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.
623 Nature Communications. 2019;10(1):1523.

- 624 29. Seabold S, Perktold J, editors. Statsmodels: Econometric and statistical modeling
625 with python. Proceedings of the 9th Python in Science Conference; 2010: Austin, TX.
- 626 30. Vallat R. Pingouin: statistics in Python. Journal of Open Source Software.
627 2018;3(31):1026.
- 628 31. Kuric M, Ebert R. plotastic: Bridging Plotting and Statistics in Python. The Journal of
629 Open Source Software. 2024.
- 630 32. Hothorn T, Lausen B. Maximally Selected Rank Statistics in R. R News. 2002;2:3-5.
- 631 33. Kuric M, Ebert R. markur4/Supplemental-INA-6-Subpopulations-and-Aggregation-
632 Detachment-Dynamics: Supplementary Data [
- 633 34. Kawano MM, Huang N, Tanaka H, Ishikawa H, Sakai A, Tanabe O, et al. Homotypic
634 cell aggregations of human myeloma cells with ICAM-1 and LFA-1 molecules. British Journal
635 of Haematology. 1991;79(4):583-8.
- 636 35. Okuno Y, Takahashi T, Suzuki A, Ichiba S, Nakamura K, Okada T, et al. In vitro
637 growth pattern of myeloma cells in liquid suspension or semi-solid culture containing
638 interleukin-6. International Journal of Hematology. 1991;54(1):41-7.
- 639 36. Blonska M, Zhu Y, Chuang HH, You MJ, Kunkalla K, Vega F, et al. Jun-regulated
640 genes promote interaction of diffuse large B-cell lymphoma with the microenvironment.
641 Blood. 2015;125(6):981-91.
- 642 37. Tai Y-T, Li X-F, Breitkreutz I, Song W, Neri P, Catley L, et al. Role of B-cell-activating
643 factor in adhesion and growth of human multiple myeloma cells in the bone marrow
644 microenvironment. Cancer Research. 2006;66(13):6675-82.
- 645 38. Standal T, Seidel C, Plesner T, Sanderson R, Waage A, Børset M, et al.
646 Osteoprotegerin is bound, internalized, and degraded by multiple myeloma cells. Blood.
647 2002;100:3002-7.
- 648 39. Van Valckenborgh E, Croucher PI, De Raeve H, Carron C, De Leenheer E, Blacher
649 S, et al. Multifunctional role of matrix metalloproteinases in multiple myeloma: a study in the
650 5T2MM mouse model. Am J Pathol. 2004;165(3):869-78.

- 651 40. Zhou F, Meng S, Song H, Claret FX. Dickkopf-1 is a key regulator of myeloma bone
652 disease: opportunities and challenges for therapeutic intervention. Blood reviews.
653 2013;27(6):261-7.
- 654 41. Sprynski AC, Hose D, Caillot L, Rème T, Shaughnessy JD, Barlogie B, et al. The role
655 of IGF-1 as a major growth factor for myeloma cell lines and the prognostic relevance of the
656 expression of its receptor. Blood. 2009;113(19):4614-26.
- 657 42. Tabolacci C, De Martino A, Mischiati C, Feriotto G, Beninati S. The Role of Tissue
658 Transglutaminase in Cancer Cell Initiation, Survival and Progression. Medical Sciences.
659 2019;7(2):19.
- 660 43. Bao L, Lai Y, Liu Y, Qin Y, Zhao X, Lu X, et al. CXCR4 is a good survival prognostic
661 indicator in multiple myeloma patients. Leukemia Research. 2013;37(9):1083-8.
- 662 44. Sarin V, Yu K, Ferguson ID, Gugliemini O, Nix MA, Hann B, et al. Evaluating the
663 efficacy of multiple myeloma cell lines as models for patient tumors via transcriptomic
664 correlation analysis. Leukemia. 2020;34(10):2754-65.
- 665 45. Dotterweich J, Schlegelmilch K, Keller A, Geyer B, Schneider D, Zeck S, et al.
666 Contact of myeloma cells induces a characteristic transcriptome signature in skeletal
667 precursor cells -Implications for myeloma bone disease. Bone. 2016;93:155-66.
- 668 46. Wadgaonkar R, Phelps KM, Haque Z, Williams AJ, Silverman ES, Collins T. CREB-
669 binding protein is a nuclear integrator of nuclear factor-kappaB and p53 signaling. The
670 Journal of Biological Chemistry. 1999;274(4):1879-82.
- 671 47. Webster GA, Perkins ND. Transcriptional Cross Talk between NF-κB and p53.
672 Molecular and Cellular Biology. 1999;19(5):3485-95.
- 673 48. Polager S, Ginsberg D. p53 and E2f: partners in life and death. Nature Reviews
674 Cancer. 2009;9(10):738-48.
- 675 49. Hose D, Rème T, Hielscher T, Moreaux J, Messner T, Seckinger A, et al. Proliferation
676 is a central independent prognostic factor and target for personalized and risk-adapted
677 treatment in multiple myeloma. Haematologica. 2011;96(1):87-95.

- 678 50. Garcés J-J, Simicek M, Vicari M, Brozova L, Burgos L, Bezdekova R, et al.
679 Transcriptional profiling of circulating tumor cells in multiple myeloma: a new model to
680 understand disease dissemination. Leukemia. 2020;34(2):589-603.
- 681 51. Brandl A, Solimando AG, Mokhtari Z, Tabares P, Medler J, Manz H, et al. Junctional
682 adhesion molecule C expression specifies a CD138low/neg multiple myeloma cell population
683 in mice and humans. Blood Advances. 2022;6(7):2195-206.
- 684 52. Wong AD, Searson PC. Mitosis-mediated intravasation in a tissue-engineered tumor-
685 microvessel platform. Cancer research. 2017;77(22):6453-61.
- 686 53. Maichl DS, Kirner JA, Beck S, Cheng W-H, Krug M, Kuric M, et al. Identification of
687 NOTCH-driven matrisome-associated genes as prognostic indicators of multiple myeloma
688 patient survival. Blood Cancer Journal. 2023;13(1):1-6.
- 689 54. Bou Zerdan M, Nasr L, Kassab J, Saba L, Ghossein M, Yaghi M, et al. Adhesion
690 molecules in multiple myeloma oncogenesis and targeted therapy. International Journal of
691 Hematologic Oncology. 2022;11(2):IJH39.
- 692 55. Evers M, Schreder M, Stühmer T, Jundt F, Ebert R, Hartmann TN, et al. Prognostic
693 value of extracellular matrix gene mutations and expression in multiple myeloma. Blood
694 Cancer Journal. 2023;13(1):43.
- 695
- 696
- 697

698 **Figure Legends**

699 **Fig. 1:** INA-6 growth conformations and survival on hMSCs. **A:** Interaction of INA-6 (green)
700 with hMSCs (black, negative staining) at different INA-6 densities (constant hMSC densities).
701 Scale bar = 200 μm . **B:** Frequency of single hMSCs (same as A) that are covered by INA-6
702 of varying group sizes. Technical replicates = three per datapoint; Single hMSCs evaluated:
703 100 per technical replicate. **C:** Interaction of INA-6 with hMSCs at different hMSC densities
704 (constant INA-6 densities). Scale bar = 300 μm . **D:** Two types of homotypic interaction:
705 Attachment after cell contact and sustained attachment of daughter cells after cell division.
706 Datapoints represent one of four independent time-lapse recordings, each evaluating 116
707 interaction events. **E:** Effects of hMSC-density on the viability (ATP, top) and apoptosis
708 (Caspase3/7 activity, bottom). INA-6:MSC ratio = 4:1; Technical replicates = four per
709 datapoint; **E left:** Signals were measured in INA-6 washed off from hMSCs and normalized
710 by INA-6 cultured in MSC-conditioned medium (= red line) (n=4). **E right:** Signals were
711 measured in co-cultures and normalized by the sum of the signals measured in hMSC and
712 INA-6 cultured separately (= red line) (n=3). **Statistics:** Paired t-test, two-factor RM-ANOVA.
713 Datapoints represent independent co-cultures with hMSCs from three (A, B, D, E right), four
714 (E left) unique donors. Confl. = Confluent.

715

716 **Fig. 2:** Time-lapse analysis of INA-6 detachment from INA-6 aggregates and hMSCs. **A:**
717 Frequency of observed INA-6 aggregates that did or did not lose INA-6 cell(s). 87 aggregates
718 were evaluated per datapoint. **B:** Example of an “disseminating” INA-6 aggregate growing on
719 fluorescently (PKH26) stained hMSC (from A-D). Dashed green lines are trajectories of
720 detached INA-6 cells. Scale bar = 50 μm . **(C-E):** Quantitative assessment of INA-6
721 detachments. 45 detachment events were evaluated per datapoint. Seeding ratio INA-6:MSC
722 = 4:1. **C:** Most INA-6 cells dissociated from another INA-6 cell and not from an hMSC
723 [F(1, 3) = 298, p-unc=4.2e-4]. **D:** Detachment frequency of aggregate size categories. **E:**
724 Detachment frequency of INA-6 cells detaching as single, pairs or more than three cells.
725 **Statistics:** (A): Paired-t-test; (C-E): Paired-t-test, Two-factor RM-ANOVA; Datapoints

726 represent three (A) or four (C-E) independent time-lapse recordings of co-cultures with
727 hMSCs from two (A) or three (C-E) unique donors.

728

729 **Fig. 3:** Detachment of INA-6 daughter cells after Cell Division. **(A-D):** INA-6 divisions in
730 interaction with confluent hMSCs. Seeding ratio INA-6:MSC = 4:20. **A:** Three examples of
731 dividing INA-6 cells generating either two MA, or one MA and one nMA daughter cells as
732 described in (G). Dashed circles mark mother cells (white), MA cell (blue), and first position
733 of nMA cell (green). Scale bar: 20 μ m. **B:** Cell division of MSC-adhering (MA) mother cell can
734 yield one mobile non-MSC-adhering (nMA) daughter cell. **C:** Frequencies of INA-6 pairs
735 defined in (A, B) per observed cell division. 65 divisions were evaluated for each of three
736 independent time-lapse recordings. **D:** Rolling duration of nMA cells after division did not
737 depend on hMSC donor [$H(2) = 5.250$, $p\text{-unc} = .072$]. Datapoints represent single nMA-cells
738 after division. **(E-G):** Adhesive and cell cycle assessment of MSC-interacting INA-6
739 subpopulations using the V-Well assay. **E:** Schematic of V-Well Assay (See Fig. S1 for
740 detailed analysis). MSC-interacting subpopulations were separated by subsequent
741 centrifugation and removal of the pellet. The pellet size was quantified by its total
742 fluorescence brightness. Adhering subpopulations were resuspended by rough pipetting. **F:**
743 Relative cell pellet sizes of adhesive INA-6 subpopulations that cycle either asynchronously
744 or were synchronized at mitosis. Gray lines in-between points connect dependent
745 measurements of co-cultures ($n=9$) that shared the same hMSC-donor and INA-6 culture.
746 Co-cultures were incubated for three different durations (1 h, 2 h, and 3 h after INA-6
747 addition). Time points were pooled, since time did not show an effect on cell adhesion
748 [$F(2,4) = 1.414$, $p\text{-unc} = 0.343$]. Factorial RM-ANOVA shows an interaction between cell
749 cycle and the kind of adhesive subpopulation [$F(1, 8) = 42.67$, $p\text{-unc} = 1.82e-04$]. Technical
750 replicates = 4 per datapoint. **G:** Cell cycles were profiled in cells gathered from the pellets of
751 four independent co-cultures ($n=4$) and the frequency of G0/G1 cells are displayed
752 depending on co-culture duration (see Fig. S3 for cell cycle profiles). Four Technical
753 replicates were pooled after pelleting. **Statistics:** (D): Kruskal-Wallis H-test. (F): Paired t-test,

754 (G): Paired t-test, two-factor RM-ANOVA. Datapoints represent INA-6 from independent co-
755 cultures with hMSCs from three unique donors.

756

757 **Fig. 4:** Separation and gene expression of INA-6 subpopulations. **A:** Schematic of “Well-
758 Plate Sandwich Centrifugation” (WPSC) separating nMA- from MA-INA6. A co-culture 96 well
759 plate is turned upside down and attached on top of a “catching plate”, forming a “well-plate
760 sandwich”. nMA-INA6 cells are collected in the catching plate by subsequent rounds of
761 centrifugation and gentle washing. MA-INA6 are enzymatically dissociated from hMSCs or by
762 rough pipetting. Subsequent RNAseq of MSC-interacting subpopulations reveals distinct
763 expression clusters [right, multidimensional scaling plot (MDS) (n=5)]. **B:** Separation was
764 microscopically tracked after each centrifugation step. **(C-E):** RT-qPCR of genes derived
765 from RNAseq results. Expression was normalized to the median of CM-INA6. Samples
766 include those used for RNAseq and six further co-cultures (n=11; non-detects were
767 discarded). **C:** Adhesion factors, ECM proteins and matrix metalloproteinases. **D:** Factors
768 involved in bone remodeling and bone homing chemokines. **E:** Factors involved in (immune)
769 signaling. **Statistics:** (C-E): Paired t-test. Datapoints represent the mean of three (B-E)
770 technical replicates. INA-6 were isolated from independent co-cultures with hMSCs from five
771 (A, B), nine (C-E) unique donors.

772

773 **Fig. 5:** Functional analysis of MSC-interacting subpopulations **(A-C):** Functional enrichment
774 analysis of differentially expressed genes (from RNAseq) using Metascape. **A:** Gene
775 ontology (GO) cluster analysis of gene lists that are unique for MA (left) or nMA (right) INA-6.
776 Circle nodes represent subsets of input genes falling into similar GO-term. Node size grows
777 with the number of input genes. Node color defines a shared parent GO-term. Two nodes
778 with a similarity score > 0.3 are linked. **B:** Enrichment analysis of pairwise comparisons
779 between MA subpopulations and their overlaps (arranged in columns). GO terms were
780 manually picked and categorized (arranged in rows). Raw Metascape results are shown in
781 Fig. S6. For each GO-term, the p-values (x-axis) and the counts of matching input genes

782 (circle size) were plotted. The lowest row shows enrichment of gene lists from the TRRUST-
783 database. **C:** Circos plots by Metascape. Sections of a circle represent lists of differentially
784 expressed genes. Purple lines connect same genes appearing in two gene lists. \cap :
785 Overlapping groups, MA: MSC-adhering, nMA: non-MSC-adhering, CM: MSC-Conditioned
786 Medium. **D:** INA-6 were co-cultured on confluent hMSC for 24 h or 48 h, separated by WPSC
787 and sub-cultured for 48 h under IL-6 withdrawal (n=6), except the control (IL-6 + INA-6)
788 (n=3). Signals were normalized (red line) to INA-6 cells grown without hMSCs and IL-6 (n=3).
789 Statistics (D): Paired t-test, two-factor RM-ANOVA. Datapoints represent the mean of four
790 technical replicates. INA-6 were isolated from independent co-cultures with hMSCs from six
791 unique donors.

792

793 **Fig. 6:** Survival of patients with multiple myeloma regarding the expression levels of
794 adhesion and bone retention genes. **A:** p-value distribution of genes associated with patient
795 survival (n=535) depending on high or low expression levels. Red dashed line marks the
796 significance threshold of $p\text{-adj}=0.05$. Histogram of p-values was plotted using a bin width of -
797 $\log_{10}(0.05)/2$. Patients with high and low gene expression were delineated using maximally
798 selected rank statistics (maxstat). **B:** Survival curves for three genes taken from the list of
799 adhesion genes shown in (A), maxstat thresholds defining high and low expression were:
800 CXCL12: 81.08; DCN: 0.75; TGM2: 0.66 normalized counts. **C:** Gene expression (RNAseq, n
801 = 873) measured in normalized counts (edgeR) of CXCL12, DCN in Bone Marrow Plasma
802 Cell (BMPC), Monoclonal Gammopathy of Undetermined Significance (MGUS), smoldering
803 Multiple Myeloma (sMM), Multiple Myeloma (MM), Multiple Myeloma Relapse (MMR), Human
804 Myeloma Cell Lines (HMCL). The red dashed line marks one normalized read count.
805 **Statistics** (A, B): Log-rank test; (C): Kruskal-Wallis, Mann–Whitney U Test. All p-values were
806 corrected using the Benjamini-Hochberg procedure.

807

808 **Fig. 7:** Proposed model of “Detached Daughter Driven Dissemination” (DDDD) in
809 aggregating multiple myeloma. **Heterotypic Interaction:** Malignant plasma cells colonize the

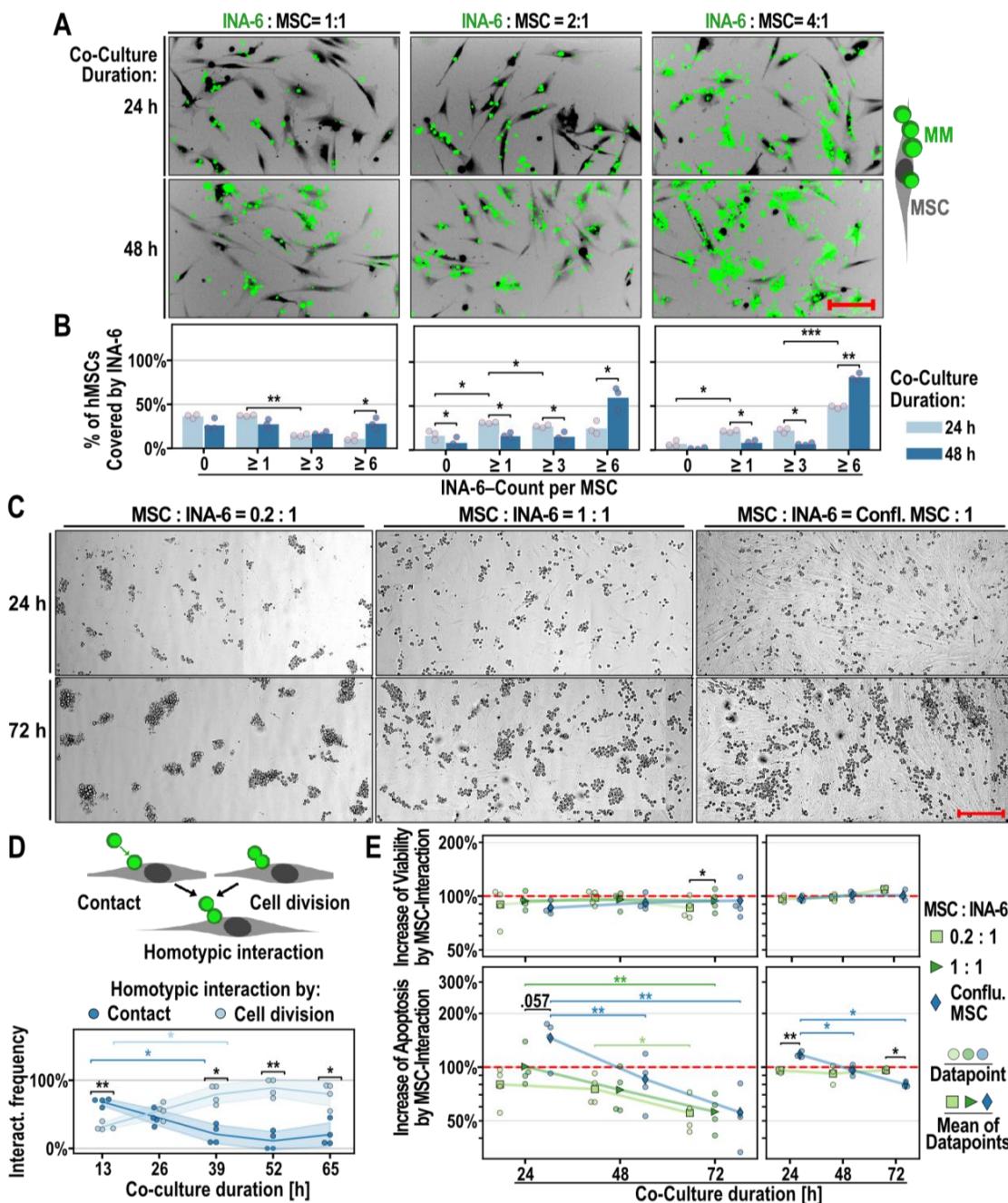
810 bone marrow microenvironment by adhering to an MSC (or osteoblast, ECM, etc.) to
811 maximize growth and survival through paracrine and adhesion mediated signaling, even if
812 contact may trigger initial apoptosis. Gene expression will focus on establishing a strong
813 anchor within the bone marrow, but also on attracting other myeloma cells (via secretion of
814 ECM factors and CXCL12/CXCL8, respectively). **Cell Division:** Cell fission can generate one
815 daughter cell that no longer adheres to the MSC (nMA). **Homotypic Interaction:** If myeloma
816 cells have the capacity to grow as aggregates, the daughter cell stays attached to their MSC-
817 adhering mother cell (MA). **Re-Adhesion:** The daughter cell “rolls around” the mother cell
818 until it re-adheres to the MSC. Our model estimates the rolling duration to be 1-10 h long.
819 **Proliferation & Saturation:** We estimate that a single myeloma cell covers one MSC
820 completely after roughly four population doublings. When heterotypic adhesion is saturated,
821 subsequent daughter cells benefit from a homotypic interaction, since they stay close to
822 growth-factor secreting MSCs and focus gene expression on proliferation (e.g. driven by
823 E2F) and not adhesion (driven by NF- κ B). **Critical Size:** Homotypic interaction is weaker
824 than heterotypic interaction, and each cell fission destabilizes the aggregate. Hence,
825 detachment of myeloma cells may depend mostly on aggregate size. **Dissemination:** After
826 myeloma cells have detached, they gained a viability advantage through IL-6-independence
827 (with unknown duration), which enhances their survival outside of the bone marrow and
828 allows them to spread throughout the body.

829

830 **Tab 1:** Adhesion and ECM genes (shown in Fig. 6A) were filtered by their association with
831 patient survival ($p\text{-adj.} < 0.01$) and was categorized as continuous downregulation during
832 disease progression. The complete list is presented in Table S2. Bone Marrow Plasma Cells
833 (BMPC), Monoclonal Gammopathy of Undetermined Significance (MGUS), moldering
834 Multiple Myeloma (sMM), Multiple Myeloma (MM), and Multiple Myeloma Relapse (MMR). $p\text{-}$
835 unc = unadjusted p -values; $p\text{-adj.}$: p -values adjusted using the Benjamini-Hochberg method
836 with 101 genes.

837

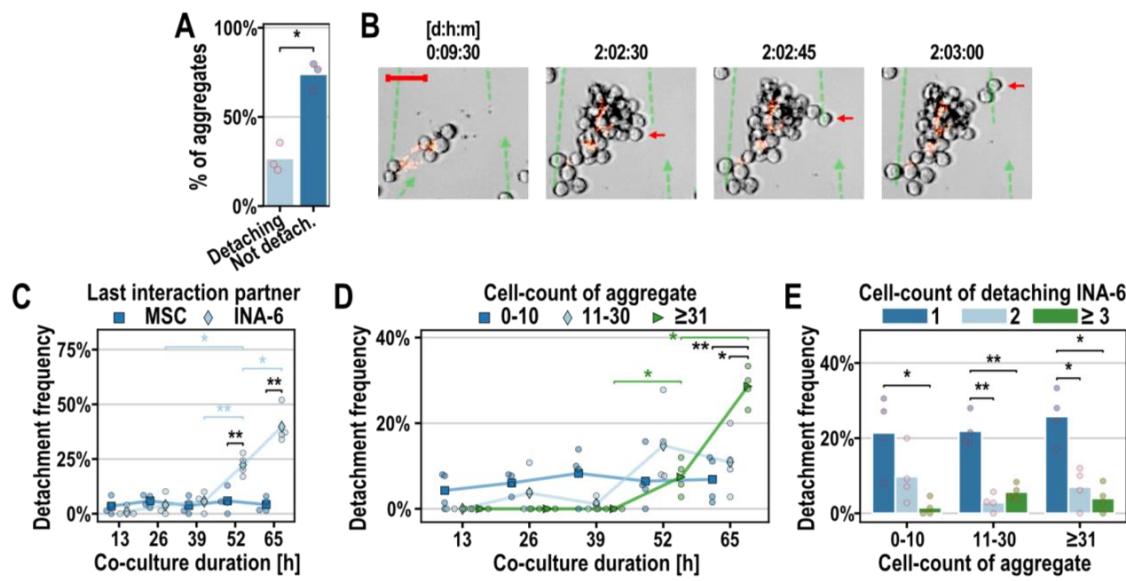
838 Fig. 1



839

840

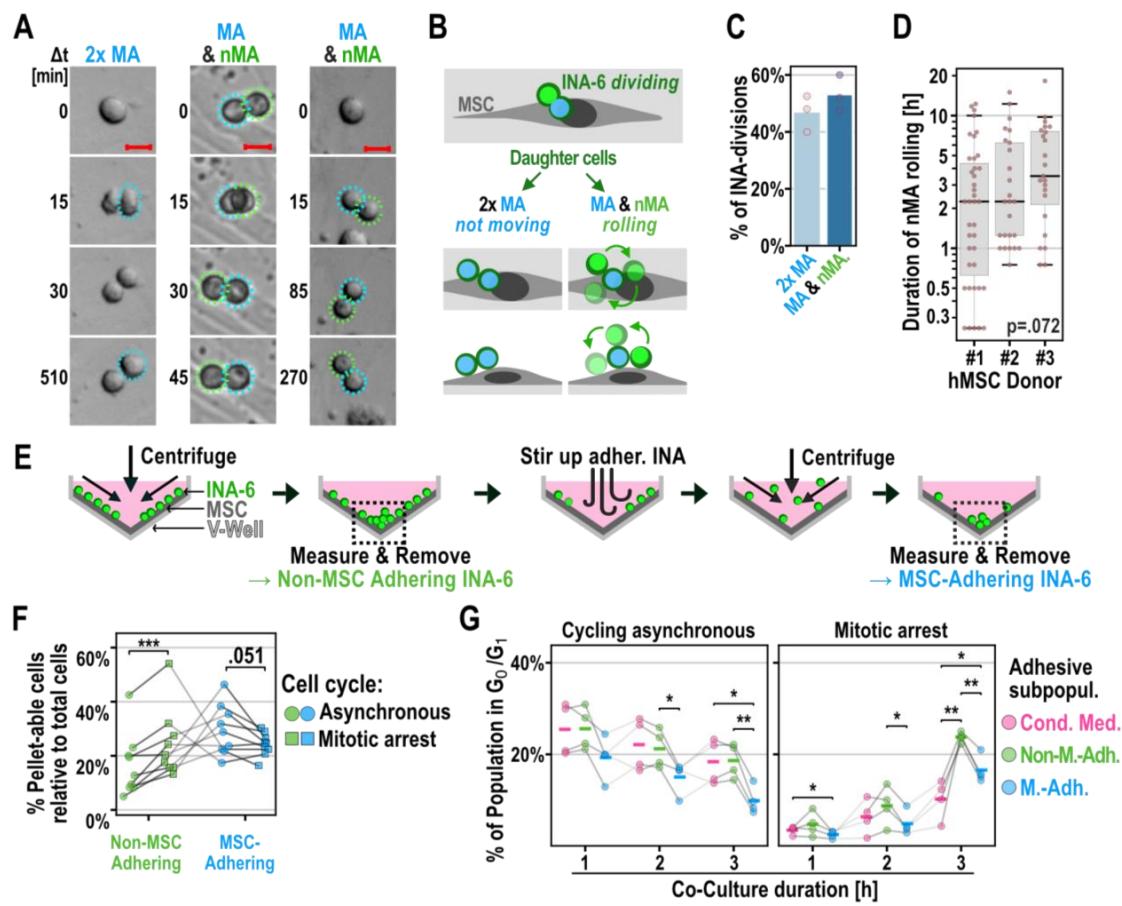
841 Fig. 2



842

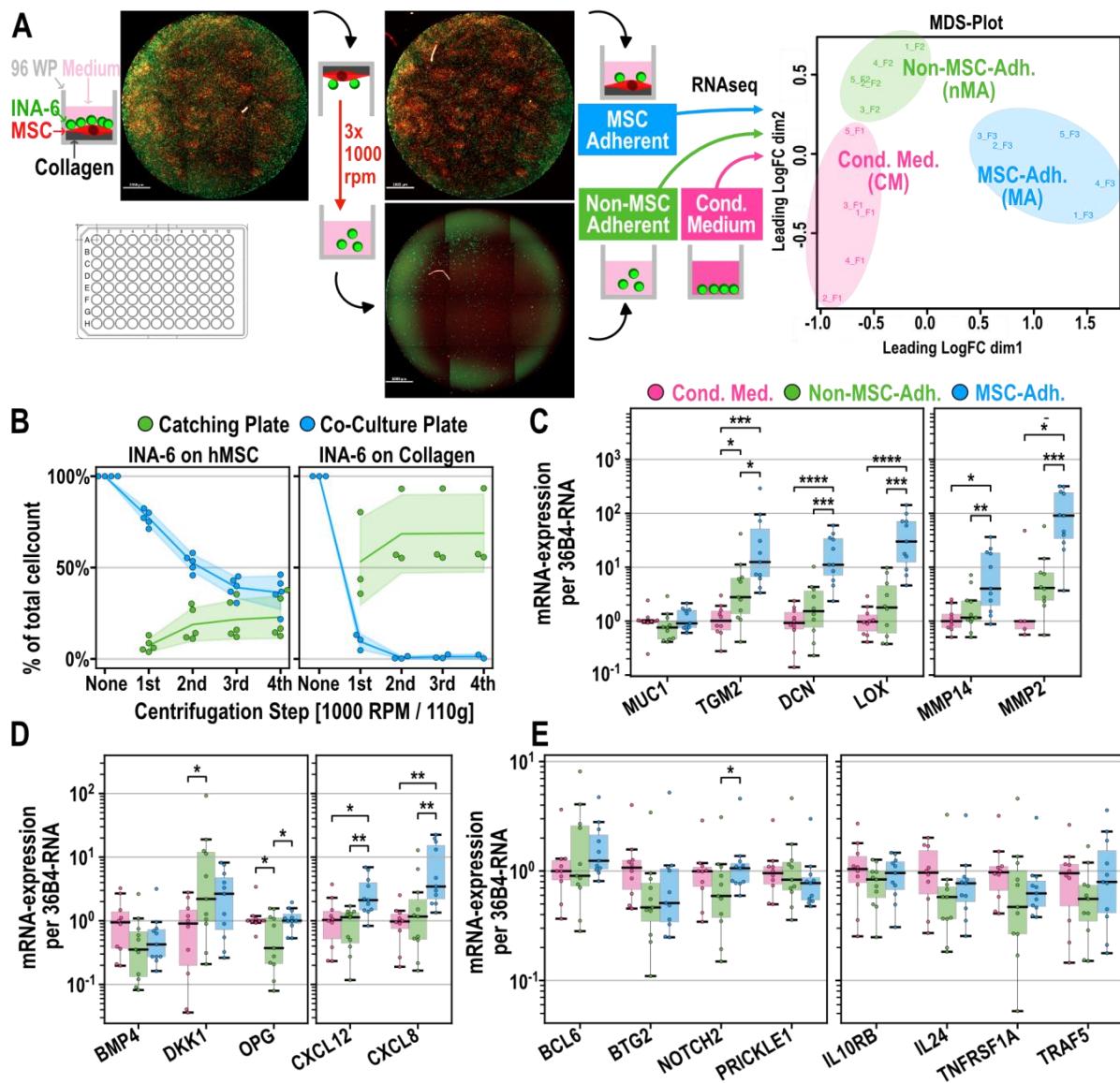
843

844 Fig. 3

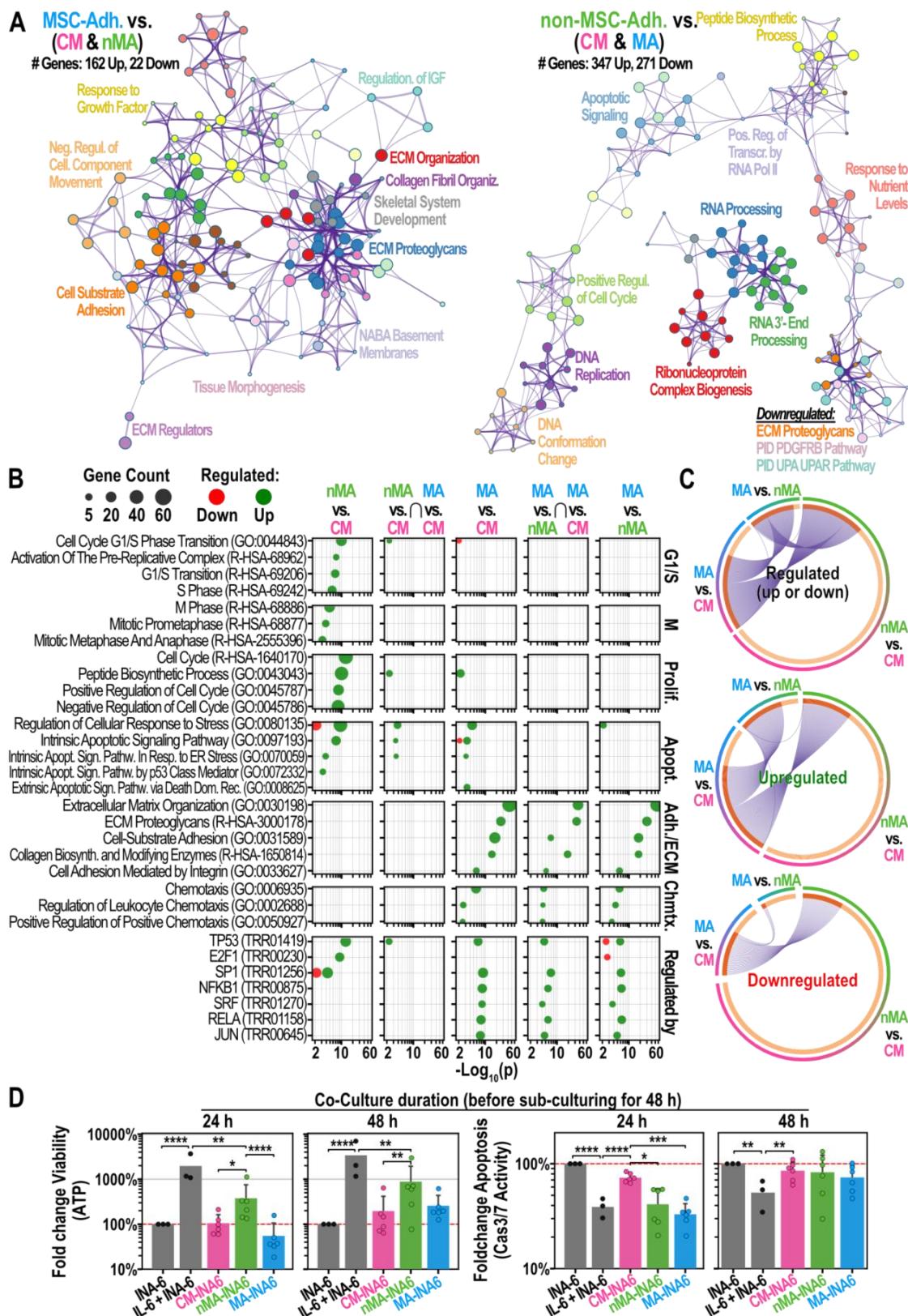


845

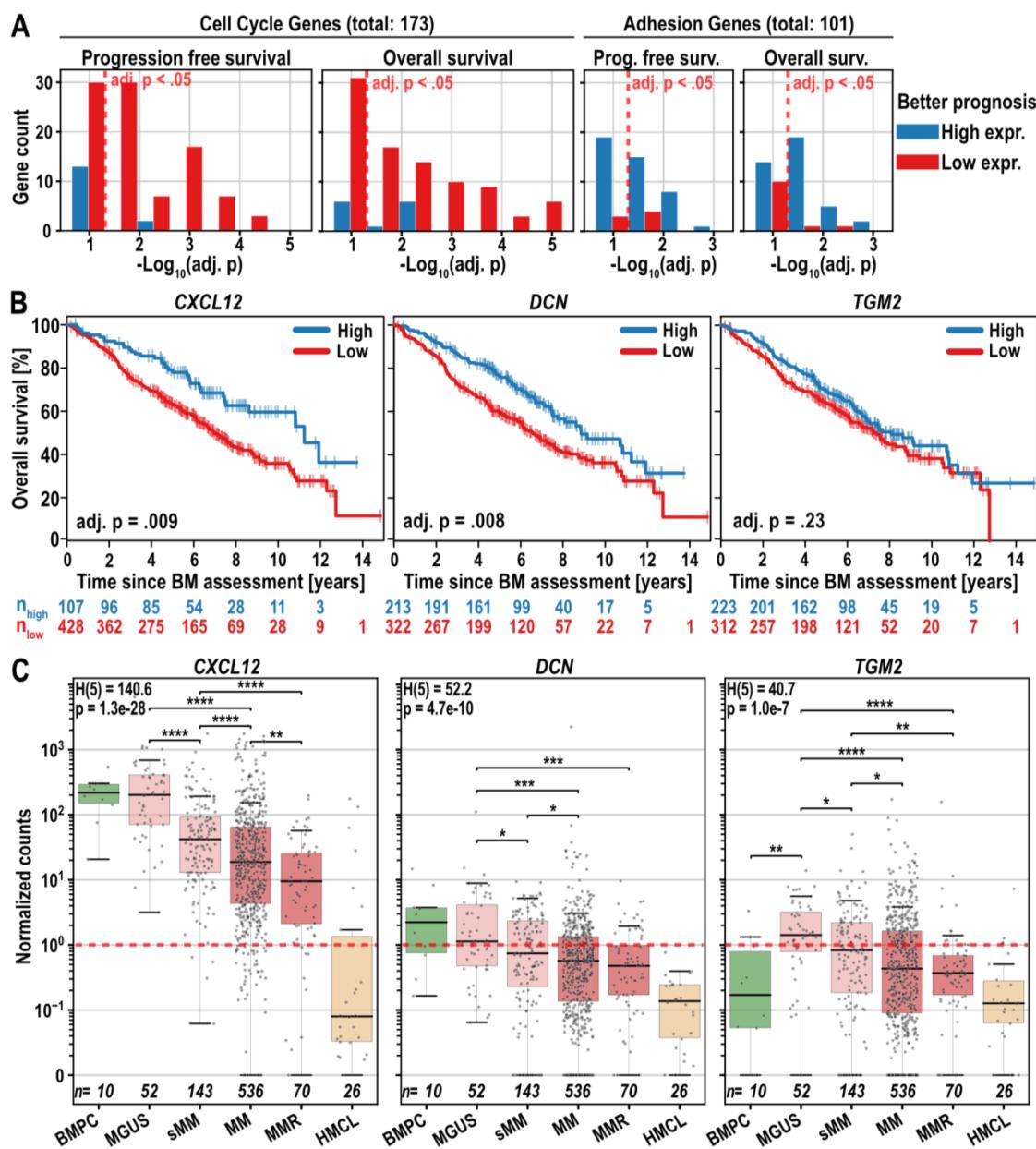
846



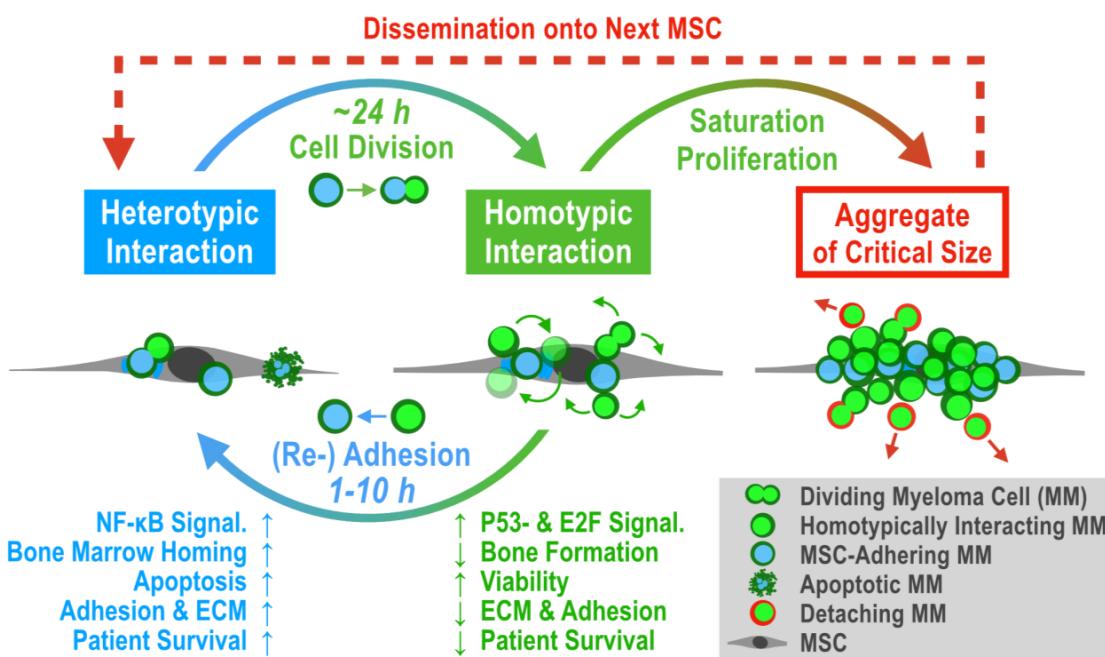
848 Fig. 5



851 Fig. 6

852
853

854 Fig. 7



855

856

857 Tab. 1

Regulation during disease progression	Gene	Ensemble ID	Progression Free / Overall Survival	Better Prognosis with high/low expression	Association of expression with survival	
					[p-unc]	[p-adj]
Not Downregulated (or overall low expression)	CCNE2	ENSG00000175305	Overall	low	5.34E-04	8.64E-03
	MMP2	ENSG00000087245	Prog. Free	high	2.29E-05	2.32E-03
	OSMR	ENSG00000145623	Prog. Free	high	5.67E-04	7.15E-03
Continuously Downregulated (BMPC > MGUS > sMM > MM > MMR)	AXL	ENSG00000167601	Overall	high	3.64E-05	1.84E-03
	COL1A1	ENSG00000108821	Prog. Free	high	3.03E-04	4.37E-03
			Overall	high	5.93E-04	8.64E-03
	CXCL12	ENSG00000107562	Prog. Free	high	1.16E-04	2.93E-03
			Overall	high	6.48E-04	8.64E-03
	CYP1B1	ENSG00000138061	Overall	high	6.84E-04	8.64E-03
	DCN	ENSG00000011465	Overall	high	2.47E-04	8.33E-03
	LRP1	ENSG00000123384	Overall	high	4.34E-04	8.64E-03
	LTBP2	ENSG00000119681	Prog. Free	high	9.03E-05	2.93E-03
	MFAP5	ENSG00000197614	Prog. Free	high	2.43E-04	4.09E-03
MYL9	ENSG00000101335	Prog. Free	high	1.46E-04	2.95E-03	
		Overall	high	1.56E-05	1.57E-03	

858

Parsed Citations

1. Blade J, Beksac M, Caers J, Jurczyszyn A, von Lilienfeld-Toal M, Moreau P, et al. Extramedullary disease in multiple myeloma: a systematic literature review. *Blood Cancer Journal*. 2022;12(3):1-10.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
2. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos M-V, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *The Lancet Oncology*. 2014;15(12):e538-48.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
3. Zeissig MN, Zannettino ACW, Vandyke K. Tumour Dissemination in Multiple Myeloma Disease Progression and Relapse: A Potential Therapeutic Target in High-Risk Myeloma. *Cancers (Basel)*. 2020;12(12).
4. Akhmetzyanova I, McCarron MJ, Parekh S, Chesi M, Bergsagel PL, Fooksman DR. Dynamic CD138 surface expression regulates switch between myeloma growth and dissemination. *Leukemia*. 2020;34(1):245-56.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
5. Garcia-Ortiz A, Rodriguez-Garcia Y, Encinas J, Maroto-Martin E, Castellano E, Teixido J, et al. The Role of Tumor Microenvironment in Multiple Myeloma Development and Progression. *Cancers (Basel)*. 2021;13(2).
6. Aggarwal R, Ghobrial IM, Roodman GD. Chemokines in multiple myeloma. *Experimental hematology*. 2006;34(10):1289-95.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
7. Alsayed Y, Ngo H, Runnels J, Leleu X, Singha UK, Pitsillides CM, et al. Mechanisms of regulation of CXCR4/SDF-1 (CXCL12)-dependent migration and homing in multiple myeloma. *Blood*. 2007;109(7):2708-17.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
8. Hu X, Villodre ES, Larson R, Rahal OM, Wang X, Gong Y, et al. Decorin-mediated suppression of tumorigenesis, invasion, and metastasis in inflammatory breast cancer. *Communications Biology*. 2021;4(1):72.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
9. Huang S-Y, Lin H-H, Yao M, Tang J-L, Wu S-J, Hou H-A, et al. Higher Decorin Levels in Bone Marrow Plasma Are Associated with Superior Treatment Response to Novel Agent- Based Induction in Patients with Newly Diagnosed Myeloma - A Retrospective Study. *PloS One*. 2015;10(9):e0137552.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
10. Katz B-Z. Adhesion molecules-The lifelines of multiple myeloma cells. *Seminars in Cancer Biology*. 2010;20(3):186-95.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)
11. Kibler C, Schermutzki F, Waller HD, Timpl R, Muller CA, Klein G. Adhesive interactions of human multiple myeloma cell lines with different extracellular matrix molecules. *Cell Adhesion and Communication*. 1998;5(4):307-23.
Pubmed: [Author and Title](#)
Google Scholar: [Google Scholar Search](#)

12. Glavey SV, Naba A, Manier S, Clauser K, Tahri S, Park J, et al. Proteomic characterization of human multiple myeloma bone marrow extracellular matrix. Leukemia. 2017;31(11):2426-34.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

13. Sanz-Rodriguez F, Ruiz-Velasco N, Pascual-Salcedo D, Teixido J. Characterization of VLA-4-dependent myeloma cell adhesion to fibronectin and VCAM-1: VLA-4-dependent Myeloma Cell Adhesion. British Journal of Haematology. 1999;107(4):825-34.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

14. Chatterjee M, Honemann D, Lentzsch S, Bommert K, Sers C, Herrmann P, et al. In the presence of bone marrow stromal cells human multiple myeloma cells become independent of the IL-6/gp130/STAT3 pathway. Blood. 2002;100(9):3311-8.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

15. Hideshima T, Mitsiades C, Tonon G, Richardson PG, Anderson KC. Understanding multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets. Nature Reviews Cancer. 2007;7(8):585-98.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

16. Frassanito MA, Cusmai A, Iodice G, Dammacco F. Autocrine interleukin-6 production and highly malignant multiple myeloma: relation with resistance to drug-induced apoptosis. Blood. 2001;97(2):483-9.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

17. Urashima M, Chauhan D, Uchiyama H, Freeman G, Anderson K. CD40 ligand triggered interleukin-6 secretion in multiple myeloma. Blood. 1995;85(7):1903-12.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

18. Solimando AG, Malerba E, Leone P, Prete M, Terragna C, Cavo M, et al. Drug resistance in multiple myeloma: Soldiers and weapons in the bone marrow niche. Frontiers in Oncology. 2022;12:973836.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

19. Ghobrial IM. Myeloma as a model for the process of metastasis: implications for therapy. Blood. 2012;120(1):20-30.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

20. Dziadowicz SA, Wang L, Akhter H, Aesoph D, Sharma T, Adjero DA, et al. Bone Marrow Stroma-Induced Transcriptome and Regulome Signatures of Multiple Myeloma. Cancers. 2022;14(4):927.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

21. Seckinger A, Delgado JA, Moser S, Moreno L, Neuber B, Grab A, et al. Target Expression, Generation, Preclinical Activity, and Pharmacokinetics of the BCMA-T Cell Bispecific Antibody EM801 for Multiple Myeloma Treatment. Cancer Cell. 2017;31(3):396- 410.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

22. Seckinger A, Hillengass J, Emde M, Beck S, Kimmich C, Dittrich T, et al. CD38 as Immunotherapeutic Target in Light Chain Amyloidosis and Multiple Myeloma-Association With Molecular Entities, Risk, Survival, and Mechanisms of Upfront Resistance. Frontiers in Immunology. 2018;9:1676.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

23. Burger R, Guenther A, Bakker F, Schmalzing M, Bernand S, Baum W, et al. Gp130 and ras mediated signaling in human plasma cell line INA-6: a cytokine-regulated tumor model for plasmacytoma. *The Hematology Journal: The Official Journal of the European Haematology Association.* 2001;2(1):42-53.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

24. Gramatzki M, Burger R, Trautman U, Marschalek R, Lorenz H, Hansen-Hagge TE, et al. Two new interleukin-6 dependent plasma cell lines carrying a chromosomal abnormality involving the IL-6 gene locus. *1994;84 Suppl. 1:173a-a.*

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

25. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England).* 2015;31(2):166-9.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

26. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

27. Zerbino DR, Achuthan P, Akanni W, Amode M R, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Research.* 2018;46(D1):D754-D61.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

28. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications.* 2019;10(1):1523.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

29. Seabold S, Perktold J, editors. *Statsmodels: Econometric and statistical modeling with python.* Proceedings of the 9th Python in Science Conference; 2010: Austin, TX.

30. Vallat R. Pingouin: statistics in Python. *Journal of Open Source Software.* 2018;3(31):1026.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

31. Kuric M, Ebert R. plotastic: Bridging Plotting and Statistics in Python. *The Journal of Open Source Software.* 2024.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

32. Hothorn T, Lausen B. Maximally Selected Rank Statistics in R. *R News.* 2002;2:3-5.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

33. Kuric M, Ebert R. markur4/Supplemental-INA-6-Subpopulations-and-Aggregation- Detachment- Dynamics: Supplementary Data [

34. Kawano MM, Huang N, Tanaka H, Ishikawa H, Sakai A, Tanabe O, et al. Homotypic cell aggregations of human myeloma cells with ICAM-1 and LFA-1 molecules. *British Journal of Haematology.* 1991;79(4):583-8.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

35. Okuno Y, Takahashi T, Suzuki A, Ichiba S, Nakamura K, Okada T, et al. In vitro growth pattern of myeloma cells in liquid suspension or semi-solid culture containing interleukin-6. *International Journal of Hematology.* 1991;54(1):41-7.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

36. Blonska M, Zhu Y, Chuang HH, You MJ, Kunkalla K, Vega F, et al. Jun-regulated genes promote interaction of diffuse large B-cell lymphoma with the microenvironment. *Blood*. 2015;125(6):981-91.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

37. Tai Y-T, Li X-F, Breitkreutz I, Song W, Neri P, Catley L, et al. Role of B-cell-activating factor in adhesion and growth of human multiple myeloma cells in the bone marrow microenvironment. *Cancer Research*. 2006;66(13):6675-82.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

38. Standal T, Seidel C, Plesner T, Sanderson R, Waage A, Borset M, et al. Osteoprotegerin is bound, internalized, and degraded by multiple myeloma cells. *Blood*. 2002;100:3002-7.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

39. Van Valckenborgh E, Croucher PI, De Raeve H, Carron C, De Leenheer E, Blacher S, et al. Multifunctional role of matrix metalloproteinases in multiple myeloma: a study in the

5. T2MM mouse model. Am J Pathol. 2004;165(3):869-78.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

40. Zhou F, Meng S, Song H, Claret FX. Dickkopf-1 is a key regulator of myeloma bone disease: opportunities and challenges for therapeutic intervention. *Blood reviews*. 2013;27(6):261-7.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

41. Sprynski AC, Hose D, Caillot L, Reme T, Shaughnessy JD, Barlogie B, et al. The role of IGF-1 as a major growth factor for myeloma cell lines and the prognostic relevance of the expression of its receptor. *Blood*. 2009;113(19):4614-26.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

42. Tabolacci C, De Martino A, Mischiati C, Feriotto G, Beninati S. The Role of Tissue Transglutaminase in Cancer Cell Initiation, Survival and Progression. *Medical Sciences*. 2019;7(2):19.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

43. Bao L, Lai Y, Liu Y, Qin Y, Zhao X, Lu X, et al. CXCR4 is a good survival prognostic indicator in multiple myeloma patients. *Leukemia Research*. 2013;37(9):1083-8.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

44. Sarin V, Yu K, Ferguson ID, Guglielmini O, Nix MA, Hann B, et al. Evaluating the efficacy of multiple myeloma cell lines as models for patient tumors via transcriptomic correlation analysis. *Leukemia*. 2020;34(10):2754-65.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

45. Dotterweich J, Schlegelmilch K, Keller A, Geyer B, Schneider D, Zeck S, et al. Contact of myeloma cells induces a characteristic transcriptome signature in skeletal precursor cells -Implications for myeloma bone disease. *Bone*. 2016;93:155-66.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

46. Wadgaonkar R, Phelps KM, Haque Z, Williams AJ, Silverman ES, Collins T. CREB- binding protein is a nuclear integrator of nuclear factor-kappaB and p53 signaling. *The Journal of Biological Chemistry*. 1999;274(4):1879-82.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

47. Webster GA, Perkins ND. Transcriptional Cross Talk between NF- κ B and p53. Molecular and Cellular Biology. 1999;19(5):3485-95.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

48. Polager S, Ginsberg D. p53 and E2f: partners in life and death. Nature Reviews Cancer. 2009;9(10):738-48.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

49. Hose D, Reme T, Hielscher T, Moreaux J, Messner T, Seckinger A, et al. Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. Haematologica. 2011;96(1):87-95.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

50. Garces J-J, Simicek M, Vicari M, Brozova L, Burgos L, Bezdekova R, et al. Transcriptional profiling of circulating tumor cells in multiple myeloma: a new model to understand disease dissemination. Leukemia. 2020;34(2):589-603.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

51. Brandl A, Solimando AG, Mokhtari Z, Tabares P, Medler J, Manz H, et al. Junctional adhesion molecule C expression specifies a CD138low/neg multiple myeloma cell population in mice and humans. Blood Advances. 2022;6(7):2195-206.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

52. Wong AD, Searson PC. Mitosis-mediated intravasation in a tissue-engineered tumor- microvessel platform. Cancer research. 2017;77(22):6453-61.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

53. Maichl DS, Kirner JA, Beck S, Cheng W-H, Krug M, Kuric M, et al. Identification of NOTCH-driven matrisome-associated genes as prognostic indicators of multiple myeloma patient survival. Blood Cancer Journal. 2023;13(1):1-6.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

54. Bou Zerdan M, Nasr L, Kassab J, Saba L, Ghossein M, Yaghi M, et al. Adhesion molecules in multiple myeloma oncogenesis and targeted therapy. International Journal of Hematologic Oncology. 2022;11(2):IJH39.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

55. Evers M, Schreder M, Stuhmer T, Jundt F, Ebert R, Hartmann TN, et al. Prognostic value of extracellular matrix gene mutations and expression in multiple myeloma. Blood Cancer Journal. 2023;13(1):43.

Pubmed: [Author and Title](#)

Google Scholar: [Google Scholar Search](#)

Chapter 2: Semi-Automation of Data Analysis

Abstract

`plotastic` addresses the challenges of transitioning from exploratory data analysis to hypothesis testing in Python’s data science ecosystem. Bridging the gap between `seaborn` and `pingouin`, this library offers a unified environment for plotting and statistical analysis. It simplifies the workflow with user-friendly syntax and seamless integration with familiar `seaborn` parameters (`y`, `x`, `hue`, `row`, `col`). Inspired by `seaborn`’s consistency, `plotastic` utilizes a `DataAnalysis` object to intelligently pass parameters to `pingouin` statistical functions. Hence, statistics and plotting are performed on the same set of parameters, so that the strength of `seaborn` in visualizing multidimensional data is extended onto statistical analysis. In essence, `plotastic` translates `seaborn` parameters into statistical terms, configures statistical protocols based on intuitive plotting syntax and returns a `matplotlib` figure with known customization options and more. This approach streamlines data analysis, allowing researchers to focus on correct statistical testing and less about specific syntax and implementations.

Introduction

The reproducibility crisis in research highlights a significant challenge in contemporary bio-sciences, where a substantial portion of studies faces reproducibility issues (Begley & Ioannidis, 2015). One critical yet often overlooked aspect contributing to this crisis is data management. The literature most often refers to *big data* as the main challenge (Gomez-Cabrero et al., 2014). However, these challenges are also present in smaller datasets, which the author refers to as *semi-big data*. This term describes datasets that, while not extensive enough to necessitate advanced computational tools typically reserved for *big data*, are sufficiently large to render manual analysis very time-intensive. *Semi-big data* is often generated by methods like automated microscopy or multiplex qPCR, which produce volumes of data that are manageable on a surface level, but pose substantial barriers for in-depth, manual reproducibility (Bustin, 2014). This is further complicated by the complexity inherent in multidimensional datasets. For example, the qPCR experiment from Chapter 1, Fig. 4 involves the analysis of 19 genes across three subpopulations, including eleven biological and three technical replicates, resulting in a total of 1881 data points that are all assigned to a complex set of experimental variables. Without a clearly documented data analysis protocol and standardized data formats, the reproduction of such analysis becomes extremely challenging, if not impossible (Bustin, 2014).

The evolving standards in data analysis advocate for the standardization of analytical pipelines, rationalization of sample sizes, and enhanced infrastructure for data storage, address-

ing some of these challenges (Goodman et al., 2016; Wilkinson et al., 2016). However, these advancements can place undue pressure on researchers, particularly those with limited training in statistics, underscoring the need for intuitive, user-friendly analytical tools (Gosselin, 2021; Armstrong, 2014; Gómez-López et al., 2019)

In this context, `plotastic` emerges as a tool designed to democratize access to sophisticated statistical analysis, offering a user-centric interface that caters to researchers across varying levels of statistical proficiency. By integrating robust statistical methodologies within an accessible framework, `plotastic` aims to contribute to enhancing the reproducibility and integrity of research in the biosciences (Gomez-Cabrero et al., 2014).

initially, the need to develop `plotastic` arose during this project. The first is to address the author's need for a tool that could handle the complex, multidimensional data generated by e.g. qPCR experiments. These experiments typically involve the analysis of multiple genes across several time points and biological replicates, resulting in datasets that are challenging to analyze manually. The author's experience with traditional statistical software, such as Prism, revealed that these tools required extensive manual input, making them unsuitable for the efficient analysis of complex, multidimensional data. - The second was to increase speed. This is required for developing methods

Since `plotastic` optimizes the analysis of *semi-big data*, we introduce the term *semi-automation* to distinguish itself from the fully automated pipelines used for *big data*. Semi-automation is defined as the following aspects:

1. **Semi-big input:** The input size is oriented towards *semi-big data*, which is characterized as being manageable by manual analysis, yet highly time inefficient, and probably impossible to re-analyse by someone else than the researcher.
2. **Standardized input** The input follows a standardized format (e.g. long-format)
3. **Minimize user configuration:** User configuration is strictly minimized. The user is never asked to pass the same parameters twice. This reduces the risk of human error and time spent on configuration.
4. **Default configuration provides acceptable results:** If the user does not provide any manual configuration, the pipeline should provide acceptable results. Options should be provided to allow a level of flexibility to adapt the pipeline to the user's needs.
5. **Small Reviewable Processing Steps:** The analysis steps are structured into small processes that can be combined to form a complete analysis pipeline. That way, each step can act as a stage for quality control to improve error detection and troubleshooting. For

a statistical analysis, that means the processing steps are separated into 3 steps, those being assumption testing, factor analysis and post-hoc testing.

6. **Isolated Steps:** Processing steps should work independently from another, in the best case only depending on the raw data input. If a processing step depends on the output from other steps, the software should tell the user what exact steps it expects.
7. **Human readable outputs:** Every processing step may provide an output that is not necessarily standardized, but is required to be human readable to ensure reviewability.

Challenges: - Reproducibility crisis? - Data is exploding - Demands for rigorous statistical analysis are increasing - Biologists are not trained in statistics

The demands are rising: (Moreno-Indias et al., 2021)

As laid out in the introduction, one can doubt if a PhD student without coding skills is at its max efficiency.

Why does Biomedicine need plotastic?: - Thorough analysis has become a standard, with assumption testing, omnibus tests and post-hoc analyses for every experiment. - But data is increasing - Example of my data? - The number of dedicated statisticians is limited - The know-how of statistics in biology is limited, for example, Some authors ignored the problem of multiple testing while others used the method uncritically with no rationale or discussion (Perneger, 1998; Armstrong, 2014)

Why did I need plotastic?

Why do biologists need plotastic? - Assays output more data in shorter time, e.g. multiplex qPCR - example: 20 genes, 3 timepoints, 11 biological replicates, (all 3 technical replicates already averaged) - $20 * 3 * 11 = 660$ data points

this is multidimensional data: 660 data points spread across two dimensions: time and gene

- in manual analysis e.g. in Excel, the user has to manually select the data, copy it, paste it into a new sheet, and then perform the statistical test. In Prism, the user has to select the data, click on the statistical test, and then select the data again. This is not only time-consuming, but also prone to

- Re-Analysis: The user has to repeat the process for every gene and timepoint. This is not only time-consuming, but also prone to errors.

shortly Describe Main Packages in more detail: - seaborn: It multidimensional data - pingouin: It's a statistical package

Statement of Need

Python's data science ecosystem provides powerful tools for both visualization and statistical testing. However, the transition from exploratory data analysis to hypothesis testing can be cumbersome, requiring users to switch between libraries and adapt to different syntaxes. `seaborn` has become a popular choice for plotting in Python, offering an intuitive interface. Its statistical functionality focuses on descriptive plots and bootstrapped confidence intervals (Waskom, 2021). The library `pingouin` offers an extensive set of statistical tests, but it lacks integration with common plotting capabilities (Vallat, 2018). `statannotations` integrates statistical testing with plot annotations, but uses a complex interface and is limited to pairwise comparisons (Charlier et al., 2022).

`plotastic` addresses this gap by offering a unified environment for plotting and statistical analysis. With an emphasis on user-friendly syntax and integration of familiar `seaborn` parameters, it simplifies the process for users already comfortable with `seaborn`. The library ensures a smooth workflow, from data import to hypothesis testing and visualization.

Example

The following code demonstrates how `plotastic` analyzes the example dataset “fmri”, similar to Waskom (2021) (Figure 1).

```
1 ### IMPORT PLOTASTIC
2 import plotastic as plst
3
4 # IMPORT EXAMPLE DATA
5 DF, _dims = plst.load_dataset("fmri", verbose = False)
6
7 # EXPLICITLY DEFINE DIMENSIONS TO FACET BY
8 dims = dict(
9     y = "signal",      # y-axis, dependent variable
10    x = "timepoint",   # x-axis, independent variable (within-subject factor)
11    hue = "event",     # color, independent variable (within-subject factor)
12    col = "region"    # axes, grouping variable
13 )
14 # INITIALIZE DATAANALYSIS OBJECT
15 DA = plst.DataAnalysis(
16     data=DF,           # Dataframe, long format
17     dims=dims,         # Dictionary with y, x, hue, col, row
18     subject="subject", # Datapoints are paired by subject (optional)
19     verbose=False,     # Print out info about the Data (optional)
20 )
21 # STATISTICAL TESTS
22 DA.check_normality() # Check Normality
23 DA.check_sphericity() # Check Sphericity
```

```

24 DA.omnibus_rm_anova() # Perform RM-ANOVA
25 DA.test_pairwise()     # Perform Posthoc Analysis
26
27 # PLOTTING
28 (DA
29 .plot_box_strip()    # Pre-built plotting function initializes plot
30 .annotate_pairwise(  # Annotate results from DA.test_pairwise()
31     include="__HUE" # Use only significant pairs across each hue
32 )
33 )

```

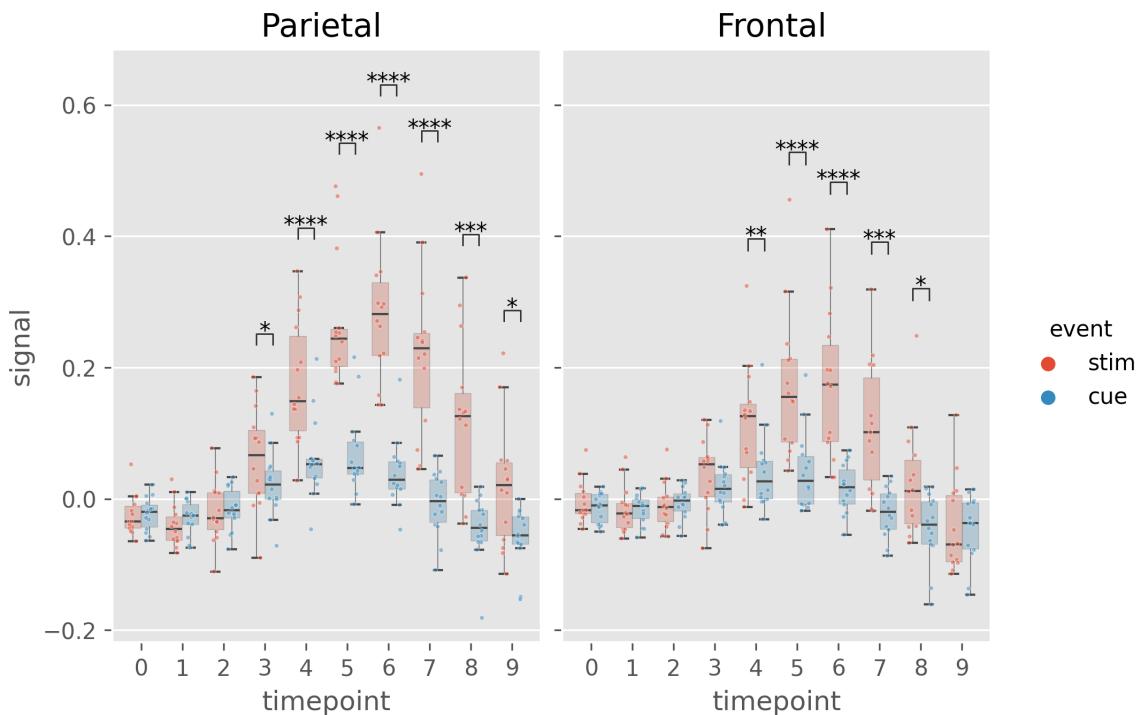


Figure 1: Example figure of plotastic (version 0.1). Image style was set by `plt.style.use('ggplot')`

Table 1: Results from `DA.check_sphericity()`. plotastic assesses sphericity after grouping the data by all grouping dimensions (hue, row, col). For example, `DA.check_sphericity()` grouped the ‘fmri’ dataset by “region” (col) and “event” (hue), performing four subsequent sphericity tests for four datasets.

‘region’, ‘event’	spher	W	chi2	dof	pval	group count	n per group
‘frontal’, ‘cue’	True	3.26e+20	-462.7	44	1	10	[14]
‘frontal’, ‘stim’	True	2.45e+17	-392.2	44	1	10	[14]
‘parietal’, ‘cue’	True	1.20e+20	-452.9	44	1	10	[14]
‘parietal’, ‘stim’	True	2.44e+13	-301.9	44	1	10	[14]

Table 2: Results of `DA.omnibus_rm_anova()`. `plotastic` performs one two-factor RM-ANOVA per axis (grouping the data by row and col dimensions) using `x` and `hue` as the within-factors. For this example, `DA.omnibus_rm_anova()` grouped the ‘fmri’ dataset by “region” (col), performing two subsequent two-factor RM-ANOVAs. Within-factors are “timepoint” (`x`) and “event” (`hue`). For conciseness, GG-Correction and effect sizes are not shown.

‘region’	Source	SS	ddof1	ddof2	MS	F	p-unc	stars
‘parietal’	timepoint	1.583	9	117	0.175	26.20	3.40e-24	****
‘parietal’	event	0.770	1	13	0.770	85.31	4.48e-07	****
‘parietal’	timepoint * event	0.623	9	117	0.069	29.54	3.26e-26	****
‘frontal’	timepoint	0.686	9	117	0.076	15.98	8.28e-17	****
‘frontal’	event	0.240	1	13	0.240	23.44	3.21e-4	***
‘frontal’	timepoint * event	0.242	9	117	0.026	13.031	3.23e-14	****

Overview

The functionality of `plotastic` revolves around a seamless integration of statistical analysis and plotting, leveraging the capabilities of `pingouin`, `seaborn`, `matplotlib` and `statannotations` (Vallat, 2018; Waskom, 2021; Hunter, 2007; Charlier et al., 2022). It utilizes long-format `pandas` `DataFrames` as its primary input, aligning with the conventions of `seaborn` and ensuring compatibility with existing data structures (Wickham, 2014; Team, 2020; McKinney, 2010).

`plotastic` was inspired by `seaborn` using the same set of intuitive and consistent parameters (`y`, `x`, `hue`, `row`, `col`) found in each of its plotting functions (Waskom, 2021). These parameters intuitively delineate the data dimensions plotted, yielding ‘faceted’ subplots, each presenting `y` against `x`. This allows for rapid and insightful exploration of multidimensional relationships. `plotastic` extends this principle to statistical analysis by storing these `seaborn` parameters (referred to as dimensions) in a `DataAnalysis` object and intelligently passing them to statistical functions of the `pingouin` library. This approach is based on the impression that most decisions during statistical analysis can be derived from how the user decides to arrange the data in a plot. This approach also prevents code repetition and streamlines statistical analysis. For example, the `subject` keyword is specified only once during `DataAnalysis` initialisation, and `plotastic` selects the appropriate paired or unpaired version of the test. Using `pingouin` alone requires the user to manually pick the correct test and to repeatedly specify the `subject` keyword in each testing function.

In essence, `plotastic` translates plotting parameters into their statistical counterparts. This translation minimizes user input and also ensures a coherent and logical connection between plotting and statistical analysis. The goal is to allow the user to focus on choosing the correct statistical test (e.g. parametric vs. non-parametric) and worry less about specific implementations.

At its core, `plotastic` employs iterators to systematically group data based on various

dimensions, aligning the analysis with the distinct requirements of tests and plots. Normality testing is performed on each individual sample, which is achieved by splitting the data by all grouping dimensions and also the x-axis (hue, row, col, x). Sphericity and homoscedasticity testing is performed on a complete sampleset listed on the x-axis, which is achieved by splitting the data by all grouping dimensions (hue, row, col) (Table 1). For omnibus and posthoc analyses, data is grouped by the row and col dimensions in parallel to the `matplotlib` axes, before performing one two-factor analysis per axis using x and hue as the within/between-factors. (Table 2).

`DataAnalysis` visualizes data through predefined plotting functions designed for drawing multi-layered plots. A notable emphasis within `plotastic` is placed on showcasing individual datapoints alongside aggregated means or medians. In detail, each plotting function initializes a `matplotlib` figure and axes using `plt.subplots()` while returning a `DataAnalysis` object for method chaining. Axes are populated by `seaborn` plotting functions (e.g., `sns.boxplot()`), leveraging automated aggregation and error bar displays. Keyword arguments are passed to these `seaborn` functions, ensuring the same degree of customization. Users can further customize plots by chaining `DataAnalysis` methods or by applying common `matplotlib` code to override `plotastic` settings. Figures are exported using `plt.savefig()`.

`plotastic` also focuses on annotating statistical information within plots, seamlessly incorporating p-values from pairwise comparisons using `statannotations` (Charlier et al., 2022). This integration simplifies the interface and enables options for pair selection in multidimensional plots, enhancing both user experience and interpretability.

For statistics, `plotastic` integrates with the `pingouin` library to support classical assumption and hypothesis testing, covering parametric/non-parametric and paired/non-paired variants. Assumptions such as normality, homoscedasticity, and sphericity are tested. Omnibus tests include two-factor RM-ANOVA, ANOVA, Friedman, and Kruskal-Wallis. Posthoc tests are t-tests, Wilcoxon, and Mann-Whitney-U.

To sum up, `plotastic` stands as a unified and user-friendly solution catering to the needs of researchers and data scientists, seamlessly integrating statistical analysis with the power of plotting in Python. It streamlines the workflow, translates `seaborn` parameters into statistical terms, and supports extensive customization options for both analysis and visualization.

Discussion

Is plotastic tested? Coverage? Does it cover every feature? What is not covered

Is plotastic USABLE for biologists? - Yes but use is limited by minimal knowledge of Python
- However, that is subject to change as Python is becoming more popular in biology and AI

assisted coding decreased the barrier to entry significantly. Tools like github copilot are able to generate code, fix bugs and suggest improvements. This is a game changer for biologists that are not familiar with programming. - Furthermore, installing and using plotastic for biologists is overestimated. These steps are needed:

- Install anaconda from the internet
- Open the terminal
- Type `pip install plotastic`
- Check Rea

The evaluation of plotastic within this thesis reflects its potential to address key challenges in the field of data analysis. The software integrates a comprehensive suite of statistical tests, such as ANOVA and t-tests, designed for adaptability and ease of use, leveraging the functionalities of pingouin.

In the context of the reproducibility crisis in scientific research, plotastic offers noteworthy contributions, though it is not positioned as a universal remedy. The tool's unique approach to integrating statistical analysis with visual representation establishes a new paradigm, promoting methodological transparency. By mandating that statistical analyses accompany relevant graphical outputs, plotastic ensures that analyses are not only conducted with proper scientific rigor but also documented in a manner that facilitates replication, provided the user possesses proficiency in Python.

Usability is a critical attribute of analytical software, particularly as researchers confront increasingly complex datasets. While the developer's intimate familiarity with plotastic may bias perceptions of its ease of use, it is recognized that novices may initially encounter challenges. Nevertheless, plotastic is distinguished by its user-friendly interface, enabling users with minimal statistical training to perform sophisticated analyses by intuitively mapping plotting concepts to statistical operations.

The transition to a new analytical framework, especially one that incorporates coding, presents a learning curve. However, the advantages of plotastic in terms of analytical clarity, speed, and depth are anticipated to outweigh these initial challenges. Support mechanisms, such as assistance from advanced AI like ChatGPT, are available to mitigate these hurdles, supporting users across varying levels of expertise.

In conclusion, plotastic is posited as a valuable tool in the landscape of scientific research, offering a means to enhance the reproducibility and efficiency of data analysis. Its development ethos emphasizes simplifying complex analytical tasks, thereby contributing to the broader goal of fostering transparent and reproducible research practices.

Summarising Discussion

Time Lapse

Lore Ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lore ipsum dolor sit amet. Lore ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lore ipsum

Myeloma

Lore Ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lore ipsum dolor sit amet. Lore ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lore ipsum

Semi-Automated Analysis Improves Agility During Establishing new *in vitro* Methods

Was plotastic useful for me? - Yes incredibly. I was able to perform the statistical tests and visualize the data in a fraction of the time that I would have needed manually. This allowed me to focus on the interpretation of the results and the writing of the manuscript. There was one particular example where my analysis was so fast, that I fed raw datatables during microscopy into python scripts and was able to adapt the experimental technique during the experiment. This allows for an agile and adaptive work environment that is not possible with manual analysis and proved invaluable during development of *in vitro* methods. - These experiments benefited from the use of plotastic, as the

Further research is needed to assess the true impact of semi-automated analysis on the agility

of establishing new *in vitro* methods.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Zheng, X. (2016, March). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems* (No. arXiv:1603.04467). arXiv. Retrieved 2024-03-07, from <http://arxiv.org/abs/1603.04467> doi: 10.48550/arXiv.1603.04467
- Alcorta-Sevillano, N., Macías, I., Infante, A., & Rodríguez, C. I. (2020, December). Deciphering the Relevance of Bone ECM Signaling. *Cells*, 9(12), 2630. Retrieved 2023-12-20, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7762413/> doi: 10.3390/cells9122630
- Armstrong, R. A. (2014, September). When to use the Bonferroni correction. *Ophthalmic & Physiological Optics: The Journal of the British College of Ophthalmic Opticians (Optometrists)*, 34(5), 502–508. doi: 10.1111/opo.12131
- Barzilay, R., Ben-Zur, T., Bulvik, S., Melamed, E., & Offen, D. (2009, May). Lentiviral delivery of LMX1a enhances dopaminergic phenotype in differentiated human bone marrow mesenchymal stem cells. *Stem cells and development*, 18(4), 591–601. doi: 10.1089/scd.2008.0138
- Begley, C. G., & Ioannidis, J. P. A. (2015, January). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research*, 116(1), 116–126. doi: 10.1161/CIRCRESAHA.114.303819
- Bianco, P. (2014). "Mesenchymal" stem cells. *Annual review of cell and developmental biology*, 30, 677–704. doi: 10.1146/annurev-cellbio-100913-013132
- Bladé, J., Beksac, M., Caers, J., Jurczyszyn, A., von Lilienfeld-Toal, M., Moreau, P., ... Richardson, P. (2022, March). Extramedullary disease in multiple myeloma: A systematic literature review. *Blood Cancer Journal*, 12(3), 1–10. Retrieved 2023-03-24, from <https://www.nature.com/articles/s41408-022-00643-3> doi: 10.1038/s41408-022-00643-3
- Bondi, A. B. (2000, September). Characteristics of scalability and their impact on performance. In *Proceedings of the 2nd international workshop on Software and performance* (pp. 195–203). New York, NY, USA: Association for Computing Machinery. Retrieved 2024-03-07, from <https://dl.acm.org/doi/10.1145/350391.350432> doi: 10.1145/350391.350432
- Boswell, D., & Foucher, T. (2011). *The Art of Readable Code: Simple and Practical Techniques for Writing Better Code*. "O'Reilly Media, Inc."
- Brooke, J. (1996, January). SUS – a quick and dirty usability scale. In (pp. 189–194).
- Bustin, S. A. (2014, December). The reproducibility of biomedical research: Sleepers awake! *Biomolecular Detection and Quantification*, 2, 35–42. Retrieved 2024-03-18, from <https://www.sciencedirect.com/science/article/pii/S2214753515000030> doi: 10.1016/j.bdq.2015.01.002
- Caplan, A. (1991). Mesenchymal stem cells. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*, 9(5), 641–50. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1870029> doi: 10.1002/jor.1100090504
- Caplan, A. I. (1994, July). The mesengenic process. *Clinics in plastic surgery*, 21(3), 429–435.
- Chacon, S., & Straub, B. (2024, March). *Git - Book*. Retrieved 2024-03-07, from <https://git-scm.com/book/de/v2>
- Charlier, F., Weber, M., Izak, D., Harkin, E., Magnus, M., Lalli, J., ... Repplinger, S. (2022, October). *Trevis-md/statannotations: V0.5*. Zenodo. Retrieved 2023-11-16, from <https://zenodo.org/record/7213391> doi: 10.5281/ZENODO.7213391
- Cooper, G. M. (2000). The Cell: A Molecular Approach. 2nd Edition. *Sinauer Associates*, Proliferation in Development and Differentiation. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK9906/>
- da Silva Meirelles, L., Chagastelles, P. C., & Nardi, N. B. (2006, June). Mesenchymal stem cells reside in virtually all post-natal organs and tissues. *Journal of cell science*, 119(Pt 11), 2204–2213. doi: 10.1242/jcs.02932

- Dominici, M., Le Blanc, K., Mueller, I., Slaper-Cortenbach, I., Marini, F., Krause, D., ... Horwitz, E. (2006). Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*, 8(4), 315–317. doi: 10.1080/14653240600855905
- Duvall, P., Matyas, S., & Glover, A. (2007). *Continuous integration: Improving software quality and reducing risk*. Pearson Education. Retrieved from <https://books.google.de/books?id=PV9qfEdv9L0C>
- Ekmekci, B., McAnany, C. E., & Mura, C. (2016, July). An Introduction to Programming for Bioscientists: A Python-Based Primer. *PLOS Computational Biology*, 12(6), e1004867. Retrieved 2024-03-10, from <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004867> doi: 10.1371/journal.pcbi.1004867
- Fazeli, P. K., Horowitz, M. C., MacDougald, O. A., Scheller, E. L., Rodeheffer, M. S., Rosen, C. J., & Klibanski, A. (2013, March). Marrow Fat and Bone—New Perspectives. *The Journal of Clinical Endocrinology and Metabolism*, 98(3), 935–945. Retrieved 2023-12-20, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3590487/> doi: 10.1210/jc.2012-3634
- Friedenstein, A., & Kuralesova, A. I. (1971, August). Osteogenic precursor cells of bone marrow in radiation chimeras. *Transplantation*, 12(2), 99–108.
- Friedenstein, A. J., Piatetzky-Shapiro, I. I., & Petrakova, K. V. (1966, December). Osteogenesis in transplants of bone marrow cells. *Journal of embryology and experimental morphology*, 16(3), 381–390.
- Gabr, M. M., Zakaria, M. M., Refaie, A. F., Ismail, A. M., Abou-El-Mahasen, M. A., Ashamallah, S. A., ... Ghoneim, M. A. (2013). Insulin-producing cells from adult human bone marrow mesenchymal stem cells control streptozotocin-induced diabetes in nude mice. *Cell transplantation*, 22(1), 133–145. doi: 10.3727/096368912X647162
- García-Ortiz, A., Rodríguez-García, Y., Encinas, J., Maroto-Martín, E., Castellano, E., Teixidó, J., & Martínez-López, J. (2021, January). The Role of Tumor Microenvironment in Multiple Myeloma Development and Progression. *Cancers*, 13(2). Retrieved 2021-02-02, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7827690/> doi: 10.3390/cancers13020217
- Glavey, S. V., Naba, A., Manier, S., Clouser, K., Tahri, S., Park, J., ... Ghobrial, I. M. (2017, November). Proteomic characterization of human multiple myeloma bone marrow extracellular matrix. *Leukemia*, 31(11), 2426–2434. Retrieved 2023-09-05, from <https://www.nature.com/articles/leu2017102> doi: 10.1038/leu.2017.102
- Gomez-Cabrero, D., Abugessaisa, I., Maier, D., Teschendorff, A., Merkenschlager, M., Gisel, A., ... Tegnér, J. (2014, March). Data integration in the era of omics: Current and future challenges. *BMC Systems Biology*, 8(2), I1. Retrieved 2024-03-18, from <https://doi.org/10.1186/1752-0509-8-S2-I1> doi: 10.1186/1752-0509-8-S2-I1
- Gómez-López, G., Dopazo, J., Cigudosa, J. C., Valencia, A., & Al-Shahrour, F. (2019, May). Precision medicine needs pioneering clinical bioinformaticians. *Briefings in Bioinformatics*, 20(3), 752–766. doi: 10.1093/bib/bbx144
- Goodman, S. N., Fanelli, D., & Ioannidis, J. P. A. (2016, June). What does research reproducibility mean? *Science Translational Medicine*, 8(341), 341ps12-341ps12. Retrieved 2024-03-18, from <https://www.science.org/doi/10.1126/scitranslmed.aaf5027> doi: 10.1126/scitranslmed.aaf5027
- Gosselin, R.-D. (2021, February). Insufficient transparency of statistical reporting in preclinical research: A scoping review. *Scientific Reports*, 11(1), 3335. Retrieved 2024-03-11, from <https://www.nature.com/articles/s41598-021-83006-5> doi: 10.1038/s41598-021-83006-5
- Gronthos, S., Graves, S. E., Ohta, S., & Simmons, P. J. (1994, December). The STRO-1+ fraction of adult human bone marrow contains the osteogenic precursors. *Blood*, 84(12), 4164–4173.
- Hunter, J. D. (2007, May). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*,

- 9(3), 90–95. Retrieved 2023-11-15, from <https://ieeexplore.ieee.org/document/4160265> doi: 10.1109/MCSE.2007.55
- Incerti, D., Thom, H., Baio, G., & Jansen, J. P. (2019, May). R You Still Using Excel? The Advantages of Modern Software Tools for Health Technology Assessment. *Value in Health*, 22(5), 575–579. Retrieved 2024-03-11, from <https://www.sciencedirect.com/science/article/pii/S1098301519300506> doi: 10.1016/j.jval.2019.01.003
- Jansen, B. J. H., Gilissen, C., Roelofs, H., Schaap-Oziemlak, A., Veltman, J. A., Raymakers, R. A. P., ... Adema, G. J. (2010, April). Functional differences between mesenchymal stem cell populations are reflected by their transcriptome. *Stem cells and development*, 19(4), 481–490. doi: 10.1089/scd.2009.0288
- Kazman, R., Bianco, P., Ivers, J., & Klein, J. (2020, December). *Maintainability* (Report). Carnegie Mellon University. Retrieved 2024-03-07, from <https://kilthub.cmu.edu/articles/report/Maintainability/12954908/1> doi: 10.1184/R1/12954908.v1
- Kibler, C., Schermutzki, F., Waller, H. D., Timpl, R., Müller, C. A., & Klein, G. (1998, June). Adhesive interactions of human multiple myeloma cell lines with different extracellular matrix molecules. *Cell Adhesion and Communication*, 5(4), 307–323. doi: 10.3109/15419069809040300
- Krekel, H., Oliveira, B., Pfannschmidt, R., Bruynooghe, F., Laugher, B., & Bruhin, F. (2004). *Pytest*. Retrieved from <https://github.com/pytest-dev/pytest>
- McKinney, W. (2010, January). Data Structures for Statistical Computing in Python. In (pp. 56–61). doi: 10.25080/Majora-92bf1922-00a
- Moreno-Indias, I., Lahti, L., Nedyalkova, M., Elbere, I., Roshchupkin, G., Adilovic, M., ... Zomer, A. L. (2021, February). Statistical and Machine Learning Techniques in Human Microbiome Studies: Contemporary Challenges and Solutions. *Frontiers in Microbiology*, 12. Retrieved 2024-03-18, from <https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.635781/full> doi: 10.3389/fmicb.2021.635781
- Muruganandan, S., Roman, A. A., & Sinal, C. J. (2009, January). Adipocyte differentiation of bone marrow-derived mesenchymal stem cells: Cross talk with the osteoblastogenic program. *Cellular and molecular life sciences : CMLS*, 66(2), 236–253. doi: 10.1007/s00018-008-8429-z
- Myers, G. J., Sandler, C., & Badgett, T. (2011). *The art of software testing* (3rd ed.). Wiley Publishing. Retrieved from <https://malenezi.github.io/malenezi/SE401/Books/114-the-art-of-software-testing-3-edition.pdf>
- Narzt, W., Pichler, J., Pirklbauer, K., & Zwinz, M. (1998, January). A Reusability Concept for Process Automation Software..
- Nowotschin, S., & Hadjantonakis, A.-K. (2010, August). Cellular dynamics in the early mouse embryo: From axis formation to gastrulation. *Current opinion in genetics & development*, 20(4), 420–427. doi: 10.1016/j.gde.2010.05.008
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... Chintala, S. (2019, December). *PyTorch: An Imperative Style, High-Performance Deep Learning Library* (No. arXiv:1912.01703). arXiv. Retrieved 2024-03-07, from <http://arxiv.org/abs/1912.01703> doi: 10.48550/arXiv.1912.01703
- Peng, R. D. (2011, December). Reproducible Research in Computational Science. *Science*, 334(6060), 1226–1227. Retrieved 2024-03-18, from <https://www.science.org/doi/10.1126/science.1213847> doi: 10.1126/science.1213847
- Perneger, T. V. (1998, April). What's wrong with Bonferroni adjustments. *BMJ : British Medical Journal*, 316(7139), 1236–1238. Retrieved 2021-11-24, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1112991/>
- Pittenger, M. F., Mackay, A. M., Beck, S. C., Jaiswal, R. K., Douglas, R., Mosca, J. D., ... Marshak, D. R.

- (1999). Multilineage Potential of Adult Human Mesenchymal Stem Cells. , 284(April), 143–148. doi: 10.1126/science.284.5411.143
- The Python Language Reference.* (n.d.). Retrieved 2024-03-07, from <https://docs.python.org/3/reference/index.html>
- R Core Team. (2018). *R: A language and environment for statistical computing* [Manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners.. Retrieved 2024-03-07, from <https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe>
- Rajkumar, S. V., & Kumar, S. (2020, September). Multiple myeloma current treatment algorithms. *Blood Cancer Journal*, 10(9), 94. Retrieved 2023-07-03, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7523011/> doi: 10.1038/s41408-020-00359-2
- Rayhan, A., & Gross, D. (2023). *The Rise of Python: A Survey of Recent Research.* doi: 10.13140/RG.2.2.27388.92809
- Sacchetti, B., Funari, A., Remoli, C., Giannicola, G., Kogler, G., Liedtke, S., ... Bianco, P. (2016). No identical "mesenchymal stem cells" at different times and sites: Human committed progenitors of distinct origin and differentiation potential are incorporated as adventitial cells in microvessels. *Stem Cell Reports*, 6(6), 897–913. Retrieved from <http://dx.doi.org/10.1016/j.stemcr.2016.05.011> doi: 10.1016/j.stemcr.2016.05.011
- Sandve, G. K., Nekrutenko, A., Taylor, J., & Hovig, E. (2013, October). Ten Simple Rules for Reproducible Computational Research. *PLoS Computational Biology*, 9(10), e1003285. Retrieved 2024-03-07, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3812051/> doi: 10.1371/journal.pcbi.1003285
- Shenghui, H., Nakada, D., & Morrison, S. J. (2009). Mechanisms of Stem Cell Self-Renewal. *Annual Review of Cell and Developmental Biology*, 25(1), 377–406. Retrieved from <https://doi.org/10.1146/annurev.cellbio.042308.113248> doi: 10.1146/annurev.cellbio.042308.113248
- Smith, A. M., Niemeyer, K. E., Katz, D. S., Barba, L. A., Githinji, G., Gymrek, M., ... Vanderplas, J. T. (2018). Journal of Open Source Software (JOSS): Design and first-year review. *PeerJ Preprints*, 4, e147. doi: 10.7717/peerj-cs.147
- Stock, P., Bruckner, S., Winkler, S., Dollinger, M. M., & Christ, B. (2014, April). Human bone marrow mesenchymal stem cell-derived hepatocytes improve the mouse liver after acute acetaminophen intoxication by preventing progress of injury. *International journal of molecular sciences*, 15(4), 7004–7028. doi: 10.3390/ijms15047004
- Tam, P. P., & Beddington, R. S. (1987, January). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development (Cambridge, England)*, 99(1), 109–126.
- Tanavalee, C., Luksanapruksa, P., & Singhatanadighe, W. (2016, June). Limitations of Using Microsoft Excel Version 2016 (MS Excel 2016) for Statistical Analysis for Medical Research. *Clinical Spine Surgery*, 29(5), 203. Retrieved 2024-03-11, from https://journals.lww.com/jspinaldisorders/fulltext/2016/06000/limitations_of_using_microsoft_excel_version_2016.5.aspx doi: 10.1097/BSD.0000000000000382
- Team, T. P. D. (2020, February). *Pandas-dev/pandas: Pandas.* Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.3509134> doi: 10.5281/zenodo.3509134
- Terpos, E., Ntanasis-Stathopoulos, I., Gavriatopoulou, M., & Dimopoulos, M. A. (2018, January). Pathogenesis of bone disease in multiple myeloma: From bench to bedside. *Blood Cancer Journal*, 8(1), 7. doi: 10.1038/s41408-017-0037-4
- Ullah, I., Subbarao, R. B., & Rho, G. J. (2015). Human mesenchymal stem cells - current trends and future prospective Bioscience Reports. doi: 10.1042/BSR20150025

- Vallat, R. (2018, November). Pingouin: Statistics in Python. *Journal of Open Source Software*, 3(31), 1026. Retrieved 2023-05-29, from <https://joss.theoj.org/papers/10.21105/joss.01026> doi: 10.21105/joss.01026
- van Rossum, G., Lehtosalo, J., & Langa, L. (2014). PEP 484 – Type Hints / peps.python.org. Retrieved 2024-03-08, from <https://peps.python.org/pep-0484/>
- Viguet-Carrin, S., Garnero, P., & Delmas, P. D. (2006, March). The role of collagen in bone strength. *Osteoporosis International*, 17(3), 319–336. Retrieved 2023-12-20, from <https://doi.org/10.1007/s00198-005-2035-9> doi: 10.1007/s00198-005-2035-9
- Waskom, M. L. (2021, April). Seaborn: Statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. Retrieved 2023-03-26, from <https://joss.theoj.org/papers/10.21105/joss.03021> doi: 10.21105/joss.03021
- Wickham, H. (2014, September). Tidy Data. *Journal of Statistical Software*, 59, 1–23. Retrieved 2023-11-15, from <https://doi.org/10.18637/jss.v059.i10> doi: 10.18637/jss.v059.i10
- Wilkins, A., Kemp, K., Ginty, M., Hares, K., Mallam, E., & Scolding, N. (2009, July). Human bone marrow-derived mesenchymal stem cells secrete brain-derived neurotrophic factor which promotes neuronal survival in vitro. *Stem cell research*, 3(1), 63–70. doi: 10.1016/j.scr.2009.02.006
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016, March). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. Retrieved 2024-03-18, from <https://www.nature.com/articles/sdata201618> doi: 10.1038/sdata.2016.18
- Xu, W., Zhang, X., Qian, H., Zhu, W., Sun, X., Hu, J., ... Chen, Y. (2004, July). Mesenchymal stem cells from adult human bone marrow differentiate into a cardiomyocyte phenotype in vitro. *Experimental biology and medicine (Maywood, N.J.)*, 229(7), 623–631.
- Yang, A., Troup, M., & Ho, J. W. (2017, July). Scalability and Validation of Big Data Bioinformatics Software. *Computational and Structural Biotechnology Journal*, 15, 379–386. Retrieved 2024-03-07, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5537105/> doi: 10.1016/j.csbj.2017.07.002

Appendix

Supplementary Figures and Methods

Keep it Together: Describing Myeloma Dissemination *in vitro* with hMSC-Interacting Subpopulations and their Aggregation/Detachment Dynamics

Martin Kuric¹, Susanne Beck², Doris Schneider¹, Wyonna Rindt³, Marietheres Evers⁴, Jutta Meißner-Weigl¹, Sabine Zeck¹, Melanie Krug¹, Marietta Herrmann⁵, Tanja Nicole Hartmann⁶, Ellen Leich⁴, Maximilian Rudert⁷, Denitsa Docheva¹, Anja Seckinger⁸, Dirk Hose⁸, Franziska Jundt³, Regina Ebert¹

¹University of Würzburg, Department of Musculoskeletal Tissue Regeneration, Würzburg, Germany

²University Hospital Heidelberg, Institute of Pathology, Heidelberg, Germany

³University Hospital Würzburg, Department of Internal Medicine II, Würzburg, Germany

⁴University of Würzburg, Institute of Pathology, Comprehensive Cancer Center Mainfranken, Würzburg, Germany

⁵University Hospital Würzburg, IZKF Research Group Tissue Regeneration in Musculoskeletal Diseases, Würzburg, Germany

⁶University of Freiburg, Department of Internal Medicine I, Faculty of Medicine and Medical Center, Freiburg, Germany

⁷University of Würzburg, Orthopedic Department, Clinic König-Ludwig-Haus, Würzburg, Germany

⁸Vrije Universiteit Brussel, Department of Hematology and Immunology, Jette, Belgium

Tab. S1: List of hMSC donors, myeloma cell lines, and their mycoplasma test status. If no unique donors were available, hMSC donors were used twice for the same experiment at different passages. WPSC: Well plate sandwich centrifugation.

Cell Type	Donor / Line	Donor Ages	Donor Sex	Date of negative Mycoplasma test	Experiment(s)	Figures
Myeloma Cell Line	INA-6	80	m	09.02.22	All	All
	U266			10.10.22	- Validation of V-Well Adhesion Assay	S1E
	MM1.S			24.02.22		
hMSC	1639	49	m	not tested	- Validation of V-Well Adhesion Assay	S1E
	1571	72	m	not tested	- Time-lapse: INA-6 on dispersed hMSC	1D; 2[A-E]
	1573	47	m	not tested	- Saturation of hMSCs	1[A-B]
	1578	82	m	not tested		
	1842	63	m	not tested	- INA-6 Viability dep. on time and hMSC adhesion surface (INA not washed off)	1E right
	1843	60	m	not tested		
	1537	77	f	not tested		
	1794	82	m	not tested		
	1779	61	m	not tested	- INA-6 Viability dep. on time and hMSC adhesion surface (INA washed off)	1[C, E left]
	1849	69	m	not tested		
	1854	80	f	not tested		
	1605	71	f	not tested	- Time-lapse: INA-6 on dispersed hMSC	1D; 2[A-E]
	1650	57	m	not tested		
	1859	64	f	not tested	- Time-lapse: INA-6 on confluent hMSC	2[G-I]
	1863	79	f	not tested		
	1861	52	f	not tested		
	1818	81	f	not tested	- Cell Cycle Profiling after V-well assay	3C
	1824	82	f	not tested	(Donor measured twice, different passages) - V-well adhesion assay of mitotically blocked INA-6 followed by Cell Cycle Profiling after V-well assay	3[B,C]
	1827	56	m	not tested	- V-well adhesion assay of mitotically blocked INA-6 followed by Cell Cycle Profiling after V-well assay	

				assay	
1501	59	m	not tested	- INA-6 AI-assisted count during WPSC (INA-6 stained with celltracker green)	4B
1643	75	f	not tested		
1718	67	m	not tested		
1720	58	m	not tested		
1653	65	m	not tested		
1591	78	m	not tested	- WPSC (MACS) followed by RNAseq, Metascape analysis and qPCR validation - WPSC (Wash) followed by qPCR-Validation and Luminescent Viability assays	4[A,C,D,E] ; 5[A-C] 4[C-E], 4F
1654	74	m	not tested	- WPSC (MACS) followed by RNAseq, Metascape analysis and qPCR validation - WPSC (Wash) followed by qPCR-Validation and Luminescent Viability assays	4[A,C,D,E] ; 5[A-C] 4[C-E], 4F
1655	78	f	not tested	- WPSC (MACS) followed by RNAseq, Metascape analysis and qPCR validation	4[A,C,D,E] ; 5[A-C]
1668	80	f	not tested		
1670	66	f	not tested		
1701	81	m	not tested	- WPSC (Wash) followed by qPCR-Validation and Luminescent Viability assays	4[C-E], 4F
1702	79	f	not tested		
1600	77	m	not tested		
1681	56	m	not tested	- WPSC (Wash) followed by Luminescent Viability assays	4F
1672	65	m	not tested	- WPSC (Wash) followed by qPCR-Validation	4[C-E]

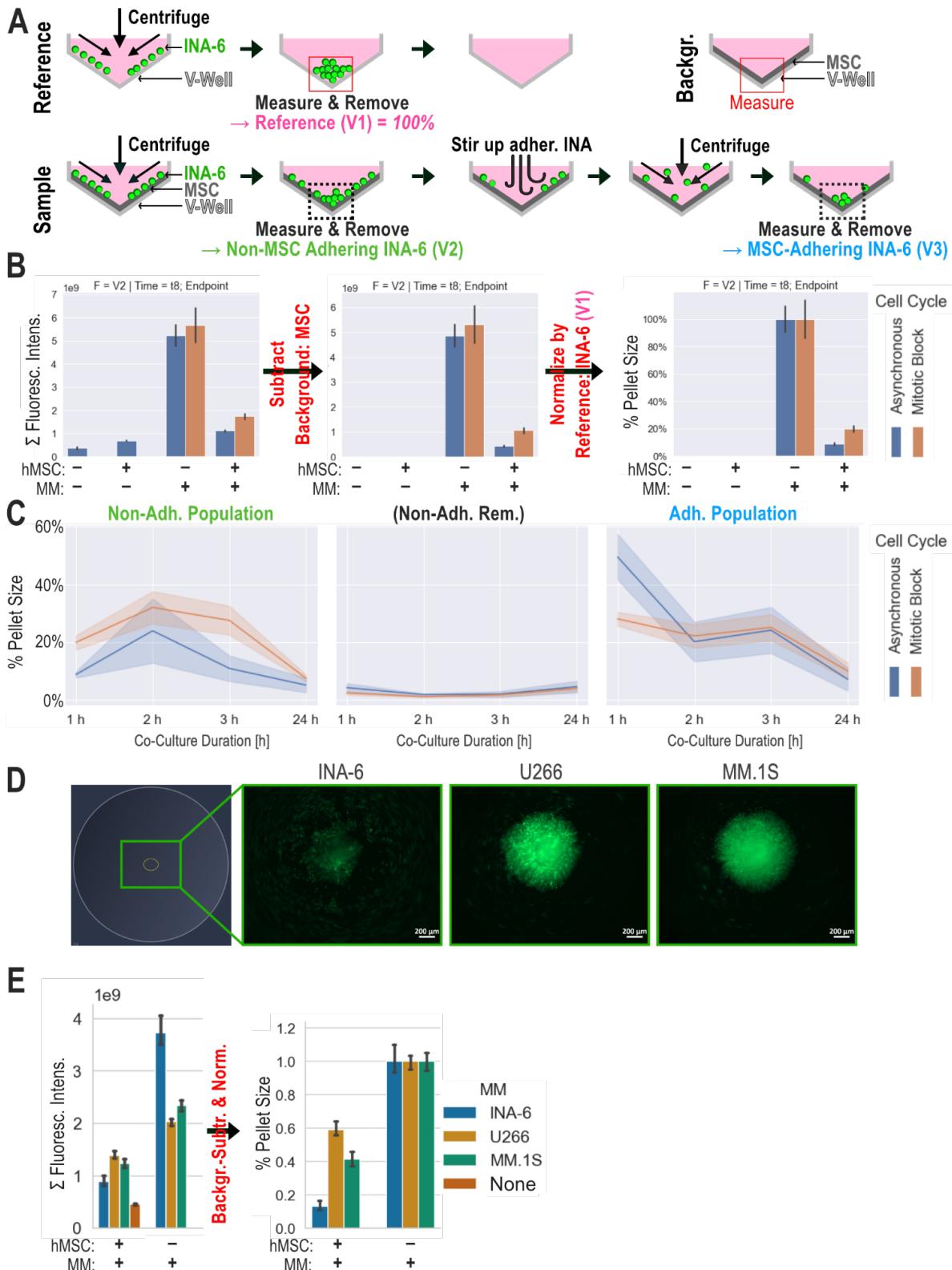


Fig. S1: Principle and quantification of the V-well adhesion assay of fluorescently labeled myeloma cells adapted by Weetall et al. 2001. **A:** Sample: Subsequent rounds of centrifugation and removal of cell pellet yielded the size of adhesive subpopulations. Fluorescently stained INA-6 cells were added to an hMSC monolayer. Non-adherent INA-6 cells (V2) were pelleted in the well-tip. Pellets were quantified by fluorescence brightness and isolated by pipetting. Immobile INA-6 cells (V3) were manually detached by forceful pipetting.

Reference: Omitting adhesive hMSC-layer yielded ~100% non-adherent cells (V1) after the first centrifugation step; Background: hMSC monolayer was used as background signal. **B:** Calculation of the population size relative to total cells starting with pellet intensity. The shown example is the pellet gained by centrifuging mobile subpopulation (V2) after 1 h of co-culture. (see Fig. 2 for context): Intensity values from pellet images were summarized. After subtracting the unlabeled hMSC signal and normalization by a full-size pellet (reference), the resulting values represented the fraction of the adhesive subpopulation. **C:** One of three biological replicates summarized in Fig. 2. Line range shows the standard deviation of four technical replicates. Non.Adh. Rem.: Fluorescence signal after removal of V2. **D:** Example images of myeloma cell lines (INA-6, U266, MM.1S) pelleted in the tip of V-wells. The leftmost image shows the recorded area in a complete V-well. Scale bar = 200 μ m. **E:** Results from (D) comparing adhesion strength of three myeloma cell lines to hMSC. Error bars represent technical deviation. MM=Multiple Myeloma.

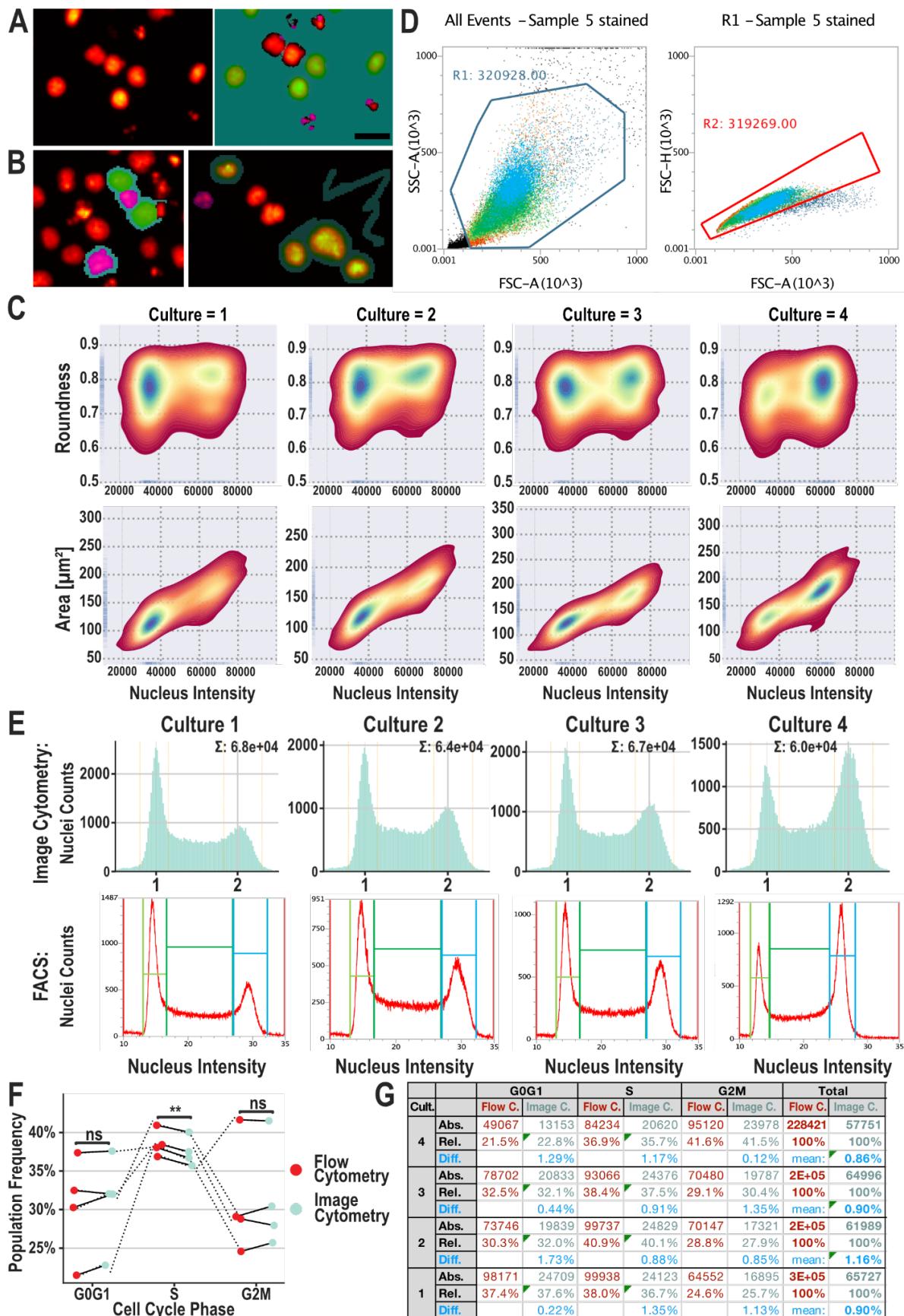


Fig. S2: Validation of image cytometric analysis of cell cycle in four INA-6 cultures **A:** Left: Example image cytometric scan: INA-6 cells were stained with Hoechst33342 and scanned by automated fluorescence microscopy. Right: The image was segmented using a convolutional neural network (ZEISS ZEN intellesis) trained to discern healthy nuclei (green) from fragmented ones (magenta). Doublets are excluded by setting an area- and roundness threshold. Scale bar: 20 μ m. **B:** Two example images from the training set. **C:** Quality of image cytometric data was ensured by plotting the distribution of nuclei brightnesses vs. the distribution of both nuclei-roundnesses and nuclei-areas. Nuclei with double fluorescence intensity have the same roundness while their area increases, as expected from a cell in G2 phase. **D:** The same samples from (C) were also measured with flow cytometry. Representative example of gating strategy: Left: Dead cells were excluded by setting a minimum threshold for side-scattering (SSC-A). Right: Doublets were excluded by setting a maximum threshold for forward scatter area (FSC-A) (sample “5” represents culture “4” in this figure). **E:** Cell cycle profiles of four independent INA-6 cultures were measured by both image cytometry (top) and flow cytometry (bottom). For both methods, frequencies of G0/G1, S, and G2M were summed up by setting fluorescence intensity thresholds. **F:** Image cytometry yields the same frequencies for G0/G1, S, and G2M when compared to flow cytometry. RM-ANOVA showed that the method has no significant effect on the frequencies of cell cycle populations [$F(1,3)=1.421$, $p\text{-unc}=.32$]. **G:** Results from (F) in tabular form. On average, frequencies for G0/G1, S, and G2M measured by Image cytometry differ by 0.95 percent points compared to flow cytometry measurement. Cult.: Culture; C.: Image cytometry; Abs.: Absolute cell count; Rel.: Relative cell count; Diff.: Difference between relative cell counts determined by flow cytometry and image cytometry.

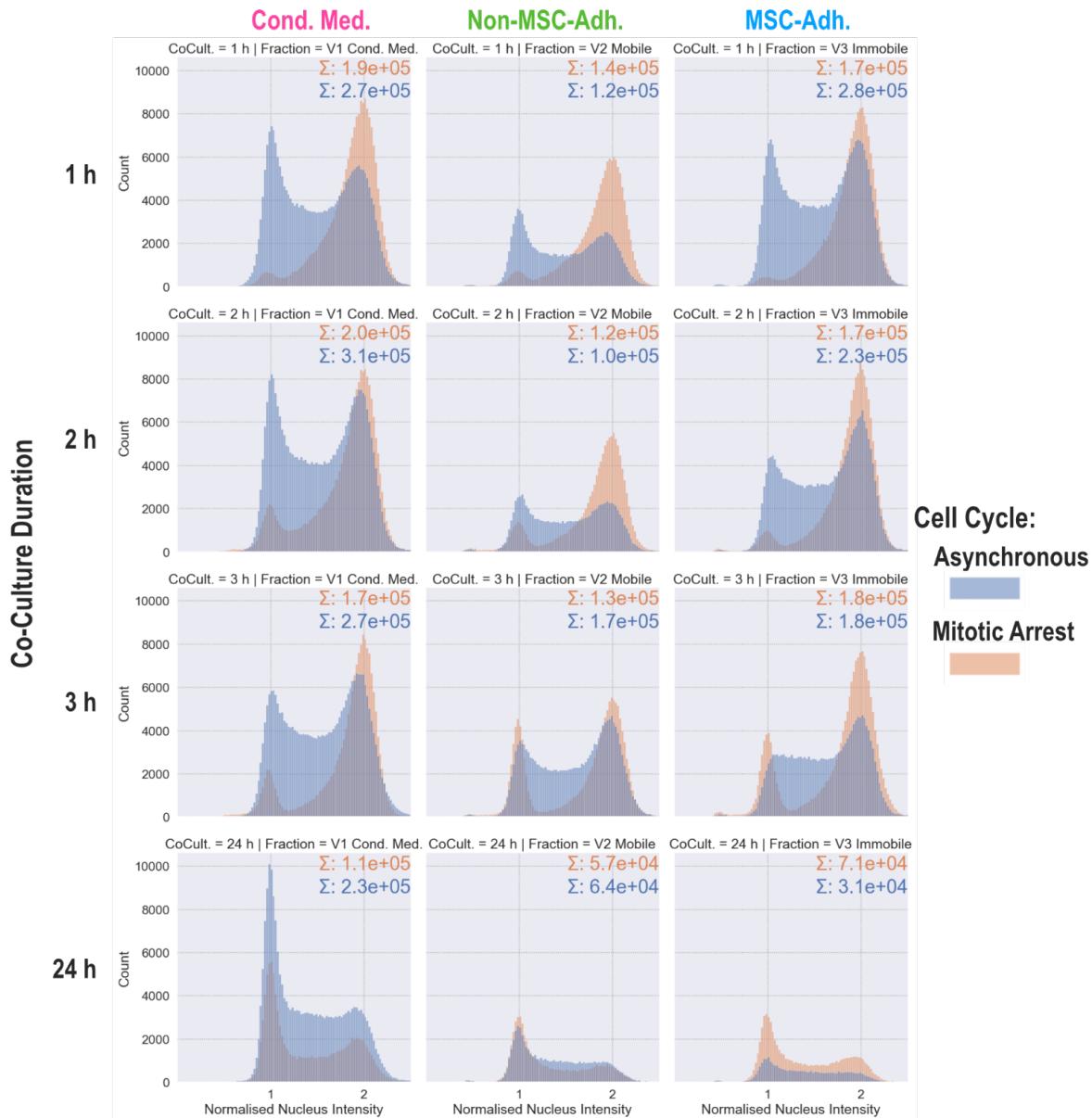
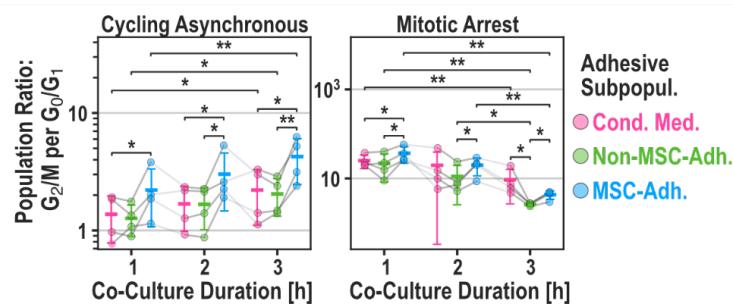
A**B**

Fig. S3: Cell cycle analysis of INA-6 pellets gained from V-Well Adhesion assay (Fig. 3). **A:** Cell cycle profiles of MSC-adhering subpopulations. INA-6 cells were synchronized by double thymidine block followed by nocodazole. Cell cycle was released directly before addition to hMSCs. Histograms were normalized and summed up across all biological replicates ($n=4$). Technical replicates (3) were pooled prior to cell cycle profiling. CoCult. = Co-culture duration. Fraction = Adhesion subpopulations. **B:** Similar figure to Fig. 3C displaying ratio of INA-6 populations (G_2/M to G_0/G_1). **Statistics:** Paired t-test (B).

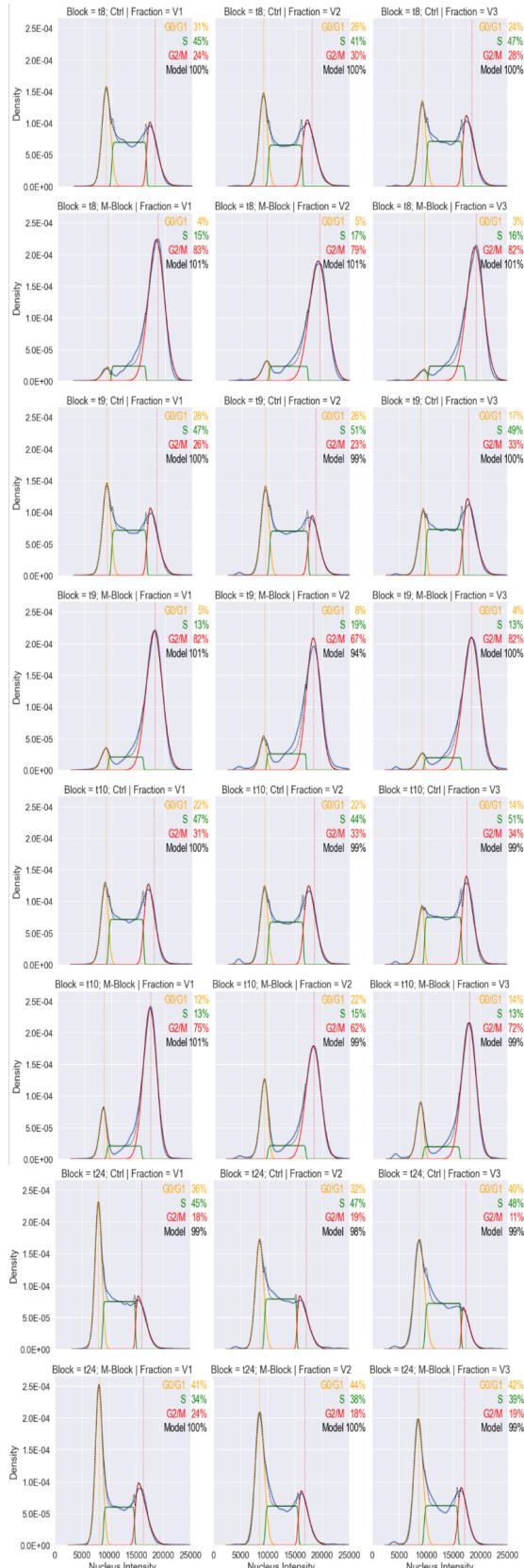


Fig. S4: Representative (one of the four independent sample sets as seen in Fig. S3) curve fitting analysis of cell cycle profiles generated by Image Cytometry. t8, t9, t10, and t24 refer to 1, 2, 3, and 24 hours after the addition of INA-6 cells to hMSCs.

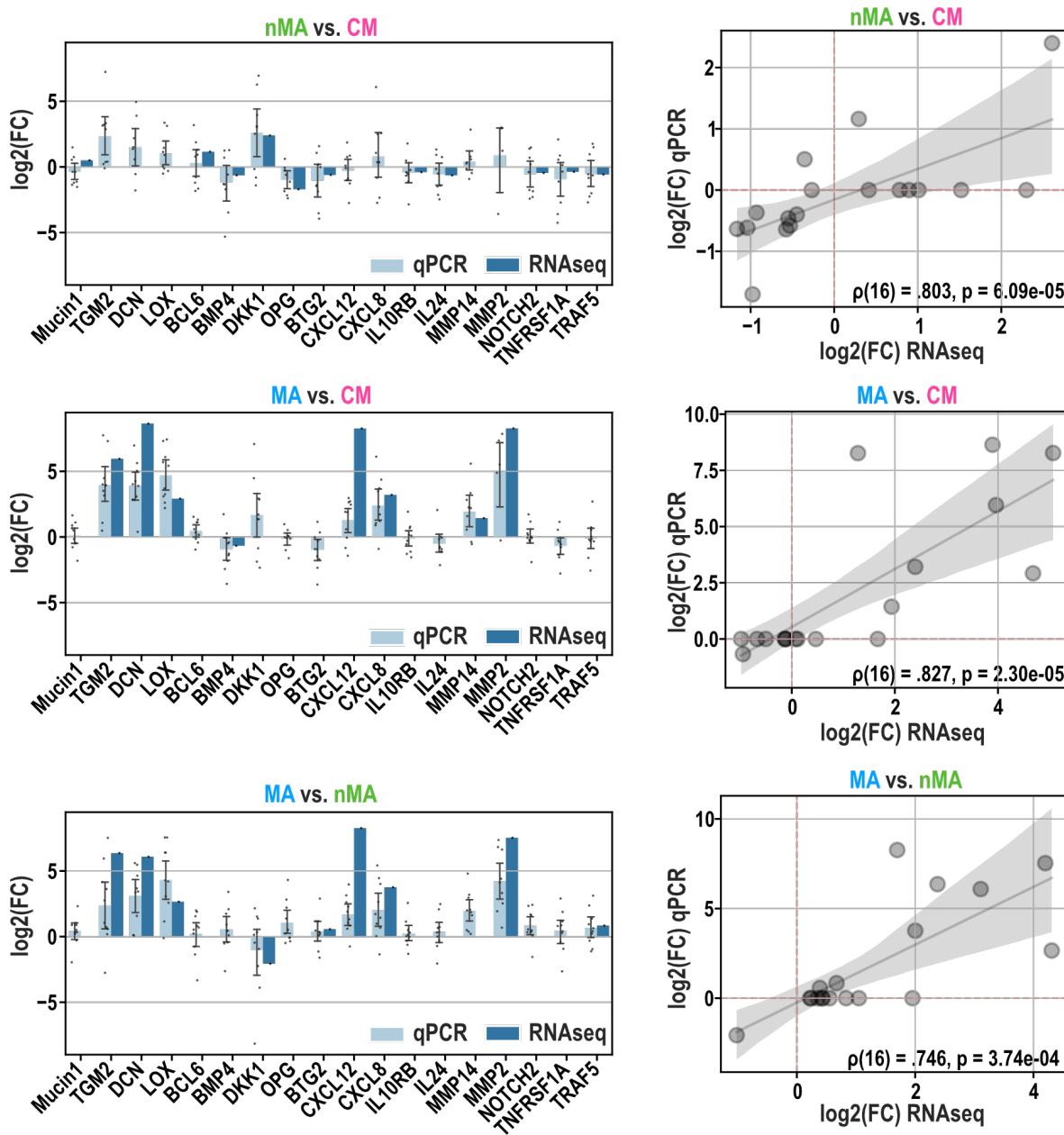


Fig. S5: Correlation of RNAseq with qPCR **Left:** Validation of RNAseq results (Fig. 3) with qPCR showing the log₂(foldchange expression) of 18 genes. For qPCR, Datapoints each represent one biological replicate ($n=10$), which is the mean of technical replicates ($n=3$). Bar height represents mean of biological replicates, error bars show standard deviation of biological replicates. **Right:** Correlation between qPCR and RNAseq in terms of log₂(mean foldchange expression per gene). Each dot represents one gene shown in the barplot to the left. Genes measured with qPCR that showed no differential expression in RNAseq were set to have a log₂(FC) = 0. Shaded area shows the confidence interval of linear regression. Correlation coefficient (ρ) was calculated using Spearman's rank. N = 18 genes. FC = fold change expression.

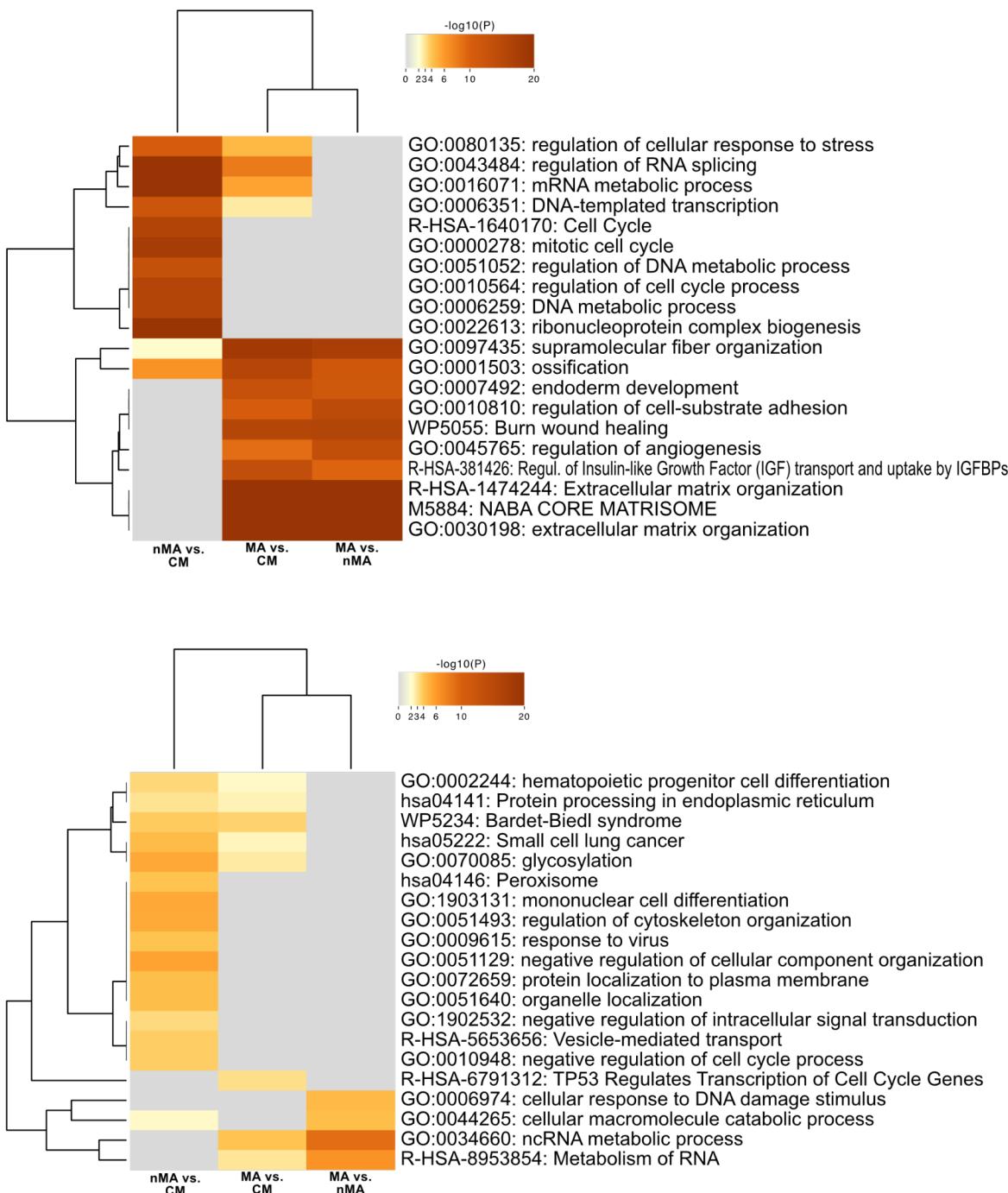


Fig. S6: Functional enrichment analysis by Metascape using genes that are differentially expressed between MSC-interacting subpopulations. **Top:** Upregulated genes. **Bottom:** Downregulated genes.

Tab. S2: Adhesion genes (from Fig. 6A) filtered by association with patient survival ($p < 0.05$) and categorized by a continuous downregulation across disease progression. The full table including non-significant associations is found in the supplementary data. Bone Marrow Plasma Cell (BMPC), Monoclonal Gammopathy of Undetermined Significance (MGUS), Smoldering Multiple Myeloma (sMM), Multiple Myeloma (MM), Multiple Myeloma Relapse (MMR). p-adj. = adjusted p-values (Benj.-Hoch.).

Regulation during disease progression	Gene	Ensemble ID	Progression on Free / Overall Survival	Better Prognosis with high/low expression	Association of expression with survival	
					[p-unc]	[p-adj]
Not downregulated or overall low expression	CCDC80	ENSG00000091986	Prog. Free	high	2.04E-03	1.58E-02
	CCN2	ENSG00000118523	Overall	high	2.89E-03	2.43E-02
	CCNE2	ENSG00000175305	Prog. Free	low	1.21E-02	4.62E-02
			Overall	low	5.34E-04	8.64E-03
	COL4A1	ENSG00000187498	Overall	high	9.47E-03	3.99E-02
	COL4A2	ENSG00000134871	Prog. Free	high	1.24E-02	4.62E-02
	F3	ENSG00000117525	Overall	high	9.18E-03	3.99E-02
	HTRA1	ENSG00000166033	Prog. Free	high	1.20E-02	4.62E-02
	IGFBP7	ENSG00000163453	Prog. Free	low	9.53E-03	4.38E-02
	MMP2	ENSG00000087245	Prog. Free	high	2.29E-05	2.32E-03
	OSMR	ENSG00000145623	Prog. Free	high	5.67E-04	7.15E-03
			Overall	high	1.29E-02	4.64E-02
Continuously downregulated (PC > MGUS > sMM > MM > MMR)	SERPINH1	ENSG00000149257	Prog. Free	low	1.83E-03	1.58E-02
			Overall	low	4.40E-03	2.61E-02
	ACTN1	ENSG00000072110	Overall	high	7.73E-03	3.55E-02
	AEBP1	ENSG00000106624	Prog. Free	high	1.08E-02	4.62E-02
	AXL	ENSG00000167601	Prog. Free	high	1.50E-03	1.51E-02
			Overall	high	3.64E-05	1.84E-03
	COL1A1	ENSG00000108821	Prog. Free	high	3.03E-04	4.37E-03
			Overall	high	5.93E-04	8.64E-03
	COL3A1	ENSG00000168542	Overall	high	1.08E-02	4.29E-02
	COL6A1	ENSG00000142156	Prog. Free	high	1.20E-02	4.62E-02
			Overall	high	1.10E-02	4.29E-02
	CXCL12	ENSG00000107562	Prog. Free	high	1.16E-04	2.93E-03
			Overall	high	6.48E-04	8.64E-03
	CYP1B1	ENSG00000138061	Prog. Free	high	8.64E-03	4.17E-02
			Overall	high	6.84E-04	8.64E-03
	DCN	ENSG00000011465	Prog. Free	high	4.83E-03	3.05E-02
			Overall	high	2.47E-04	8.33E-03
	FBLN1	ENSG00000077942	Prog. Free	high	2.68E-03	1.93E-02
			Overall	high	3.73E-03	2.61E-02
	GNB3	ENSG00000111664	Prog. Free	high	3.75E-03	2.52E-02
			Overall	high	5.73E-03	3.05E-02
	IGFBP4	ENSG00000141753	Prog. Free	high	8.68E-03	4.17E-02
			Overall	high	7.09E-03	3.41E-02
	ITGAX	ENSG00000140678	Prog. Free	high	6.72E-03	3.60E-02
			Overall	high	3.12E-03	2.43E-02
	LAMB2	ENSG00000172037	Overall	high	1.35E-03	1.39E-02
	LRP1	ENSG00000123384	Prog. Free	high	6.46E-03	3.60E-02
			Overall	high	4.34E-04	8.64E-03
	LTBP2	ENSG00000119681	Prog. Free	high	9.03E-05	2.93E-03
			Overall	high	1.17E-02	4.36E-02
	MAP3K8	ENSG00000107968	Prog. Free	high	9.58E-04	1.08E-02
	MFAP5	ENSG00000197614	Prog. Free	high	2.43E-04	4.09E-03
			Overall	high	4.27E-03	2.61E-02
	MMP14	ENSG00000157227	Prog. Free	high	6.93E-05	2.93E-03
			Overall	high	6.69E-03	3.38E-02
	MYL9	ENSG00000101335	Prog. Free	high	1.46E-04	2.95E-03
			Overall	high	1.56E-05	1.57E-03
	NRP1	ENSG00000099250	Prog. Free	high	1.89E-03	1.58E-02
			Overall	high	2.21E-03	2.03E-02
	TGFBI	ENSG00000120708	Overall	high	4.30E-03	2.61E-02
	TNC	ENSG00000041982	Prog. Free	high	1.28E-02	4.62E-02
			Overall	high	4.75E-03	2.67E-02
	TPM1	ENSG00000140416	Overall	high	1.37E-03	1.39E-02
	TUBA1A	ENSG00000167552	Prog. Free	low	6.78E-03	3.60E-02

Supplementary Materials and Methods

Isolation and Culturing of Primary Human Bone Marrow-Derived Mesenchymal Stromal Cells

Primary human MSCs were obtained from the femoral head of patients (S. Tab. 1) undergoing elective hip arthroplasty. Material was collected with the informed consent of all patients and the procedure was approved by the local Ethics Committee of the University of Würzburg (186/18). In brief, bone marrow was washed with MSC-Medium [Dulbecco's modified Eagle's medium (DMEM/F12) (Thermo Fisher Scientific, Darmstadt, Germany) supplemented with 10% Fetal Calf Serum (FCS) (Bio&Sell GmbH, Feucht, Germany (Fernandez-Rebollo et al., 2017), 100 U/ml penicillin, 0.1 mg/ml streptomycin, 50 µg/ml ascorbate and 100 nmol/l sodium selenite (both Sigma-Aldrich GmbH, Munich, Germany)) and centrifuged at 250 g for 5 min. The pellet was washed four times with MSC-medium and resulting supernatants containing released cells were collected. Cells were pelleted and cultured at a density of 1×10^9 cells per 175 cm² culture flask. After two days non-attached cells were washed away and adherent ones were cultivated in MSC-Medium until confluence. Then, they were either frozen in liquid nitrogen or directly utilized for experiments. hMSC cultures were sustained for a maximum of two passages. All cells were cultured at 37 °C and at 5% CO₂.

Culturing of Myeloma Cell Lines

The plasmacytoma cell line INA-6 [RRID:CVCL_5209; DSMZ, Braunschweig, Germany, authenticated by DSMZ 2014 (see supplemental); (Burger et al., 2001; Gramatzki et al., 1994) was cultivated in RPMI 1640 medium (Life Technologies GmbH) supplemented with 20 % (v/v) FCS, 100 µg/ml gentamicin, 2 mmol/l L-glutamine (both Life Technologies GmbH), 1 mmol/l sodium pyruvate, 100 nmol/l sodium selenite (both Sigma Aldrich GmbH, Schnelldorf, Germany) and 2 ng/ml recombinant human interleukin-6 (IL-6; Miltenyi Biotec, Bergisch Gladbach, Germany). INA-6 were passaged three times per week by diluting them to 1×10^5 , 2×10^5 , or 4×10^5 cells/mL for 3, 2 and 1 days of culturing, respectively. MM.1S (RRID:CVCL_8792) (Greenstein et al., 2003), and U266 cells (CVCL_0566), (Nilsson et al., 1970) were propagated and cultivated in RPMI1640 medium comprising 10 % (v/v) FCS, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mmol/l L-glutamine, and 1 mmol/l sodium pyruvate. All cells were cultured at 37 °C and at 5% CO₂.

Co-Culturing of Primary hMSCs and INA-6 and MSC-Conditioning of Medium

For each co-culture, hMSCs were seeded out 24 h prior to INA-6 addition to generate MSC-conditioned medium (CM). CM from different donors was collected separately and used immediately when adding INA-6. To ensure that CM was free of hMSCs, it was strained (40 µm) and centrifuged for 15 minutes at 250 g. INA-6 cells were washed with PBS (5 min, 1200 rpm), resuspended in MSC-medium and added to hMSCs such that co-culture comprised 33% (v/v) of CM gathered directly from the respective hMSC-donor. Co-cultures did not contain IL-6 (Chatterjee et al., 2002).

Collagen I Coating

Collagen I solution (isolated from rat tail, Corning, NY, USA) was diluted 1:2 (75 ng/mL) in acetic acid (0.02 N), applied to 96 well plates (30 µL in each well) and incubated for 2 h at room temperature. Acetic acid was removed and wells were washed once with 100 µL of PBS. Coated plates were stored dry at 4 °C.

Fluorescent Staining of Cells

For each live staining, cells were strained (70 µm) to remove clumps and washed (5 min, 250 g) once with the respective media (without FCS) and then resuspended in staining reagents.

For CellTracker™ Green CMFDA Dye and CellTracker™ Deep Red Dye (Thermo Fisher Scientific) staining, 1 mL staining solution for a maximum of 1×10^6 cells was prepared. Staining was done at RT for 15 min using 5 µM CMFDA (5-Chlormethyl-fluoresceindiacetat) and 5 min of 1-2 µM DeepRed. To reduce background, stained cells were pelleted, resuspended in cell medium (containing FCS), incubated for 30 min (37 °C, 5% CO₂), washed in cell medium, resuspended in 100 µL - 1 mL and counted.

For PKH26 staining (Sigma Aldrich), a maximum of 1×10^4 cells was resuspended in 500 µL diluent C before swiftly adding 500 µL of staining solution (1 µL diluted in 500 µL diluent C) and incubating cells for 5 min at RT. The staining reaction was stopped by adding 1 mL of FCS-containing medium and adding 3 mL of FCS-free medium. Cells were washed with 10 mL of FCS-containing medium, resuspended in 100 µL - 1 mL cell medium, and counted.

For Calcein-AM (Calcein-O,O'-diacetat-tetrakis-(acetoxymethyl)-ester) (Thermo Fisher Scientific) staining, end concentrations of 0.5 µM were used. 12.5 µL of diluted stock solution (2.5 µM) was carefully added to 50 µL of the co-culture and incubated for 10 minutes at 37 °C.

For Hoechst33342 staining, cells were washed once with PBS, resuspended in a maximum of 500 µL of PBS, and fixed with 5 mL of ice-cold ethanol (70% v/v) by vigorously pipetting up and down to dissociate aggregates. Cells were washed once with PBS and stained with 2.5 µg/mL Hoechst33342 (Thermo Fisher Scientific) diluted in PBS for 1 h at 37 °C.

Automated Fluorescence Microscopy

To remove clumps for microscopic applications, we cultured cells in 40 µm strained FCS. To reduce background fluorescence and phototoxicity, we used phenol-red free versions of the respective medium, if available.

All microscopy equipment was acquired from ZEISS. The microscope was an Axio Observer 7 with confocal Apotome.2 equipped with a motorized reflector revolver and motorized scanning table (130x100 mm). The microscope was mounted on an Antivibrations-Set [Axio Observer (D)] with two antivibration carrier plates, each equipped with two vibration dampening feet. The light source was a microLED 2 for transmission light and (for fluorescence) Colibri 7 (R[G/Y]B-UV) for five channels of incident light (385, 475, 555, 590, 630 nm). For excitation (EX) and emission (EM) light filtering and beam splitting (BS) we used the following reflectors: 96 HE BFP shift free (E) (EX: 390/40, BS: 420, EM: 450/40), 43 HE Cy 3 shift free (E) (EX: 550/25, BS: 570, EM: 605/70), 38 HE eGFP shift free (E) (EX: 470/40, BS: 495, EM: 525/50) and 90 HE LED (E) (EX: 385, 475, 555 und 630 nm, BS: 405 + 493 + 575 + 653, EM: 425/30 + 514/30 + 592/30 + 709/100). We used the black and white camera Axiocam 506 mono (D) and if not stated otherwise, 2x2 binning was used for fluorescence imaging. For mosaic acquisitions (“tiles”) we used a tiling overlap of 8-10% and image tiles were not stitched. Images were magnified 5x and 10x (Fluar 5x/0.25 M27 and EC Plan-Neofluar 10x/0.3 Ph1 M27).

Cell Viability and Apoptosis Assay

To examine cell viability and apoptosis, cells were seeded in a 96-well plate (1×10^4 cells per well) to be measured inside culture well after respective incubation time immediately. ATP-amount and Caspase 3/7 activity were used as a proxy for viability and apoptosis rates, respectively. They were assessed using the CellTiter-Glo Luminescent Cell Viability Assay and the Caspase-Glo 3/7 Assay, respectively (Promega GmbH, Mannheim, Germany), according to the manufacturer's instructions.

Luminescence was measured with an Orion II Luminometer (Berthold Detection Systems, Pforzheim, Germany).

Microscopic Characterization of MSC Saturation

For saturating hMSC with INA-6, hMSCs were stained with CellTracker Green, plated out on 384-Well plates (Greiner) at 5×10^3 hMSC/cm² and cultured for 24 h. INA-6 cells were stained with CellTracker DeepRed, resuspended in MSC-medium, added to adhering hMSCs in different amounts (5×10^3 , 1×10^3 , 2×10^3 INA-6/cm²) and co-cultured for 24 h and 48 h. The complete co-culture was scanned and the number of INA-6 cells adhering on one MSC was counted manually for 100 MSCs for each technical replicate. Fluorescent images were digitally re-stained (INA-6 green, hMSC inverse black).

Analysis INA-6 Survival and Aggregation Depending on hMSCs Confluence

To describe aggregate growth and survival of INA-6 depending on hMSC density, unstained hMSCs were seeded out into 96-well plates (white, clear bottom, Greiner) at different densities (Tab. S3). To ensure nutrient supply, we used lower cell densities for longer co-culturing durations while maintaining constant ratios of INA-6 to adhesion surface provided by hMSCs. Those plates that are to be assessed after 72 h of co-culturing received further 100 µL of fresh MSC-medium after 24 h of co-culturing (total volume of 300 µL), and after 48 h of co-culturing, 100 µL was removed gently from the co-culture and (carefully not to stir up co-culture on bottom) replaced with fresh MSC-Medium after 48 h of co-culturing.

To describe aggregate growth, complete wells were scanned using 10x magnification, phase contrast, 2x2 binning, and autofocus focusing on each tile both before and after harvesting. Afterwards, INA-6 cells were harvested for measuring viability and apoptosis.

Tab. S3: Seeding densities for describing growth and survival of INA-6 depending on hMSC density. Co-cult. dur. = Co culturing duration; MSC-adh. surface = adhesion surface provided by hMSCs; vol. = volume.

Co- cult. dur. [h]	hMSC density [1000 hMSC/cm ²]			INA-6 density [1000 INA-6/cm ²]	Ratios INA : MSC (adh. surface)			Seeding vol. [μ L]	End vol. [μ L]
24	2	10	40	10	1 : 0.2	1 : 1	1 : confluent	200	200
48	1	5	40	5	1 : 0.2	1 : 1	1 : confluent	200	200
72	1	5	40	5	1 : 0.2	1 : 1	1 : confluent	200 [after 24 h: + 100] [after 48 h: exchange 100]	300

For luminescent assessment of cell survival, INA-6 were harvested by removing co-culture medium, adding 150 μ L of MSC-Medium, and then stirred by strongly pipetting up and down twice while aiming the pipette tip at the upper corner, lower left and lower right of the well bottom ('Mercedes star'). Washing and stirring was repeated once before washing wells again with 150 mL MSC-Medium. Harvested INA-6 cells were strained (40 μ m), pelleted, and resuspended in 200 μ L MSC-Medium. Cells were counted using Neubauer chambers, re-distributed into 96-well plates (white, clear bottom) with 1×10^5 INA-6 cells per well, and then subjected to viability and apoptosis assays.

To minimize the loss of sensitive apoptotic cells, another approach was used to measure viability and apoptosis without harvesting INA-6 cells. hMSCs and INA-6 were seeded out individually in parallel to the co-cultures (S. Tab. 02). Prior to measuring viability and apoptosis, culture volume was adjusted to 150 μ L by removing 50 μ L or 150 μ L for the timepoints 48 h or 72 h, respectively (carefully not to stir up culture on bottom). 100 μ L of luminescent reagents were then added directly to 150 μ L of co-culture. The fold change of viability or apoptosis that is due to MSC interaction ($FC_{MSC\ interaction}$) was then calculated using the following formula, with L being the mean of four technical replicates measured in relative luminescent units per seconds [RLU/s], $L_{Co\ Culture}L_{MSC}$, $L_{INA\ 6}$ the luminescence measured in the co-culture, hMSCs alone and INA 6 alone, respectively.

$$FC_{MSC\ Interaction} = \frac{L_{Co\ Culture}}{L_{MSC} + L_{INA\ 6}}$$

Time-Lapse Characterization of INA-6 Aggregation, Detachment and Division

In order to record the aggregation and detachment of INA-6 in contact with hMSCs, hMSCs (5e3 cells/cm²) were fluorescently stained with PKH26 and plated onto 8-well μ -Slides (ibidi, Gräfelfing,

Germany). hMSCs were incubated for 24 h before being placed into an ibidi Stage Top Incubation System and were equilibrated to the incubation system for a minimum of 3 h (80% humidity and 5% CO₂). INA-6 cells (2×10^4 cells/cm²) were washed and resuspended in 33% (v/v) MSC-conditioned medium before adding them directly before acquisition start in a small volume (10 µL). Brightfield and fluorescence images of 13 mm² of co-culture were acquired every 15 minutes for 63 h. Movement speed of the motorized table was adjusted to the lowest setting that allows acquisition of the complete region within 15 minutes.

Respective events of interest were analyzed manually and categorized into defined event parameters. Events were binned across the time axis using these boundaries: [0.0, 12.85, 25.7, 38.55, 51.4, 64.25]. We collected a minimum of events per recording and analysis so that each time bin contained at least 5 values, except when analyzing detachment events, since these did not appear before 20 h of incubation for some replicates. For each recording and event parameter, the event count was normalized by dividing by the total number of events per time bin.

We determined the frequency and the cause of aggregation by looking for two interacting INA-6 cells and went backward in time to see if they were two daughter cells or if two independent INA-6 cells had collided.

We determined the frequency of aggregates with detaching cells by tracing their growth across the complete time-lapse and looking for detachment events. We picked random 100 aggregates by including aggregates from both the border and center of the well.

We characterized detachment events by noting multiple parameters manually: Time point of detachment, aggregate size (at the time of detachment), the last interaction partner, and the number of detaching INA-6 cells.

For characterizing cell division events, we recorded a new set of time-lapse videos using unstained hMSCs that were grown to confluence for 24 h (4×10^4 hMSCs/cm²) to provide for unlimited adhesion surface. We categorized daughter cells in terms of their mobility (mobility being the speed of putative movements or “rolling”). The mobility criteria were met if one INA-6 daughter cell moved farther than half a cell radius within one frame (15 min) relative to the MSC-adherent INA-6 cell which was required to stand still in-between respective frames. We measured the “rolling” duration by subtracting the time point of the last perceived movement from the time point of division. We

excluded those division events from the measurement of rolling duration, if INA-6 cells underwent apoptosis shortly after division.

Cell Cycle Synchronization at M-Phase

INA-6 cells were arrested at mitosis by double thymidine (2 mM) treatments followed by 5 h of nocodazole (500 ng/mL) incubation. In detail: 3×10^5 /mL INA-6 in 4 mL were treated with 2 mM thymidine (Sigma) for 16.5 h. Cells were released by washing them in INA-6 medium once and allowed to cycle for 9 h before treating them with 2 mM thymidine for 18 h a second time. Afterwards, cells were released and allowed to cycle for 2 h before treating them with 100 ng/ mL nocodazole (Sigma) for 5 h. Arrested INA-6 were released by washing them once and resuspending them in MSC-medium with 33% MSC-conditioned medium. Cell cycle profile was checked using image cytometry (Fig S2).

V-Well Adhesion Assay

This assay was modified from (Weetall et al., 2001). 96 v-well plates were coated with collagen I (rat tail, Corning). Collagen coating ensures that confluent hMSCs withstand centrifugation even after hMSCs in the well tip were removed. hMSCs (4×10^4 cells/cm²) were seeded out and grown to confluence for 24 h in collagen-coated v-well plates. To ensure that only INA-6 are pelleted in the v-well tip, hMSCs were removed from the well-tip by touching the well-ground with a 10 µL pipette and roughly pipetting hMSCs away.

Arrested INA-6 (1×10^4 cells/cm²) were released by washing them once in PBS and resuspending them in 33% (v/v) MSC-conditioned medium before adding them on top of confluent hMSCs. INA-6 adhered for 1, 2, 3 and 24 h before the complete co-culture was stained with 0.5 µM Calcein-AM (10 min at 37 °C).

Non-adherent INA-6 were pelleted by centrifugation using a Hettich 1460 rotor ($r = 124$ mm) at 2000 rpm (555 g) for 10 min.

The well tip was imaged by fluorescence microscopy with 5x magnification, 96 HE emission filter, autofocus configured for maximum signal intensity, 2x2 binning and 14 bit grayscale depth. Pellet brightness was analyzed in ZEN 2.6 (Zeiss) by summing up pixel brightnesses across the complete pellet image. Background brightness was acquired from a cell culture with only hMSCs. Reference brightness was acquired from a cell culture with only INA-6, defining 100% pellet brightness without

adhesion. Background intensity was subtracted before normalizing by reference. Outliers were removed from technical replicates ($n=4$) if their z-score was larger than 1.5σ technical variation.

After measuring pellet brightnesses, the cell pellet was removed by pipetting 10 μL from the well tip.

Pellets of the same technical replicates were pooled, washed in PBS, resuspended in 200 μL PBS, added to 1.8 mL ice-cold 70% ethanol and stored at -20 °C.

Remaining non-MSC-adhering INA-6 cells were removed by replacing culture medium with 100 μL of medium. MSC-adherent INA-6 were manually detached by rapid pipetting and equally pelleted, analyzed, and isolated.

Cell Cycle Profiling

INA-6 cells were fixed in 70% ice-cold ethanol, washed, resuspended in PBS, distributed in 96-well plates and stained with Hoechst-33342 (2.5 $\mu\text{g/mL}$ in PBS) for 1 h at 37 °C.

For image cytometric cell cycle profiling, plates were scanned completely using automated fluorescence microscopy with 5x magnification, 96 HE emission filter, 1x1 binning, 14 bit depth and an illumination time that fills 70% of grayscale range. The autofocus was configured to re-adjust every second tile. A pre-trained convolutional neural network (“DeepFeatures 2 reduced”, Intellesis, Zeiss) was fine-tuned to segment scans into background, single nuclei and fragmented nuclei. Nuclei were filtered to exclude fragmented nuclei and those nuclei with extreme size (within the range of 50-500 μm^2) and roundness (within the range of 0.4-1.0). Cell cycle profiles were normalized by the mode of the nucleus intensities within the G0/G1 peak. To retrieve frequencies of cells cycling in G0/G1, S, and G2 phase, the brightness distribution of all single nuclei was fitted to the sum of three Gaussian curves (“Skewed Gaussian Model” for G0G1 and G2 phase, and “Rectangle Model” for S phase) using the python package LMFIT (Newville et al., 2014) (Fig. S4). The gaussian curves were used to calculate the cell frequencies for each cell cycle phase by integration using the composite trapezoidal rule implemented by numpy.trapz (Harris et al., 2020).

For validation of image cytometry, 5 mL of INA-6 stock culture was removed and ethanol fixed as described above. Flow cytometry analyses were performed using an Attune Nxt Flow Cytometer (Thermo Fisher, USA). Data analyses were performed using FlowJo V10 software (TreeStar, USA).

Protocol: Well Plate Sandwich Centrifugation (WPSC)

96 well plates (flat bottom, clear) were coated with collagen I (rat tail, Corning). Collagen coating ensures that confluent hMSCs withstand centrifugation and repeated washing. hMSCs (2×10^4 cells/cm²) were seeded out and grown to confluence for 72 h in collagen-coated 96-well plates.

To remove aggregates from the medium and prevent clogging of magnetic columns, we strained any FCS-containing fluid with a 40 µm cell strainer.

Collect MSC-conditioned medium and add INA-6:

1. Collect hMSC-conditioned medium (CM) from the well plates and replace it with 100 µL of fresh hMSC medium. Collect CM from different donors separately
2. Strain CM (40 µm) and centrifuge it for 15 minutes at 250 g to ensure that CM does not contain hMSCs
3. Dilute CM by mixing 2 parts of CM with 1 part of MSC-medium (dilute 1.5 fold)
4. Count INA-6 cells and retrieve enough cells to fill all 96 wells with 2×10^4 INA-6/cm² (6.8×10^4 cells per well, covering ~65% of the well bottom).
5. Centrifuge INA-6 (5 min, 250 g) and resuspend them in a volume of diluted CM to reach a concentration of 6.8×10^5 INA-6/mL
6. Add 100 µL INA-6 suspension to hMSCs (end volume: 200 µL; end concentration: 33% (v/v) hMSC-conditioned medium)
7. Incubate for 24 h at 37 °C and 5% CO₂

Prepare CM-INA6 reference:

8. Add 100 µL of fresh MSC-medium into each well of an empty 96-well plate (not coated)
9. Add 100 µL of INA-6 suspension (6.8×10^5 INA-6/mL in diluted CM)
10. Incubate for 24 h at 37 °C and 5% CO₂

Collect CM-INA6 and nMA-INA6

11. Pre-warm well plate centrifuge to 37 °C
12. Prepare a counter-weight by filling 200 µL of water into all wells of an empty 96-well plate
13. Prepare well-plate sandwiches:

- a. Turn an empty 96-well plate (“catching plate”) upside down and place one on top of the co-culture-plate, the CM-IN6 reference plate, and the counter-weight so that all well openings align.
 - b. Fix well plates using tape with reusable adhesive (e.g. Leukofix)
14. Turn both plates around. Medium will spill from the co-culture plate into the catching plate
 15. Centrifuge plate for 40 seconds at 1000 rpm with the catching plate facing the ground
 16. Remove the adhesive tape and the co-culture plate.
 17. Turn the co-culture plate around and add 30 µL of washing medium (MSC-Medium 0% FCS, 3 mM EDTA) gently by touching the wall of each well and pressing the pipette slowly.
 - a. *Work quickly to ensure that co-culture does not dry. We recommend using a multipette (Eppendorf).*
 - b. *Many nMA-IN6 are removed by physical force applied by adding 30 µL of medium and not just by centrifugation. Hence, it is critical to apply the same dispensing technique across all replicates. We recommend using a multipette (Eppendorf) that can apply 30 µL with controllable pressure, since its push-button retains a long pushing path even for dispensing small volumes, unlike push-buttons from the usual 100 µL pipettes that reduce the pushing-path for smaller volumes.*
 - c. *Centrifugation minimizes technical variability by replacing one step of manual pipetting. Also, it ensures that confluent MSCs remain unharmed. Manual pipetting on the other hand would require touching the well-bottom to remove all fluids which damages the adhesive hMSC layer.*
 18. Turn the co-culture plate upside down, place it onto the catching plate and re-apply adhesive tape to fix the wellplate sandwich
 19. Repeat steps 14-18 two more times until the catching plate contains 290 µL of medium in each well
 20. Pool CM-IN6 from the catching plate that was fixed to the reference plate
 21. Pool nMA-IN6 from the catching plate that was fixed to the co-culture plate
 22. Collect remaining IN6 by adding 100 µL of PBS into each well of the catching plates, collect and pool with CM-IN6 or nMA-IN6.

23. Strain CM-INA6 and nMA-INA6 using 40 µm cell strainer
 24. Isolate MA-INA6 by continue with either accutase dissociation or rough pipetting
- Collect MA-INA6 by accutase dissociation followed by MAC sorting
25. Block 2 mL tubes with sorting buffer (PBS, 2 mM EDTA, 1% BSA) for 1 h at 4 °C
 26. Dilute accutase (Sigma A6964) (400-600 units/mL) 4-fold in cold PBS. Always keep accutase on ice, since accutase loses activity at room temperature.
 27. Add 50 µL of cold accutase (directly after the last centrifugation step) and incubate co-culture plate for 5 minutes at 37 °C.
 28. Place a co-culture plate onto a shaker and shake for 1 minute at 300 rpm.
 29. Collect cell suspension from wells and stop the reaction by adding 500 µL of FCS to pooled cell suspension.
 30. Evaluate presence of adherent INA-6 cells and the integrity of confluent hMSCs under the microscope.
 31. Repeat steps 24-27 until all INA-6 cells have dissociated or until confluent hMSCs start to tear.
 32. Strain cell suspension (30 µm). This yields MA-MSC.
 33. Pellet MA-INA6, nMA-INA6 and CM-INA6 (1200 rpm, 10 min).
 34. Resuspend MA-INA6 in 86 µL sorting buffer (PBS, 2 mM EDTA, 1% BSA)
 35. Resuspend CM-INA6 and nMA-INA6 in 300 µL cold diluted accutase and incubate for 3 min at 37 °C to ensure equal treatment for all samples.
 36. Stop accutase by adding 200 µL of FCS (100%)
 37. Pellet CM-INA6 and nMA-INA6 (1200 rpm, 10 min) and resuspend in 86 µL sorting buffer (PBS, 2 mM EDTA, 1% BSA).
 38. Transfer samples into 2 mL tubes that were blocked with sorting buffer
 39. Add 10 µL of CD45 coated magnetic beads (Miltenyi Biotec B.V. & Co. KG, Bergisch Gladbach)
 40. Place tubes into rotator and incubate for 15 minutes at 4 °C
 41. Continue with MAC sorting according to the manual. Use an MS column and wash 3 times.

42. Improve purity of eluted MA-INA6 by straining eluate (30 µm) (wash strainer using 1 mL of sorting buffer) and applying it onto an MS column a second time. Wash three times.

43. Collect 20 µL per eluate and apply it onto a 96-well plate to evaluate purity

a. Incubate plate for 24 h

b. Count the number of adherent cells (hMSCs) per INA-6 using phase contrast microscopy

c. *We reached a mean purity of 3.2×10^4 ($\pm 2.2 \times 10^4$) hMSCs per MA-INA6.*

d. *hMSC contamination did not have an impact on RNAseq, since those genes that are highly expressed in hMSCs (VCAM1, ALPL, FGF5, FGFR2), did not appear as differentially expressed in MA-INA6 (Data not shown). RNAseq detected 0.44 ± 0.16 CPM-normalized counts of VCAM1 transcripts in MA-INA6, however, it was excluded like all genes with less than 1 count in at least 2 of 5 replicates.*

44. Count cells using a Neubauer chamber

45. Pellet samples (250 g for 5 min)

46. Resuspend in respective medium or lysis buffer (e.g. RA1 for RNA extraction)

Collect MA-INA6 by rough pipetting (no MAC sorting)

47. After the last centrifugation step, add hMSC-medium to each well of the co-culture plate to reach a volume of 150 µL

a. *Since the yield of MA-INA6 was large, we dissociated MA-INA-6 cells from hMSCs by vigorous pipetting (for further samples after RNAseq, see Tab. S1). Since no enzymatic digestion is used, we reckoned that there would be no need for MAC sorting. Confluent hMSCs withstand this procedure and don't dissociate as single cells, which can be removed by straining cells (30 µm). We reached similar purities as for MAC-sorting (Data not shown).*

48. Using a multi-channel pipette (100 µL), gently raise 90 µL into the tips

49. Lean pipette tip on the upper well-border and roughly pipette up and down once

50. Repeat step 48 at the lower right and lower left well border (Total of 3 pipetting steps “Mercedes Star”)

51. Attach a catching plate onto the co-culture and centrifuge for 40 seconds at 500 rpm (28 g)

52. Repeat steps 46-50 until a sufficient amount of MA-INA6 is removed
53. Control purity of MA-INA-6 by placing out aliquot onto an empty 96-well plate.
54. Collect MA-INA6 from catching plate
55. Remove hMSCs by straining cell suspension (30 µm)
56. Count cells using a Neubauer chamber
57. Pellet MA-INA6 (250 g for 5 min)
58. Resuspend in respective medium or lysis buffer

Centrifugal force: We used a Hettich 1460 rotor ($r = 124$ mm) (Hettich GmbH & Co. KG, Tuttlingen, Germany). For calculating the centrifugal force that acts onto the co-culture within well plate sandwiches, we subtracted the height of the catching plate (14.4 mm, Greiner 96 well plate) and the depth of each well (10.9 mm). This yields a radius of 98.7mm, which translates to the following centrifugal forces: 500 rpm: 28 g; 1000 rpm: 110 g; 2000 rpm: 441 g.

Washing medium with EDTA: EDTA removes calcium from integrins which are required for adhesion. It is not strong enough to dissociate INA-6 from hMSCs, but could help with removing INA-6 from other INA-6. For generating samples for RNAseq, we added 3 mM of EDTA to washing medium. For further samples, we did not add EDTA to the washing medium, since we found that it does not increase yield for all biological replicates consistently (Data not shown). We suspect that integrin-mediated adhesion depends on hMSC donor or internal variance of INA-6. We recommend using 3 mM of EDTA, however, this requires further optimizations like including an incubation time at 37 °C after the addition of washing medium to account for biological variance. However, this could take long incubation times of up to 60 minutes (Lai et al., 2022).

Track Cell Number During WPSC

To track the cell count during WPSC, INA-6 were stained with CellTracker green and both in co-culturing- and catching plates were scanned after each centrifugation step. For each round of centrifugation, an empty catching plate was used. A pre-trained convolutional neural network (Intellesis, Zeiss) was fine-tuned to segment the scans into background, cells, and cell borders. Single cells were counted and the cumulative sum for each catching plate was calculated.

Sub-Culturing After WPSC of MSC-Interacting INA-6 Subpopulations

After CM-INA6, nMA-INA6, and MA-INA6 were isolated, they were counted with a Neubauer chamber using all nine quadrants and diluted to 10^5 cells/mL in MSC-medium (10% FCS, no IL-6 except for control). 100 μ L of cell suspension was applied to 96-well plates, incubated for 48 h at 37 °C and 5% CO₂ and then subjected to viability and apoptosis assays.

RNA Isolation

Total RNA was isolated from INA-6 cells by using the NucleoSpin RNA II Purification Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions.

RNAseq, Differential Expression and Functional Enrichment Analysis of INA-6 cells

FASTQ files were merged to the respective sample. The quality of FASTQ files was assessed with FastQC (Andrews, 2010) tool, and a joint report was created with MultiQC (Ewels et al., 2016) tool. Fastq files were aligned with STAR (Dobin et al., 2013) to the GRCh38 reference genome build (Zerbino et al., 2018). Quality and alignment statistics of final BAM files were assessed with samtools stats (Li et al., 2009), and a joint report with FastQC reports by MultiQC was generated.

Raw read counts were generated with HTSeq (Anders et al., 2015) with the union method. HTSeq runs internally in STAR. Differential gene expression analysis was done with edgeR (Robinson et al., 2010) in R 3.6.3 (R Core Team, 2018), according to the edgeR manual.

Counts were merged and genes with zero counts in all samples were removed (number of genes: 36380).

The whole count table was annotated with R Bioconductor (Gentleman, n.d.) (Gentleman et al. 2004) human annotation data package org.Hs.eg.db (Carlson, 2016).

A DGEList Element was created with the raw counts, gene information, i.e. Ensembl GenelIDs, HUGO Symbol, Genename, and ENTREZ GenelIDs and a sample grouping meta data table.

```
y <- DGEList(counts=ct2[,-1:4], group=meta.data$group, genes=ct2[,1:4])
```

Counts were filtered to keep only those genes which have at least 1 read per million in at least 2 samples (number of genes: 14136). Afterwards normalization factors were recalculated.

```
keep <- rowSums(cpm(y)>1) >=2  
y <- y[keep, , keep.lib.size = FALSE]  
y1 <- calcNormFactors(y)
```

A design matrix was created with grouping factor by treatment condition (group=F1, F2, F3, which are abbreviations for CM-INA6, nMA-INA6, MA-INA6, respectively)

```
design = model.matrix(~0+group)
```

Dispersion was estimated, the resulting coefficient of biological variation (BCV) is 0.135, i.e. BCV expression values vary up and down by 13.5% between samples.

```
y1.1 <- estimateDisp(y1, design)  
BCV <- sqrt(model.F$y1.1$common.dispersion)
```

A generalized linear (glmQLFit function) model was fitted.

```
fit <- glmQLFit(y1.1, design)
```

and pairwise comparisons were made, e.g.

```
F1vsF2 <- glmQLFTest(fit, contrast = makeContrasts(groupF1 - groupF2,  
levels = design))
```

top significant differential expressed genes were written to a table

```
DE.F1vsF2 <- topTags(F1vsF2, n=nrow(F1vsF2), p.value = 0.05)
```

Afterwards, gene list of differentially expressed genes were used for functional enrichment analysis with metascape (Zhou et al., 2019).

RT-qPCR

For cDNA synthesis 1 µg of total RNA was reverse transcribed with Oligo(dT)15 primers and Random Primers (both Promega GmbH, Mannheim, Germany) and Superscript IV reverse transcriptase (Thermo Fisher Scientific) according to the manufacturer's instructions. For quantitative PCR the cDNA was diluted 1:10 and qPCR was performed in 20 µl by using 2 µl of cDNA and 10 µl of GoTaq qPCR Master Mix (Promega GmbH) and 5 pmol of sequence-specific primers obtained from biomers.net GmbH (Ulm, Germany) or Qiagen GmbH (Hilden, Germany) (see Tab. S4 for primer sequences and PCR conditions). qPCR conditions were as follows: 95°C for 3 min; 40 cycles: 95°C for 10 s; respective annealing temperature for 10 s; 72°C for 10 s; followed by melting curve analysis for the specificity of qPCR products by using the qPCR thermal cycler Professional Thermocycler Biometra (Analytik Jena AG, Jena, Germany). Samples that showed unspecific byproducts were discarded. Ct values were measured in three technical replicates (triplicates). Non-detects were discarded. One of three technical replicates was treated as an outlier and excluded if

its z-score crossed 1.5σ technical variation. We normalized expression by the housekeeping gene 36B4. Efficiencies were determined in each reaction by linear regression of log transformed amplification curve (Ramakers et al., 2003). Differential expression was calculated based on a modified $\Delta\Delta Ct$ formula that separated exponents to apply individual efficiencies to each Ct value:

$$\text{Fold Change} = \frac{E_{tar}^{\Delta Ct_{tar}(co-treated)}}{E_{ref}^{\Delta Ct_{ref}(co-treated)}} = \frac{E_{tar,co}^{Ct_{tar,co}} : E_{tar,treated}^{Ct_{tar,treated}}}{E_{ref,co}^{Ct_{ref,co}} : E_{ref,treated}^{Ct_{ref,treated}}}$$

$E_{tar,co}$ = Efficiency of the target gene measured in the control sample

$Ct_{tar,co}$ = Ct value of the target gene measured in the control sample

tar = Target Gene; ref = Reference Gene

$treated$ = Treated sample; co = Control Sample

Fold change expression was normalized by the median of CM-INA6 (and not samplewise, as commonly used in $\Delta\Delta Ct$) since some genes were not expressed without direct MSC contact s (e.g. MMP2), and also in order to display variation of CM-INA6 next to nMA-INA6 and MA-INA6.

Tab. S4: List of primers. Some primers required a melting step to be performed before fluorescent readout to remove byproducts.

Primer	Sequence 5' - 3'	base pairs [bp]	annealing temp. [°C]
36B4_s	tgcatcagtacccatttatcat	122	60
36B4_as	aggcagatggatcagccaaga		
BCL6_s	tagagcccataaaacggtcctcat	221	55 + Melting Step at 77 °C
BCL6_as	cgc当地点attgagccgagatgtgt		
BMP4_s	tacatgcgggatcttaccg	132	58
BMP4_as	atgttcttcgtggtaaagc		
BTG2_s	gtattcttgtagggccacactaa	264	60 + Melting Step at 78 °C
BTG2_as	tcttaagggtgattcggtttggaa		
CXCL8_s	actgagagtgattgagagtggacc	251	55 + Melting Step at 77 °C
CXCL8_as	ccctacaacagacccacacaatac		
CXCL12_s	gattcttcgaaagccatgtgcc	119	56

<i>CXCL12_as</i>	caatgcacacttgtctgttgtgt		
<i>DCN_s</i>	caacaacaagcttaccagagtacct	160	57
<i>DCN_as</i>	tgaaaagactcacacccgaataaga		
<i>DKK1_s</i>	gcactgatgagtagtactgcgctag	129	56
<i>DKK1_as</i>	ttttgcagtaattccggggc		
<i>IL10RB_s</i>	gagtgagcctgtctgtgagcaa	139	55
<i>IL10RB_as</i>	cttgtaaacgcaccacagcaag		
<i>IL24_s</i>	caaacagttggacgtagaaggc	149	55
<i>IL24_as</i>	tgaaatgacacagggAACAAACCA		
<i>LOX_s</i>	ctgctcagattccccaaag	125	57
<i>LOX_as</i>	tggcatcaagcaggtaatcgat		
<i>MMP2_s</i>	ttgtatttgatggcatcgctcaga	155	56
<i>MMP2_as</i>	cgtataccgcatcaatctttccg		
<i>MMP14_s</i>	cgacaagattgatgctgctc	140	57
<i>MMP14_as</i>	tcccttcccagactttgatg		
<i>MUC1_s</i>	gcagcctctcgatataacctg	200	58
<i>MUC1_as</i>	gtaggtgggtactcgctca		
<i>NOTCH2_s</i>	gtgcttgttgaacacttgtgcc	185	55
<i>NOTCH2_as</i>	cactcgcatctgtatccaccaatg		
<i>OPG (TNFRSF11B)</i>	no sequence available (Proprietary primers from Qiagen: QT00014294 TNFRSF11B_1_SG)		60
<i>PRICKLE1_s</i>	cagaggtatatcatgaaggacggc	102	56
<i>PRICKLE1_as</i>	gtcccacaccaatatgttccccac		
<i>TGM2_s</i>	caaccttctcatcgagtagttccg	100	58
<i>TGM2_as</i>	tcatccacgactccacccag		
<i>TNFRSF1A_s</i>	ctccttcaccgcttcagaaaacc	153	55
<i>TNFRSF1A_as</i>	ttcactccaataatgccggtaactg		
<i>TRAF5_s</i>	tgcctgttagataaagaggcatca	177	56
<i>TRAF5_as</i>	aacactgcacaggttcaaataagc		

Statistics

For molecular analyses, each data point represents one biological replicate, which we define as the mean of all technical replicates of co-cultures that were seeded out from the same batch of hMSCs and/or INA-6 cells on the same day. For analyses of time-lapse recordings, each datapoint represents the normalized event count from a recording of one co-culture. We prioritized unique hMSCs for each biological replicate or recording (Tab. S1). Bars and lines represent the mean and error bars represent the standard deviation of all hMSC donors or recordings (= all biological replicates).

Metric, normal distributed, dependent data was analyzed using factorial RM-ANOVA and paired Student's t-test. Results of RM-ANOVA are reported as such: $[F(df_1, df_2) = F; p = p\text{-value}]$, with df_1 being the degrees of freedom of the observed effect, df_2 being the degrees of freedom of the error and F being the F-statistic (Vallat, 2018). If sphericity was met, p-values were not corrected with the Greenhouse-Geisser method ($p\text{-unc}$).

$$df_1 = k - 1 \quad k = \text{The number of groups (of a factor, if factorial RM-ANOVA)}$$

$$df_2 = (k - 1)(n - 1)n = \text{The number of samples in each group}$$

$$F = \frac{SS_{Effect} \div df_1}{SS_{Error} \div df_2} \quad SS = \text{Sums of squares for effect or error}$$

If datapoints within dependent sample pairs were missing, such pairs were excluded from paired t-test while other pairs of the same subject remained.

Metric non-normal distributed, independent data was analyzed using Kruskal-Wallis H-test and Mann-Whitney U tests. Results of Kruskal-Wallis H-test was reported as such: $[H(df) = H]$, with df being the degrees of freedom and H being the Kruskal-Wallis H statistic, corrected for ties (Vallat, 2018).

$$df = k - 1 \quad k = \text{The number of groups}$$

Metric bivariate non-normal distributed data was correlated using spearman's rank correlation and reported as such: $[\rho(df) = \rho, p = p\text{-value}]$, with ρ being Spearman's rank correlation coefficient. df is calculated as such:

$$df = n - 2 \quad n = \text{The number of observations}$$

These test were applied using the python (3.10) -packages pingouin (0.5.1). For three-factor RM-ANOVA we used statsmodels (0.14.0) (Seabold & Perktold, 2010; Vallat, 2018). Data was plotted

using seaborn (Waskom, 2021). Sphericity was ensured by Mauchly's test. Normality was checked with the Shapiro-Wilk test for $n > 3$.

Datapoints were log10 transformed to convert the scale from multiplicative ("foldchange") to additive, or in order to fulfill sphericity requirements.

P-values derived from patient survival data were corrected using the Benjamini-Hochberg procedure. For other post-hoc analyses, p-values were not adjusted for family-wise error rate in order to minimize type I errors. To prevent type II errors, the same conclusions were validated by different experimental setups and through varying hMSCs donors across experiments (Tab. S1).

Significant p-values from pairwise tests were annotated as stars between data groups (p-value: 0.05 > * > 0.01 > ** > 10^{-3} > *** 10^{-4} > ****). If too many significant pairs were detected, we annotated only those pairs of interest.

No power calculation was performed to determine sample size since samples were limited by availability of primary hMSC donors. Experiments were repeated until a minimum of three biological replicates were gathered.

Patient Cohort, Analysis of Survival and Expression

Patient samples ($n=873$) were collected at the UKHD and processed as described (Seckinger et al., 2017, 2018), and are available at the European Nucleotide Archive (ENA) via accession numbers PRJEB36223 and PRJEB37100. Consecutive patients with monoclonal gammopathy of unknown significance (MGUS) ($n = 62$), asymptomatic ($n = 259$), symptomatic, therapy-requiring ($n = 764$), and relapsed/refractory myeloma ($n = 90$), as well as healthy donors ($n = 19$) as comparators were included in the study approved by the ethics committee (#229/2003, #S-152/2010) after written informed consent.

Gene expression was measured by RNA sequencing as previously described (Seckinger et al., 2018). Gene expression is defined as the log2 transformed value of normalized counts + 1 (as pseudocount). Progression-free (PFS) and overall survival (OS) was analyzed for the subset of previously untreated symptomatic MM patients. For delineating "high" and "low" expression of target adhesion ($n=101$) and cell cycle ($n=173$) genes, thresholds per gene were calculated with maximally selected rank statistics by the maxstat package in R (Hothorn & Lausen, n.d.). PFS and OS were

analyzed for high vs. low expression with the Kaplan-Meier method (Kaplan & Meier, 1958). Significant differences between the curves were analyzed with log-rank tests (Harrington & Fleming, 1982). P-values were corrected for multiple testing by the Benjamini-Hochberg method. Analyses were performed with R version 3.6.3 (R Core Team, 2018).

References

- Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)*, 31(2), 166–169.
<https://doi.org/10.1093/bioinformatics/btu638>
- Andrews, S. (2010). FASTQC. A quality control tool for high throughput sequence data.
- Burger, R., Guenther, A., Bakker, F., Schmalzing, M., Bernand, S., Baum, W., Duerr, B., Hocke, G. M., Steininger, H., Gebhart, E., & Gramatzki, M. (2001). Gp130 and ras mediated signaling in human plasma cell line INA-6: A cytokine-regulated tumor model for plasmacytoma. *The Hematology Journal: The Official Journal of the European Haematology Association*, 2(1), 42–53. <https://doi.org/10.1038/sj.thj.6200075>
- Carlson, M. (2016). Org.Hs.eg.db. *Bioconductor*. <https://doi.org/10.18129/B9.bioc.Org.Hs.eg.db>
- Chatterjee, M., Hönenmann, D., Lentzsch, S., Bommert, K., Sers, C., Herrmann, P., Mathas, S., Dörken, B., & Bargou, R. C. (2002). In the presence of bone marrow stromal cells human multiple myeloma cells become independent of the IL-6/gp130/STAT3 pathway. *Blood*, 100(9), 3311–3318. <https://doi.org/10.1182/blood-2002-01-0102>
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–21. <https://doi.org/10.1093/bioinformatics/bts635>
- Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048.
<https://doi.org/10.1093/bioinformatics/btw354>
- Fernandez-Rebollo, E., Mentrup, B., Ebert, R., Franzen, J., Abagnale, G., Sieben, T., Ostrowska, A., Hoffmann, P., Roux, P.-F., Rath, B., Goodhardt, M., Lemaitre, J.-M., Bischof, O., Jakob, F., & Wagner, W. (2017). Human Platelet Lysate versus Fetal Calf Serum: These Supplements Do Not Select for Different Mesenchymal Stromal Cells. *Scientific Reports*, 7, 5132. <https://doi.org/10.1038/s41598-017-05207-1>
- Gentleman. (n.d.). *Bioconductor—BiocViews*. Retrieved June 9, 2023, from <https://bioconductor.org/packages/3.17/BiocViews.html>

- Gramatzki, M., Burger, R., Trautman, U., Marschalek, R., Lorenz, H., Hansen-Hagge, T. E., Baum, W., Bartram, C. R., Gebhart, E., & Kalden, J. R. (1994). *Two new interleukin-6 dependent plasma cell lines carrying a chromosomal abnormality involving the IL-6 gene locus.* 84 Suppl. 1, 173a–173a.
- Greenstein, S., Krett, N. L., Kurosawa, Y., Ma, C., Chauhan, D., Hidemitsu, T., Anderson, K. C., & Rosen, S. T. (2003). Characterization of the MM.1 human multiple myeloma (MM) cell lines: A model system to elucidate the characteristics, behavior, and signaling of steroid-sensitive and -resistant MM cells. *Experimental Hematology*, 31(4), 271–282.
[https://doi.org/10.1016/s0301-472x\(03\)00023-7](https://doi.org/10.1016/s0301-472x(03)00023-7)
- Harrington, D. P., & Fleming, T. R. (1982). A Class of Rank Test Procedures for Censored Survival Data. *Biometrika*, 69(3), 553–566. <https://doi.org/10.2307/2335991>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), Article 7825.
<https://doi.org/10.1038/s41586-020-2649-2>
- Hothorn, T., & Lausen, B. (n.d.). *Maximally Selected Rank Statistics in R* [Computer software].
<http://cran.r-project.org/web/packages/maxstat/index.html>.
- Kaplan, E. L., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, 53(282), 457–481.
<https://doi.org/10.1080/01621459.1958.10501452>
- Lai, T.-Y., Cao, J., Ou-Yang, P., Tsai, C.-Y., Lin, C.-W., Chen, C.-C., Tsai, M.-K., & Lee, C.-Y. (2022). Different methods of detaching adherent cells and their effects on the cell surface expression of Fas receptor and Fas ligand. *Scientific Reports*, 12(1), Article 1.
<https://doi.org/10.1038/s41598-022-09605-y>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
<https://doi.org/10.1093/bioinformatics/btp352>

- Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. (2014). *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python* [Computer software]. Zenodo.
<https://doi.org/10.5281/zenodo.11813>
- Nilsson, K., Bennich, H., Johansson, S. G., & Pontén, J. (1970). Established immunoglobulin producing myeloma (IgE) and lymphoblastoid (IgG) cell lines from an IgE myeloma patient. *Clinical and Experimental Immunology*, 7(4), 477–489.
- R Core Team. (2018). *R: A language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramakers, C., Ruijter, J. M., Deprez, R. H. L., & Moorman, A. F. M. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters*, 339(1), 62–66. [https://doi.org/10.1016/S0304-3940\(02\)01423-4](https://doi.org/10.1016/S0304-3940(02)01423-4)
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, 26(1), 139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- Seckinger, A., Delgado, J. A., Moser, S., Moreno, L., Neuber, B., Grab, A., Lipp, S., Merino, J., Prosper, F., Emde, M., Delon, C., Latzko, M., Gianotti, R., Lüoend, R., Murr, R., Hosse, R., J., Harnisch, L. J., Bacac, M., Fauti, T., ... Vu, M. D. (2017). Target Expression, Generation, Preclinical Activity, and Pharmacokinetics of the BCMA-T Cell Bispecific Antibody EM801 for Multiple Myeloma Treatment. *Cancer Cell*, 31(3), 396–410.
<https://doi.org/10.1016/j.ccr.2017.02.002>
- Seckinger, A., Hillengass, J., Emde, M., Beck, S., Kimmich, C., Dittrich, T., Hundemer, M., Jauch, A., Hegenbart, U., Raab, M.-S., Ho, A. D., Schönland, S., & Hose, D. (2018). CD38 as Immunotherapeutic Target in Light Chain Amyloidosis and Multiple Myeloma-Association With Molecular Entities, Risk, Survival, and Mechanisms of Upfront Resistance. *Frontiers in Immunology*, 9, 1676. <https://doi.org/10.3389/fimmu.2018.01676>
- Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, 3(31), 1026.
<https://doi.org/10.21105/joss.01026>
- Weetall, M., Hugo, R., Maida, S., West, S., Wattanasin, S., Bouhel, R., Weitz-Schmidt, G., Lake, P., & Friedman, C. (2001). A Homogeneous Fluorometric Assay for Measuring Cell

- Adhesion to Immobilized Ligand Using V-Well Microtiter Plates. *Analytical Biochemistry*, 293(2), 277–287. <https://doi.org/10.1006/abio.2001.5140>
- Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C. G., Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O. G., Janacek, S. H., Juettemann, T., To, J. K., ... Flórek, P. (2018). Ensembl 2018. *Nucleic Acids Research*, 46(D1), D754–D761. <https://doi.org/10.1093/nar/gkx1098>
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., Benner, C., & Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*, 10(1), Article 1. <https://doi.org/10.1038/s41467-019-09234-6>



Statement of individual author contributions and of legal second publication rights to manuscripts included in the dissertation

Manuscript 1: Research Article (submitted, under revision)

Martin Kuric (MK), Susanne Beck, Doris Schneider, Wyonna Rindt, Marietheres Evers, Jutta Meißner-Weigl, Sabine Zeck, Melanie Krug, Marietta Herrmann, Tanja Nicole Hartmann, Ellen Leich, Maximilian Rudert, Denitsa Docheva, Anja Seckinger, Dirk Hose, Franziska Jundt, Regina Ebert (RE) (2024): Keep it Together: Describing Myeloma Dissemination *in vitro* with hMSC-Interacting Subpopulations and their Aggregation/Detachment Dynamics, **Cancer Research Communications**

Participated in	Author Initials , Responsibility decreasing from left to right				
Study Design	<u>MK</u>	Regina Ebert	Wyonna Rindt		
Methods Development	<u>MK</u>	Doris Schneider			
Data Collection	<u>MK</u>	Doris Schneider			
Data Analysis and Interpretation	<u>MK</u>	Susanne Beck	Regina Ebert		
Manuscript Writing Writing of Introduction Writing of Materials & Methods Writing of Discussion Writing of First Draft	<u>MK</u>	Regina Ebert			

Explanations: The content of this publication exceeds the usual scope (~29 pages Supplemental). It includes not only research findings but also survival data and protocols of new, established methods and their validations. The contribution of Martin Kuric was pivotal and predominant in all aspects of this work. Doris Schneider assisted in the experimental procedures. Susanne Beck analyzed the raw data from RNAseq and survival data, which were interpreted, depicted, and summarized by Martin Kuric.

Manuscript 2: Data Analysis Software (submitted, passed peer-review, under revision)

Martin Kuric (MK), Regina Ebert (2024): plotastic: Bridging Plotting and Statistics in Python, **Journal of Open Source Software**

Participated in	Author Initials , Responsibility decreasing from left to right				
Idea, Architectural Design	<u>MK</u>				
Software Development Feature Implementation Testing	<u>MK</u>				
Distribution of Software Documentation Version Control (GitHub) Deployment (PyPi)	<u>MK</u>				
Manuscript Writing Writing of Statement of Need Writing of Example Writing of Overview	<u>MK</u>	Regina Ebert			

Explanations: The software was entirely created by Martin Kuric, comprising more than 8000 total lines (including ~2000 testable lines) and is comparable in size to a typical web application. The release of this software involved version control using GitHub, packaging and deployment on PyPi. Regina Ebert gave feedback on submitted manuscript.

Manuscript 3: Research Letter (published)

Daniela Simone Maichl, Julius Arthur Kirner, Susanne Beck, Wen-Hui Cheng, Melanie Krug, Martin Kuric (MK), Carsten Patrick Ade, Thorsten Bischler, Franz Jakob, Dirk Hose, Anja Seckinger, Regina Ebert & Franziska Jundt (2023): Identification of NOTCH-driven matrisome-associated genes as prognostic indicators of multiple myeloma patient survival, **Blood Cancer Journal 13:134**

Participated in	Author Initials, Responsibility decreasing from left to right				
Study Design Methods Development	Daniela Simone		Franziska Jundt		
Data Collection	Daniela Simone		Franziska Jundt		
Data Analysis and Interpretation	Daniela Simone	Susanne Beck	Franziska Jundt	<u>MK</u>	
Manuscript Writing Writing of Introduction Writing of Materials & Methods Writing of Discussion Writing of First Draft	Daniela Simone		Franziska Jundt	<u>MK</u>	

Explanations: This co-authorship is not a chapter in this dissertation. Martin Kuric produced figures of processed but complex-to-visualize data and gave feedback on submitted manuscript.

Manuscript 4: Research Paper (under peer-review)

Wyonna Rindt, Melanie Krug, Shuntaro Yamada, Franziska Sennefelder, Louisa Belz, Wen-Hui Cheng, Azeem Muhammad, Martin Kuric (MK), Marietheres Evers, Ellen Leich, Tanja Nicole Hartmann, Ana Rita Pereira, Marietta Herrmann, Jan Hansmann, Mohammed Ahmed Yassin, Kamal Mustafa, Regina Ebert, and Franziska Jundt (2024): A 3D bioreactor model to study osteocyte differentiation and mechanobiology under perfusion and compressive mechanical loading, **Acta Biomaterialia**

Participated in	Author Initials, Responsibility decreasing from left to right				
Study Design Methods Development	Wyonna Rindt	Franziska Jundt		<u>MK</u>	
Data Collection	Wyonna Rindt	Franziska Jundt		<u>MK</u>	
Data Analysis and Interpretation	Wyonna Rindt	Franziska Jundt		<u>MK</u>	
Manuscript Writing Writing of Introduction Writing of Materials & Methods Writing of Discussion Writing of First Draft	Wyonna Rindt	Franziska Jundt		<u>MK</u>	

Explanations: This co-authorship is not a chapter in this dissertation. Martin Kuric contributed by counseling during weekly meetings in tight collaboration with Franziska Jundt's group, assisting Wyonna Rindt during laboratory experiments, image analysis and giving feedback on submitted manuscript.

Manuscript 5: Research Paper (under revision)

Marietta Herrmann, Jutta Schneidereit, Susanne Wiesner, Martin Kuric (MK), Maximilian Rudert, Martin Lüdemann, Mugdha Srivastava, Norbert Schütze, Regina Ebert, Denitsa Docheva, Franz Jakob (2024): Peripheral blood cells enriched by adhesion to CYR61 are heterogenous myeloid modulators of tissue regeneration with early endothelial progenitor characteristics, **European Cells and Materials**

Participated in	Author Initials, Responsibility decreasing from left to right				
Study Design Methods Development	Marietta Herrmann				
Data Collection	Marietta Herrmann			<u>MK</u>	

Data Analysis and Interpretation	Marietta Herrmann				
Manuscript Writing Writing of Introduction Writing of Materials & Methods Writing of Discussion Writing of First Draft	Marietta Herrmann			<u>MK</u>	

Explanations: This co-authorship is not a chapter in this dissertation. Martin Kuric contributed by establishing and measuring large automated microscopy scans of stained cells for quantifying osteogenic differentiation and giving feedback on submitted manuscript.

Manuscript 6: Research Letter (published)					
Participated in	Author Initials, Responsibility decreasing from left to right				
Study Design Methods Development	Marietheres Evers				
Data Collection	Marietheres Evers				
Data Analysis and Interpretation	Marietheres Evers			<u>MK</u>	
Manuscript Writing Writing of Introduction Writing of Materials & Methods Writing of Discussion Writing of First Draft	Marietheres Evers			<u>MK</u>	

Explanations: This co-authorship is not a chapter in this dissertation. Martin Kuric contributed by counseling during regular meetings with Ellen Leich's group and giving feedback on submitted manuscript.

If applicable, the doctoral researcher confirms that she/he has obtained permission from both the publishers (copyright) and the co-authors for legal second publication.

The doctoral researcher and the primary supervisor confirm the correctness of the above mentioned assessment.

Würzburg

Doctoral Researcher's Name	Date	Place	Signature
----------------------------	------	-------	-----------

Würzburg

Primary Supervisor's Name	Date	Place	Signature
---------------------------	------	-------	-----------



Statement of individual author contributions to figures/tables of manuscripts included in the dissertation

Manuscript 1: Research Article (submitted, under revision)

Martin Kuric (MK), Susanne Beck, Doris Schneider, Wyonna Rindt, Marietheres Evers, Jutta Meißner-Weigl, Sabine Zeck, Melanie Krug, Marietta Herrmann, Tanja Nicole Hartmann, Ellen Leich, Maximilian Rudert, Denitsa Docheva, Anja Seckinger, Dirk Hose, Franziska Jundt, Regina Ebert1 (2024): Keep it Together: Describing Myeloma Dissemination *in vitro* with hMSC-Interacting Subpopulations and their Aggregation/Detachment Dynamics, **Cancer Research Communications**

Figure	Author Initials , Responsibility decreasing from left to right				
1	<u>MK</u>	Doris Schneider			
2	<u>MK</u>	Doris Schneider			
3	<u>MK</u>	Doris Schneider	Sabine Zeck	Wyonna Rindt	Melanie Krug
4	<u>MK</u>	Doris Schneider	Susanne Beck		
5	<u>MK</u>	Susanne Beck			
6	<u>MK</u>	Susanne Beck			
7	<u>MK</u>				
S1	<u>MK</u>	Doris Schneider	Sabine Zeck	Wyonna Rindt	Melanie Krug
S2	<u>MK</u>	Doris Schneider	Marietta Herrmann		
S3	<u>MK</u>	Doris Schneider	Sabine Zeck		
S4	<u>MK</u>				
S5	<u>MK</u>				
S6	<u>MK</u>	Susanne Beck			
Table	Author Initials , Responsibility decreasing from left to right				
1	<u>MK</u>	Susanne Beck			
2	<u>MK</u>	Susanne Beck			
S1	<u>MK</u>	Doris Schneider			
S2	<u>MK</u>	Susanne Beck			
S3	<u>MK</u>				
S4	<u>MK</u>	Doris Schneider			

Manuscript 2: Data Analysis Software (submitted, passed peer-review, under revision)

Martin Kuric, Regina Ebert (2024): plotastic: Bridging Plotting and Statistics in Python, **Journal of Open Source Software**

Figure	Author Initials , Responsibility decreasing from left to right				
1	<u>MK</u>				
Table	Author Initials , Responsibility decreasing from left to right				
1	<u>MK</u>				
2	<u>MK</u>				

Documentation	Author Initials , Responsibility decreasing from left to right				
README	<u>MK</u>				
Example Gallery	<u>MK</u>				
Features	<u>MK</u>				
Testing	Author Initials , Responsibility decreasing from left to right				
Test-Code (Pytest)	<u>MK</u>				
Continuous Integration	<u>MK</u>				

Explanations: All files are available on GitHub (<https://github.com/markur4/plotastic>) and installable via pypi.com. Documentations are found in the Readme, including example gallery and feature explanation. Software tests was written using pytest. Coverage of code by tests is reviewable with codecov (<https://app.codecov.io/gh/markur4/plotastic>). Continuous Integration is implemented using GitHub actions.

Manuscript 3: Research Letter (published)

Daniela Simone Maichl, Julius Arthur Kirner, Susanne Beck, Wen-Hui Cheng, Melanie Krug, Martin Kuric, Carsten Patrick Ade, Thorsten Bischler, Franz Jakob, Dirk Hose, Anja Seckinger, Regina Ebert & Franziska Jundt (2023): Identification of NOTCH-driven matrisome-associated genes as prognostic indicators of multiple myeloma patient survival, **Blood Cancer Journal** **13:134**

Figure	Author Initials , Responsibility decreasing from left to right				
1 a	Daniela Simone			Susanne Beck	
1 b	Daniela Simone			Susanne Beck	
1 c	Daniela Simone			Susanne Beck	<u>MK</u>
1 d	Daniela Simone			Susanne Beck	<u>MK</u>
Table	Author Initials , Responsibility decreasing from left to right				
1	Daniela Simone				

Explanations: Martin Kuric plotted multidimensional diagrams using python and fine-adjusted them using professional design software (Affinity Publisher, Serif Ltd).

Manuscript 4: Research Paper (under peer-review)

Wyonna Rindt, Melanie Krug, Shuntaro Yamada, Franziska Sennefelder, Louisa Belz, Wen-Hui Cheng, Azeem Muhammad, Martin Kuric (MK), Marietheres Evers, Ellen Leich, Tanja Nicole Hartmann, Ana Rita Pereira, Marietta Hermann, Jan Hansmann, Mohammed Ahmed Yassin, Kamal Mustafa, Regina Ebert, and Franziska Jundt (2024): A 3D bioreactor model to study osteocyte differentiation and mechanobiology under perfusion and compressive mechanical loading, **Acta Biomaterialia**

Figure	Author Initials , Responsibility decreasing from left to right				
1	Wyonna Rindt				<u>MK</u>
2	Wyonna Rindt				
3	Wyonna Rindt				
4	Wyonna Rindt				
5	Wyonna Rindt				<u>MK</u>
6	Wyonna Rindt				<u>MK</u>
7	Wyonna Rindt				<u>MK</u>

Explanations: Martin Kuric contributed by counseling on experimental procedures and data analysis, such as quantifying normalized fluorescence intensity of immunohistochemistry and qPCR.

Manuscript 5: Research Paper (under revision)

Marietta Herrmann, Jutta Schneidereit, Susanne Wiesner, Martin Kuric (MK), Maximilian Rudert, Martin Lüdemann, Mugdha Srivastava, Norbert Schütze, Regina Ebert, Denitsa Docheva, Franz Jakob (2024): Peripheral blood cells enriched by adhesion to CYR61 are heterogenous myeloid modulators of tissue regeneration with early endothelial progenitor characteristics, **European Cells and Materials**

Figure	Author Initials , Responsibility decreasing from left to right				
1	Marietta Herrmann				
2	Marietta Herrmann				
3	Marietta Herrmann				
4	Marietta Herrmann				
5	Marietta Herrmann				
6	Marietta Herrmann				
7	Marietta Herrmann				<u>MK</u>

Explanations: Martin Kuric scanned osteogenically differentiated MSCs in Fig. 7 for quantification of alizarin red staining.

Manuscript 6: Research Letter (published)

Marietheres Evers, Martin Schreder, Thorsten Stühmer, Franziska Jundt, Regina Ebert, Tanja Nicole Hartmann, Michael Altenbuchinger, Martina Rudelius, Martin Kuric (MK), Wyonna Darleen Rindt, Torsten Steinbrunn, Christian Langer, Sofia Catalina Heredia-Guerrero, Hermann Einsele, Ralf Christian Bargou, Andreas Rosenwald, Ellen Leich (2023): Prognostic value of extracellular matrix gene mutations and expression in multiple myeloma, **Blood Cancer J.** 13(1):43

Figure	Author Initials , Responsibility decreasing from left to right				
1	Marietheres Evers				
2	Marietheres Evers				

Explanations: Martin Kuric contributed indirectly through counseling and feedback on submitted manuscript.

I also confirm my primary supervisor's acceptance.

Doctoral Researcher's Name

Date

Place

Signature

Affidavit

I hereby confirm that my thesis entitled "Development and Semi-Automated Analysis of an in vitro Model for Myeloma Cells Interacting with Mesenchymal Stromal Cells" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg
Place, Date

Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation "Entwicklung und semi-automatisierte Analyse eines in vitro-Modells für Myelomzellen in Interaktion mit mesenchymalen Stromazellen" eigenständig, d.h. insbesondere selbstständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg
Ort, Datum

Unterschrift

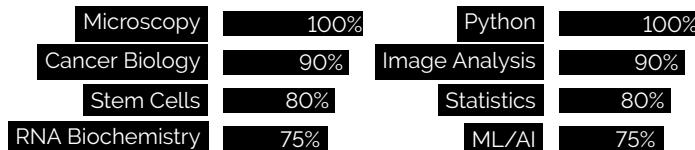
MARTIN KURIC

Cell Biologist | Data Scientist



WHO AM I?

As a cancer cell biologist with a strong passion for data analysis and machine learning, I am seeking a position where I can utilize my creativity to automate tasks, solve complex problems or handle big data.



SELECTED PROJECTS

2024	Python Software “ImageP” Accelerates batch processing of images of different sizes and types by >100%. <code>numpy</code> / <code>skimage</code> / <code>scipy</code>	GitHub Repository ↗
2020-2024	Python Software “plotastic” Published a statistical library that self-configures based on intuitive plotting parameters. <code>pandas</code> / <code>matplotlib</code> / <code>pingouin</code> / <code>seaborn</code>	Journal of Open Source Software GitHub Repository ↗
2018-2024	Cancer Research Project Worked in a team with up to three technical assistants and published a list of genes with relevance for survival of myeloma patients (under peer-review). <code>Time-Lapse Microscopy</code> / <code>RNAseq</code> / <code>Analysis of Patient Survival</code>	Journal: Cancer Research Communications
26.05.2022	Deep-Learning Assisted Image Cytometry Measurement of per-cell parameters from large automated microscopy scans. <code>Convolutional Neural Networks</code> / <code>Image Segmentation</code>	Poster at "Achilles Conference"

EDUCATION

28.01.2019 – 2024	Dr. rer. nat. in Biomedicine Research focus: Dissemination of multiple myeloma & mesenchymal stromal cell interactions	Prof. Dr. Regina Ebert University of Würzburg
01.04.2017 – 2024 parallel to M.Sc. & PhD	Elite Biological Physics Interdisciplinary & international study program for exceptional students of physics or biology.	University of Bayreuth
01.10.15 – 15.08.18	M.Sc. in Biochemistry & Molecular Biology Research focus: RNA biochemistry, small RNAseq, stem cells & piRNAs in <i>S. mediterranea</i>	Prof. Dr. Claus-D. Kuhn University of Bayreuth
01.10.12 – 14.12.15	B.Sc. in Biochemistry Research focus: Cell biology, mitochondrial inheritance in <i>S. cerevisiae</i>	Prof. Dr. Benedikt Westermann University of Bayreuth

LANGUAGES

German, English - C2
Slovakian - passive
French, Spanish - A2

SOFT SKILLS

Quality Management
Project Management
Violent Free Communication

HOBBIES

Coding - Python
Music - Piano & Guitar
Gym - Lift. Grow. Repeat

Würzburg

05.03.2024

Location

Date

Martin Kuric

Signature