Exploring the regulation of centriole size by CP110/Cep97 and their interaction network



Laura E. Hankins

Merton College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Hilary Term 2021

Acknowledgements

I would like to start by thanking my supervisor, Jordan Raff. I couldn't have asked for a better mentor. His door was always open and, over the course of my DPhil, he consistently supported both my science and my personal development. This included not only nurturing my confidence as a researcher, but also encouraging me to apply for a lot of opportunities that I am grateful for taking.

I am also grateful to all members of the Raff Lab, past and present, for their help and guidance over the last few years. I would particularly like to thank everyone who collaborated on some of the work presented in this thesis. The nature of the lab encourages cooperation, so it is no surprise that my training has been enriched by this. It has been a privilege to work in such a warm and friendly environment with a really dedicated group of people, and most importantly it's been a lot of fun! Special thanks go to my collaborators: Mustafa Aydogan, Alan Wainman, Saroj Saurya, Thomas Steinacker, Zsofi Novak, Mo Mofatteh and Ines Alvarez Rodrigo. I also thank Felix Zhou from the Lu Lab for his work on the GFP-Cep97 distancing algorithm. Finally, I would like to thank Jonathan Bohlen, Anna Caballe, Violet Feng, Joey Fisher, Lisa Gartenmann, Yanis Guimard, Seb Haddock, Andi Haensele, Nic Liew, Iona Manley, Nada Mohamad, João Monteiro, Ed Rea, Helio Roque, Judy Sayers, Lewis Timms, Eline van Houtum, Zach Wilmott, and Isaac Wong. You all helped make my time in the lab brilliant.

The Sir William Dunn School of Pathology has been an excellent department to complete my DPhil in. Thank you to everyone who has made it such a nice place to work. I am grateful to my Thesis Committee, Clive Wilson and Pedro Carvalho, for their feedback and advice, and to Lucinda Risius and Louise Samson for all their help. I am indebted to the Wellcome Trust for their generous funding of my DPhil. In particular, I would like to thank them for their response to the pandemic that struck during the final year of my studies; the funding extension they provided made it possible to complete this thesis. I also thank the British Society for Cell Biology for their support, and in particular for funding my attendance at the 2019 ASCB/EMBO meeting. I was privileged to take part in the Physiology Course at Woods Hole during my third year, and the generosity of Merton College helped to fund my place on the programme. I met some amazing scientists from across the world who also happen to be a great bunch of people; it was certainly an adventure, and I hope we'll remain in touch.

I would also like to thank my friends who have supported me throughout my studies. To Doğukan, Ines, Jack, Menelaos, Miles and Nehir for all the fun adventures. To Ed and Ines, guardians of the void. To Tanya Wilson, whose many talents include offering moral support whilst punting. To Branoc Richards – keep those wind turbines spinning, and don't forget you'll need a tray. To Aravinth Kulanthaivelu, who has kept the cynical flame burning bright since we first met at the start of our DPhil journeys. To Jack Ashton, for our mutual love of '70s clothes, 70s music, and 90s alternative comedy'. To Gregory Thompson, college husband and expert in dry wit. May my short stature provide you with amusement for years to come.

Michael, this thesis would be a far lesser work without your unwavering love and support. Thank you for making me smile, even through the tough times, and for always being there to listen, even when I was busy constructing mountains from molehills.

The greatest thanks are reserved for my family, who have supplied me with constant encouragement and love. To my brothers, Matthew and Joshua, my harshest critics and strongest allies, who always know exactly how to make me laugh. To my grandparents, who I hope would be proud (Papa, I'm sorry this is a thesis and not a novel, although it's almost as long as one!). And, most of all, to my parents, Mark and Jane. Thank you for everything. You've been my greatest advocates since day one and I am so lucky to be your daughter. It is no exaggeration to say that this thesis wouldn't exist without your support, so it seems fitting that I'm dedicating it to you.

COVID-19 Impact Statement

During the final year of my DPhil, the coronavirus pandemic resulted in our laboratory being shut down. Later, when we were able to return, reduced occupancy rules and time constraints meant that certain experiments were prioritised over others in order to complete this Thesis. Experiments that I was unable to finish included validating the expression level of Klp10A in *klp10A* mutant heterozygotes and in Ubq-Klp10A embryos (Chapter 5) and imaging eCep97-GFP, eCP110-GFP, Ubq-Cep97-GFP and Ubq-CP110-GFP in their respective mutant backgrounds (Chapter 3). I investigated the recruitment of CP110-PBM-GFP to centrioles (Chapter 6) but was unable to fully extend this to the Cep97-PBM-GFP fly line that I had generated. Finally, as discussed in Chapter 4, I would have preferred to generate a larger dataset of cycle 12 Ubq-Cep104-GFP embryos in order to further characterise the mode of Cep104 centriolar recruitment (Figure 4.11). Despite these challenges I was able to complete my DPhil, and I am grateful to the Wellcome Trust, the Sir William Dunn School of Pathology, and to my supervisor and colleagues for their support during this unprecedented time.

Abstract

Exploring the regulation of centriole size by CP110/Cep97 and their interaction network

Laura E. Hankins, Merton College
A thesis submitted for the degree of Doctor of Philosophy, Hilary Term 2021

Centrioles are cylindrical organelles that serve as templates for the assembly of centrosomes and cilia. They exhibit 9-fold symmetry and comprise a central cartwheel structure surrounded by nine microtubule (MT) blades. Within the same cell type, centriole length is consistent, suggesting that the growth of these organelles is regulated. While recent advances have been made in understanding how the growth of the cartwheel is controlled, less is known about the mechanisms mediating centriole MT length. The proteins CP110 and Cep97 localise at the distal-ends of centrioles, and are thought to 'cap' centriole MTs to inhibit their over-elongation. However, it is not clear whether this regulation takes place during the growth of the centriole, or afterwards to maintain proper centriole size.

In this thesis, I aimed to explore the regulation of centriole size by CP110/Cep97 and their interaction network. Development of an assay to track the dynamic growth of centriole MTs in early *Drosophila* embryos revealed that CP110/Cep97 overexpression slowed centriole MT growth. However, this may not be due to the direct effect of these proteins on MTs; rather, CP110/Cep97 play an unexpected role in affecting the growth rate and period of the underlying cartwheel structure, which is likely separate from their role in capping centriole MTs. Cep104 and Klp10A also localise to the distal-ends of centrioles and have been shown to interact with CP110/Cep97, yet I found that they do not affect daughter centriole growth dynamics. Indeed, Cep104 appears to act downstream of CP110/Cep97. These findings lead me to propose a model whereby centriole growth dynamics in *Drosophila* are primarily driven by the assembly of the cartwheel structure, with centriole MTs tracking the cartwheel's growth. Once centrioles reach their final size, proper centriole length must be maintained, and this likely involves CP110/Cep97 (in their capping role) alongside Klp10A.

Publications resulting from this thesis

Hankins, L. E.*, Aydogan, M. G*., Steinacker, T. L., Mofatteh, M., Saurya, S., Wainman, A., Wong, S. S., Lu, X., Zhou, F., Raff, J. W. (manuscript in preparation). Centriole distal-end proteins CP110 and Cep97 influence cartwheel growth at the proximal-end of centrioles.

*These authors contributed equally.

List of Figures

Figure 1.1 – Centriole structure in *Drosophila*. 4

Figure 1.2 – Centriole structure in vertebrates....7

Figure 1.3 – Centrosome structure in *Drosophila* embryos....9

Figure 1.4 – Cilia structure. 12

Figure 1.5 – The centrosome cycle. 16

Figure 1.6 – The centriole duplication pathway is conserved across model organisms. 19

Figure 1.7 – Plk4 oscillates at the centriole. 22

Figure 1.8 – Schematic summary of the Cep104 interaction network. 35

Figure 1.9 – The centrosome cycle in early *Drosophila* embryos. 40

Figure 1.10 – An assay to monitor the cartwheel growth dynamics of daughter centrioles. 42

Figure 2.1 – Crossing scheme for transient gRNA expression strategy. 52

Figure 2.2 – Crossing scheme for transgenic gRNA integration strategy. 53

Figure 3.1 – CP110/Cep97 localisation is consistent with them binding to the centriole MTs. 91

Figure 3.2 – CP110/Cep97 are required for each other's recruitment to centrioles. 93

Figure 3.3 – Development of an assay to track the dynamic growth of centriole MTs. 96

Figure 3.4 – Summary of the number of embryos in each dataset that were fitted with model *(1)* or model *(2)*. 99

Figure 3.5 – Centriole MTs grow slower and possibly for a longer period when CP110/Cep97 are overexpressed. 100

Figure 3.6 – Further examples of MT distancing analysis. 102

Figure 3.7 – Retraining the GFP-Cep97 distancing algorithm on CP110-GFP data. 105

Figure 3.8 – CP110/Cep97 levels oscillate at the centriole. 107

Figure 3.9 – The peak of the CP110/Cep97 oscillation correlates with S-phase length. 110

Figure 3.10 – The CP110/Cep97 oscillation occurs predominantly on the daughter centriole. 112

Figure 3.11 – Schematic summarising the predicted effect of CP110/Cep97 overexpression on the GFP-Cep97 oscillation, if the oscillation influences growth. 114

Figure 3.12 – Western blots to assess the expression of eCep97-GFP and eCP110-GFP. 115

Figure 3.13 – Comparing the recruitment dynamics of eCep97-GFP with Ubq-GFP-Cep97, and eCP110-GFP with Ubq-CP110-GFP. 119

Figure 3.14 – Overexpression of both CP110 and Cep97 leads to faster, and higher levels of, centriole recruitment. 121

Figure 3.15 – The growth rate of the MTs does not correlate with the amplitude of the GFP-Cep97 signal. 123

Figure 3.16 – Dynamics of cartwheel growth in embryos expressing varying doses of CP110/Cep97. 125

Figure 3.17– Dynamics of Plk4 centriolar recruitment in embryos expressing varying doses of CP110/Cep97.....127

Figure 3.18 – Western blots comparing CP110/Cep97 expression in embryos under different Plk4 conditions. 129

Figure 3.19 – Immunoprecipitation assay using embryo extract to test for an interaction between Plk4 and CP110/Cep97.....131

Figure 3.20 – Centriolar recruitment of Plk4, CP110 and Cep97 in nuclear cycle 14. 136

Figure 4.1 – Cep104 is recruited to both mother and daughter centrioles in *Drosophila*.....146

Figure 4.2 – Generation of a *cep104* knockout fly line using CRISPR/Cas9.....148

Figure 4.3 – Crossing scheme for transient gRNA expression strategy. 150

Figure 4.4 – Crossing scheme for transgenic gRNA integration strategy. 151

Figure 4.5 – Western blots confirm the absence of full-length Cep104 in the deletion stock. 154

Figure 4.6 – Expression of a shortened version of Cep104 was not detectable via Western blotting. 155

Figure 4.7 – Negative geotaxis assays to assess the coordination of *cep104* mutant flies. ...157

Figure 4.8 – Micrographs demonstrating that CP110/Cep97 localise to centrioles in the absence of full-length Cep104. 160

Figure 4.9 – The parameters of the GFP-Cep97 oscillation in varying doses of Cep104. 163

Figure 4.10 – Graphs showing the recruitment dynamics of Cep104-GFP over cycles 11-13. 165

Figure 4.11 – Comparisons of the Cep104 and Cep97 peaks in cycle 12. 169

Figure 4.12 – Micrographs demonstrate that Cep104 recruitment is dramatically reduced in the absence of either CP110 or Cep97. 171

Figure 4.13 – Cartwheel growth dynamics in embryos heterozygous for the *cep104* deletion. 173

Figure 4.14 – Comparisons of Sas-6-GFP and Sas-6-mNG incorporation dynamics in cycle 12.....177

Figure 4.15 – Cartwheel growth dynamics in varying doses of Cep104. 179

Figure 4.16 – Alignment of *D. melanogaster* and *H. sapiens* Cep104 protein sequences. ...186

Figure 4.17 – Predicted tertiary structure of *D. melanogaster* Cep104. 187

Figure 5.1 – The parameters of the GFP-Cep97 oscillation are not significantly perturbed in embryos heterozygous for the *klp10A* mutation. 195

Figure 5.2 – Attempts to visualise Klp10A-GFP at wing disc centrioles. 197

Figure 5.3 – Cartwheel growth dynamics under various expression levels of Klp10A. 199

Figure 6.1 – Schematic summarising the CP110/Cep97 regulatory network. 209

Figure 6.2 – Localisation of CP110 and Cep97 mutants that cannot bind Polo. 212

Figure 6.3 – Western blot to compare expression levels of CP110-WT-GFP and CP110-PBM-GFP. 214

Figure 6.4 – The CP110-PBM-GFP oscillation in WT embryos. 216

Figure 6.5 – The CP110-PBM-GFP oscillation in embryos lacking endogenous CP110. 218

Figure 6.6 – Western blot to examine CP110 expression in embryos expressing CP110-WT-GFP or CP110-PBM-GFP in the full CP110 mutant background. 219

Figure 6.7 – Schematic showing predicted CDK phosphorylation sites in CP110/Cep97. ...221

Figure 6.8 – Design and injection of the CP110-del-GFP construct into *Drosophila* embryos. 223

Figure 6.9 – Centriole length in *cp110* null wing discs overexpressing CP110-WT-GFP or CP110-PBM-GFP. 226

Figure 6.10 – The Cep97 oscillation continues under S-phase arrest. 228

Figure 6.11 – The CP110-GFP signal is lost from centrioles under mitotic arrest. 230

Figure 7.1 – Schematic illustrating important considerations for the '***cartwheel-driven***' model of daughter centriole biogenesis. 238

Figure 7.2 – Schematic illustrating the concept of the '***mechanical model***' of daughter centriole biogenesis. 239

List of Tables

Table 1.1 – Key proteins required for centriole duplication in flies, listed alongside their functional homologues in other species……...19

Table 1.2 – Further proteins implicated in centriole biology that are referenced in this thesis……28

Table 2.1 – *Drosophila* lines used in this thesis……...46

Table 2.2 – gRNAs used in this thesis……50

Table 2.3 – Primers used in this thesis……55

Table 2.4 – The microlysis PCR programme……...57

Table 2.5 – Plasmids used in this thesis……...60

Table 2.6 – Buffers used for immunoprecipitation……..67

Table 2.7 – Primary antibodies used in this thesis……...69

Table 2.8 – Secondary antibodies used in this thesis……...70

Table 2.9 – Primers used for amplification of cyclin genes……72

Abbreviations

3D-SIM	3D structured illumination microscopy
A.U.	Arbitrary units
Ala	Alanine
Ana	Anastral spindle
ANOVA	Analysis of variance
AP	Average precision
APC	Anaphase-promoting complex
Asl	Asterless
Asp	Aspartate
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
β-TrCP	Beta-transducin repeats-containing protein
Cas	CRISPR-associated protein
CC	Coiled-coil
CCO	Cdk/cyclin oscillator
CDK	Cyclin-dependent kinase
CDK5RAP	Cdk5 regulatory subunit-associated protein
Cep	Centrosomal protein
CFF	Cambridge Fly Facility
CLASP	CLIP-associating protein
CLEM	Correlative light and electron microscopy
CLIP	Cytoplasmic linker protein
CLR	Cilium-like region
Cnn	Centrosomin
CNN	Convolutional neural network
CP110	Centrosomal protein of 110 kDa
CPAP	Centrosomal P4.1-associated protein
CRISPR	Clustered regularly interspaced short palindromic repeats
CS	Centrosome separation
CV	Coefficient of variation
del	Deletion
DEPC	Diethylpyrocarbonate
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
ds	Double-stranded
DTT	Dithiothreitol
e	Endogenous promoter
EB	End-binding protein
EDTA	Ethylenediamine-tetraacetic acid
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
EM	Electron microscopy
FAP	Flagellar-associated protein

FCS	Fluorescence correlation spectroscopy
FLP	Flippase
FRAP	Fluorescence recovery after Photobleaching
GFP	Green fluorescent protein
Glu	Glutamate
GMM	Gaussian mixture model
gRNA	Guide RNA
GSC	Germline stem cell
γ-TuRC	γ-tubulin ring complex
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
HF	High fidelity
HRP	Horse radish peroxidase
IF	Immunofluorescence
IFT	Intraflagellar transport complex
IP	Immunoprecipitation
IQR	Interquartile range
Kif	Kinesin superfamily protein
Klp	Kinesin-like protein
LSM	Laser scanning microscope
MBD	Microtubule-binding domain
MBP	Maltose binding protein
MCAK	Mitotic centromere-associated kinesin
MPP	M-phase phosphoprotein
mRNA	Messenger ribonucleic acid
miRNA	Micro ribonucleic acid
MSE	Mean-squared error
MTs	Microtubules
Mut	Mutant
Na-acetate	Sodium acetate
NEB	Nuclear envelope breakdown
NG	NeonGreen
O/E	Overexpression
PAGE	Polyacrylamide gel electrophoresis
PB	Polo-box
PBM	Polo-binding mutant
PBS	Phosphate-buffered saline
PBST	1x PBS + 0.1% Triton
PCM	Pericentriolar material
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PIPES	Piperazine-N, N′-bis
Plk/PLK	Polo-like kinase
POC	Proteome of centriole
PPP1R35	Protein phosphatase 1 regulatory subunit 35
Pro	Proline
PSIPRED	PSI-blast based secondary structure PREDiction
qRT-PCR	Quantitative reverse-transcription polymerase chain reaction
Rcd	Reduction in Cnn dots
RELU	Rectified linear units

RFP	Red fluorescent protein
RKA	Reduced kinase activity
RNAi	RNA interference
ROI	Region of interest
RPE	Retinal pigment epithelium
RT	Room temperature
RTTN	Rotatin
S2	Schneider 2
Sas/SAS	Spindle assembly abnormal
SCF	Skp, Cullin, F-box
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SELU	Scaled exponential linear units
SEM	Standard error of the mean
Ser	Serine
SMLM	Single-molecule localisation microscopy
SNP	Single nucleotide polymorphism
SNR	Signal:noise ratio
Spd	Spindle-defective
SPICE	Spindle and centriole-associated protein
STAN	STIL/Ana2
STIL	SCL/TAL1 interrupting locus
TAE	Tris-acetate-EDTA
TEM	Transmission electron microscopy
Thr	Threonine
TOG	Tumour overexpressed gene
TBS	Tris-buffered saline
TTBK	Tau tubulin kinase
TZ	Transition zone
Ubq	Ubiquitin promoter
USP	Ubiquitin specific peptidase
WB	Western blotting
WT	Wild type
XMAP	*Xenopus* microtubule-associated protein
Zeo	Zeocin
ZNF	Zinc finger
ZYG	Zygotic defective

Table of Contents

Acknowledgements ..... **i**

COVID-19 Impact Statement ..... **iii**

Abstract ..... **v**

Publications resulting from this thesis ..... **vi**

List of Figures ..... **vii**

List of Tables ..... **xiii**

Abbreviations ..... **xv**

Chapter 1: Introduction **1**

1.1. Centrioles, centrosomes and cilia **1**

1.1.1. Centriole structure 1

1.1.1.1. The central cartwheel 2

1.1.1.2. The surrounding centriole microtubules 5

1.1.2. Centrosomes 7

1.1.3. Cilia 10

1.2. The centrosome cycle **12**

1.3. The centriole duplication pathway **17**

1.3.1. Cartwheel assembly 20

1.3.2. Centriole MT assembly 23

1.3.3. Factors involved in centriole size control 25

1.4. CP110/Cep97 ..... **29**

1.4.1. Regulation of CP110/Cep97 levels 31

1.4.2. CP110/Cep97 interaction partners at the distal-end 32

1.4.2.1. Cep104 33
1.4.2.2. Klp10A 36
1.4.2.3. Kif24 38
1.5. Tracking centriole protein dynamics in syncytial *Drosophila* embryos 39
1.5.1. The centrosome cycle in *Drosophila* blastoderms 39
1.5.2. An assay for tracking centriole cartwheel growth 41
1.6. Thesis outline 43
Chapter 2: Materials and Methods 45
2.1. Fly husbandry 45
2.1.1. Fly handling 45
2.1.2. Embryo handling for live imaging 45
2.1.3. *Drosophila melanogaster* stocks 46
2.1.4. Transgenic line generation 49
2.1.5. Recombinant line generation 50
2.1.6. CRISPR knockout line generation 50
2.1.6.1. Strategy 1: transient expression of gRNAs in nos-Cas9 flies 51
2.1.6.2. Strategy 2: transgenic expression of gRNAs 52
2.1.7. Climbing assays 54
2.2. Molecular Biology 54
2.2.1. PCR reactions 54
2.2.1.1. DNA amplification via PCR 54
2.2.1.2. Diagnostic PCRs for *cep104* deletion events 57
2.2.1.3. Diagnostic PCRs to sequence the *cep104* deletion 57
2.2.2. Construction of plasmids 58
2.2.2.1. Construction of plasmids for generation of transgenic *D. melanogaster* lines 58
2.2.2.2. Construction of plasmids for CRISPR gRNA expression 59

2.2.2.3. Construction of plasmids for *in vitro* mRNA generation 59

2.2.2.3.1. Generation of mRNA for injection into *D. melanogaster* embryos..... 60

2.3. Biochemistry and Cell Biology 62

2.3.1. Fixation of *D. melanogaster* embryos for western blotting 62

2.3.2. Preparation of *D. melanogaster* brains for western blotting.....62

2.3.3. SDS-PAGE and western blotting 62

2.3.3.1. Quantification from western blot to assess global CP110/Cep97 levels under different Plk4 conditions..... 64

2.3.4. Fixation of *D. melanogaster* embryos for immunoprecipitation 64

2.3.5. Immunoprecipitation 65

2.3.5.1. Coupling the α-GFP antibody to Dynabeads 65

2.3.5.2. Preparing the embryo extract.....65

2.3.5.3. Starting the IP 66

2.3.5.4. Western blot 67

2.3.6. Fixation and staining of *D. melanogaster* wing discs for imaging..... 68

2.3.7. Antibody purification..... 68

2.3.8. Primary antibodies used in this study 69

2.3.9. Secondary antibodies used in this study 70

2.3.10. Injection experiments in living embryos..... 71

2.3.10.1. Injection of mRNA into living embryos..... 71

2.3.10.2. Drug injection into living embryos 71

2.3.10.3. RNA interference in living embryos 71

2.4. Microscopy 72

2.4.1. Live imaging of *D. melanogaster* embryos.....72

2.4.1.1. Spinning disc confocal microscopy.....73

2.4.1.2. Zeiss 880 LSM with Airyscan detector 74

2.4.1.3. OMX V3 Blaze with 3D-structured illumination (3D-SIM) 74
2.4.2. Imaging of fixed and stained *D. melanogaster* wing discs 75
2.4.2.1. OMX V2 with 3D-SIM/single-molecule localisation microscopy (SMLM) 75
2.4.3. Transmission electron microscopy (TEM) 75
2.5. Image processing and analysis 76
2.5.1. Processing spinning disc images 76
2.5.1.1. Centriole tracking with TrackMate to generate fluorescence intensity curves 76
2.5.1.2. Non-linear regression models 78
2.5.1.3. Correlation analysis 80
2.5.1.4. Correcting the CP110/Cep97 oscillation centre for S-phase length 80
2.5.1.5. Quantification of Ubq-CP110-GFP/Ubq-GFP-Cep97 in *cep97*/*cp110* null embryos 80
2.5.2. Processing Zeiss 880 LSM Airyscan images 81
2.5.2.1. Airyscan processing 81
2.5.2.2. Quantification of Ubq-CP110-GFP/Ubq-GFP-Cep97 on mother/daughter centrioles 81
2.5.2.3. Analysis of centriole MT growth dynamics via machine learning 81
2.5.3. Processing OMX V2 and OMX V3 Blaze images 82
2.5.3.1. Reconstruction and alignment 82
2.5.3.2. Measurements of CP110/Cep97 ring radius 82
2.5.4. Image presentation for figures 83
2.6. Statistics 83
2.7. Bioinformatics 84
2.7.1. Identification of putative phosphorylation sites 84
2.7.2. Identification of cyclin-dependent kinase (CDK) sites 84
2.7.3. Structural predictions 85

Chapter 3: The role of CP110 and Cep97 in centriole MT length control 87

3.1. Introduction..... 87

3.1.1. Centriole length is regulated 87

3.1.2. CP110/Cep97 are implicated in centriole MT length control.....87

3.1.3. Key findings in this chapter..... 89

3.2. Results 90

3.2.1. CP110/Cep97 are required for each other's recruitment to *Drosophila* centrioles 90

3.2.2. Centriole MTs grow linearly during S-phase and stop growing before the onset of mitosis..... 94

3.2.2.1. Live imaging of centriole MT growth in *Drosophila* early embryos 94

3.2.2.2. Perturbing the global level of CP110/Cep97 affects the growth rate and period of centriole MTs 97

3.2.2.3. Attempts to use CP110-GFP as a marker to assay MT growth..... 103

3.2.3. CP110/Cep97 oscillate at the centriole 106

3.2.3.1. CP110/Cep97 levels oscillate at the centriole, peaking in S-phase..... 106

3.2.3.2. The peak of the CP110/Cep97 oscillation correlates with S-phase length..... 108

3.2.3.3. CP110/Cep97 predominantly oscillate on the growing daughter centriole..... 110

3.2.4. The properties of the CP110/Cep97 oscillation do not appear to regulate centriole MT growth..... 113

3.2.4.1. Generation of a fly line expressing CP110-GFP at endogenous levels 114

3.2.4.2. Higher levels of CP110/Cep97 expression lead to faster, and higher levels of, centriole recruitment 116

3.2.5. CP110/Cep97 play a surprising role in cartwheel growth regulation 123

3.2.6. CP110/Cep97 may interact with Plk4 at the centriole 126

3.3. Discussion..... 132

Chapter 4: Exploring the CP110/Cep97 Interaction Network – Cep104 139

4.1. Introduction 139

4.1.1. The mechanism by which CP110/Cep97 are recruited to centrioles is unknown 139

4.1.2. CP110/Cep97 are implicated in MT length control yet have no known MT-binding domains 140

4.1.3. CP110/Cep97 interact with other distal-end-binding proteins, including Cep104 ... 140

4.1.4. Key findings in this chapter 142

4.2. Results 143

4.2.1. Cep104 localises to *Drosophila* centrioles 143

4.2.2. CP110/Cep97 are still recruited to centrioles in the absence of Cep104 146

4.2.2.1. Generation of a *cep104* CRISPR knockout *Drosophila* line 146

4.2.2.1.1. Strategy 1: transient expression of gRNAs in nos-Cas9 flies 149

4.2.2.1.2. Strategy 2: transgenic expression of gRNAs 150

4.2.2.2. *cep104* knockout flies are viable, fertile and coordinated 156

4.2.2.3. GFP-Cep97 and CP110-GFP are still recruited to centrioles in the absence of Cep104 158

4.2.2.4. The parameters of the GFP-Cep97 oscillation are not perturbed by the absence of full-length Cep104 160

4.2.3. Like CP110/Cep97, Cep104 is recruited to centrioles in an oscillatory manner 163

4.2.3.1. Levels of Cep104-GFP oscillate at the centriole over the course of the cell cycle, peaking in S-phase 163

4.2.3.2. Cep104-GFP fluorescence peaks later in S-phase than RFP-Cep97 166

4.2.3.3. The peak of the Cep104 oscillation correlates with S-phase length 169

4.2.4. Cep104 recruitment to centrioles is dramatically reduced in the absence of CP110/Cep97 170

4.2.5. Cep104 levels have little effect on cartwheel growth during centriole biogenesis ... 172

4.2.5.1. Centriole cartwheel growth is not perturbed by in *cep104* mutant heterozygotes 172

4.2.5.2. Cartwheel growth rate is faster in Sas-6-mNeonGreen cartwheels than Sas-6-GFP....173

4.2.5.3. Using Sas-6-mNeonGreen to assay cartwheel growth in different doses of Cep104....177

4.2.5.4. Preliminary experiments were carried out into the effect of Cep104 dose on centriole MT distancing 180

4.3. Discussion 181

Chapter 5: Exploring the CP110/Cep97 Interaction Network – Klp10A....189

5.1. Introduction 189

5.1.1. CP110/Cep97 interact with the MT depolymerase Klp10A.... 189

5.1.2. Klp10A is implicated in centriole length control 190

5.1.3. Key findings in this chapter.... 191

5.2. Results 193

5.2.1. The Cep97 oscillation is not perturbed in embryos heterozygous for a *klp10A* mutation 193

5.2.2. Altering the dose of Klp10A does not affect centriole cartwheel growth.... 197

5.2.3. Preliminary experiments were carried out on the effect of Klp10A dose on centriole MT distancing 200

5.3. Discussion....201

Chapter 6: CP110/Cep97 and the Cell Cycle 205

6.1. Introduction 205

6.1.1. CP110/Cep97 levels and recruitment dynamics vary across different cell cycle stages 205

6.1.2. CP110/Cep97 have been identified as putative PLK1 interactors 207
6.1.3. Key findings in this chapter 210
6.2. Results 211
6.2.1. Polo binding is not important for the recruitment of CP110/Cep97 211
6.2.1.1. Generation of fly lines expressing CP110/Cep97 mutants that cannot bind Polo 211
6.2.1.2. CP110-GFP mutants that cannot bind Polo are still recruited to the centriole, where they exhibit normal dynamics 214
6.2.1.3. Design of CP110/Cep97 with mutations to putative CDK phosphorylation sites 220
6.2.1.4. CP110 recruitment to the centriole is not perturbed by deletion of regions that contain a high density of putative phosphorylation sites 221
6.2.2. Polo binding is not important for CP110's MT capping function in wing discs 224
6.2.3. The Cep97 oscillation persists under S-phase arrest 227
6.2.4. CP110 is lost from centrioles under mitotic arrest 229
6.3. Discussion 231
Chapter 7: General Discussion 235
7.1. Summary of Key Findings 235
7.2. Proposed model for daughter centriole growth 236
7.3. Centriole size maintenance 241
7.4. Future directions 243
References 247
Appendix 267
A1. Automated measurements of GFP-Cep97 distancing 267
A1.1. Centriole detection 268

A1.2. CNN to automatically predict centriole pairs 274

A1.3. Pairing of individual mother and daughter centrioles and measuring the distance between the centres of the pairs 278

A1.4. Validating the GFP-Cep97 distancing output 283

A1.5. Fitting the resulting distancing data 283

A1.6. Extracting the GFP-Cep97 oscillation data from images used for distancing measurements 286

Supplemental References 289

Chapter 1: Introduction

1.1. Centrioles, centrosomes and cilia

Centrioles are cylindrical organelles which exhibit 9-fold radial symmetry. Their structure is highly conserved, and centrioles are thought to have been present in the last common ancestor of eukaryotes, though they have subsequently been lost from some lineages (Carvalho-Santos *et al.*, 2011, Nabais *et al.*, 2020). Centrioles duplicate once per cell cycle, during interphase, when the older ('mother') centriole grows a new ('daughter') centriole (details of the centrosome cycle are reviewed in Nigg and Holland, 2018). The daughter grows to the same size as the mother, and the mechanisms regulating this size control are an area of active research (Aydogan *et al.*, 2018, Breslow and Holland, 2019, Sharma *et al.*, 2021). Centriole size control is particularly important given that centrioles organise centrosomes and cilia, two key structures in the cell, and defects in these have been implicated in diseases such as ciliopathies and cancer (reviewed in Godinho and Pellman, 2014, Nigg and Holland, 2018, Nigg and Raff, 2009, Reiter and Leroux, 2017, Valente *et al.*, 2014). Below, I introduce our current knowledge on centriole size control, before summarising the new findings detailed in this thesis.

1.1.1. Centriole structure

Centrioles are composed of a central structure called the 'cartwheel' and the centriole microtubules that surround it. The assembly of these structures involves a network of conserved proteins.

1.1.1.1. The central cartwheel

The central cartwheel (shown in green, **Figure 1.1 A**) is thought to establish the 9-fold symmetry of the centriole's overall structure and is largely composed of a protein called Spindle ASsembly abnormal protein 6 (Sas-6), which forms homodimers via its coiled-coil domain (**Figure 1.1 D**; Cottee *et al.*, 2015, Kitagawa *et al.*, 2011, Nakazawa *et al.*, 2007, van Breugel *et al.*, 2011). Nine of these homodimers then assemble into a ring through interactions between Sas-6's N-terminal head domains; these head domains form the central 'hub', while coiled-coil domains point outwards to form the cartwheel's 'spokes' (**Figure 1.1 C-E**; Cottee *et al.*, 2015, Kitagawa *et al.*, 2011, van Breugel *et al.*, 2011, van Breugel *et al.*, 2014). Sas-6 is incorporated at the proximal end of the daughter centriole (Aydogan *et al.*, 2018), and the stacking of these Sas-6 rings results in the elongation of the cartwheel structure *in vivo* (Guichard *et al.*, 2012, Guichard *et al.*, 2013).

Sas-6 interacts with another protein, Anastral spindle 2 (Ana2; the human homologue is known as STIL – see **Section 1.3.1** for more details on how these proteins interact during cartwheel assembly) which oligomerises and may act to reinforce the 9-fold symmetric structure of the Sas-6 ring during cartwheel formation (Cottee *et al.*, 2015, David *et al.*, 2016, Slevin *et al.*, 2014, Stevens *et al.*, 2010b). Sas-6 rings have been shown to stack in the cartwheel as pairs (**Figure 1.1 E**; Guichard *et al.*, 2012, Guichard *et al.*, 2013), and this interaction with Ana2/STIL can help explain how this is mediated (Cottee *et al.*, 2015). Cep135 also seems important for the stacking process (Guichard *et al.*, 2017). In *Caenorhabditis elegans,* the Ana2/STIL functional homologue SAS-5 also oligomerises and is thought to aid SAS-6 assembly (Qiao *et al.*, 2012, Rogala *et al.*, 2015, Shimanovskaya *et al.*, 2013). A key difference is that SAS-6 in *C. elegans* forms a spiral

structure known as the ‘central tube’, although this still exhibits 9-fold symmetry (Hilbert *et al.*, 2013).

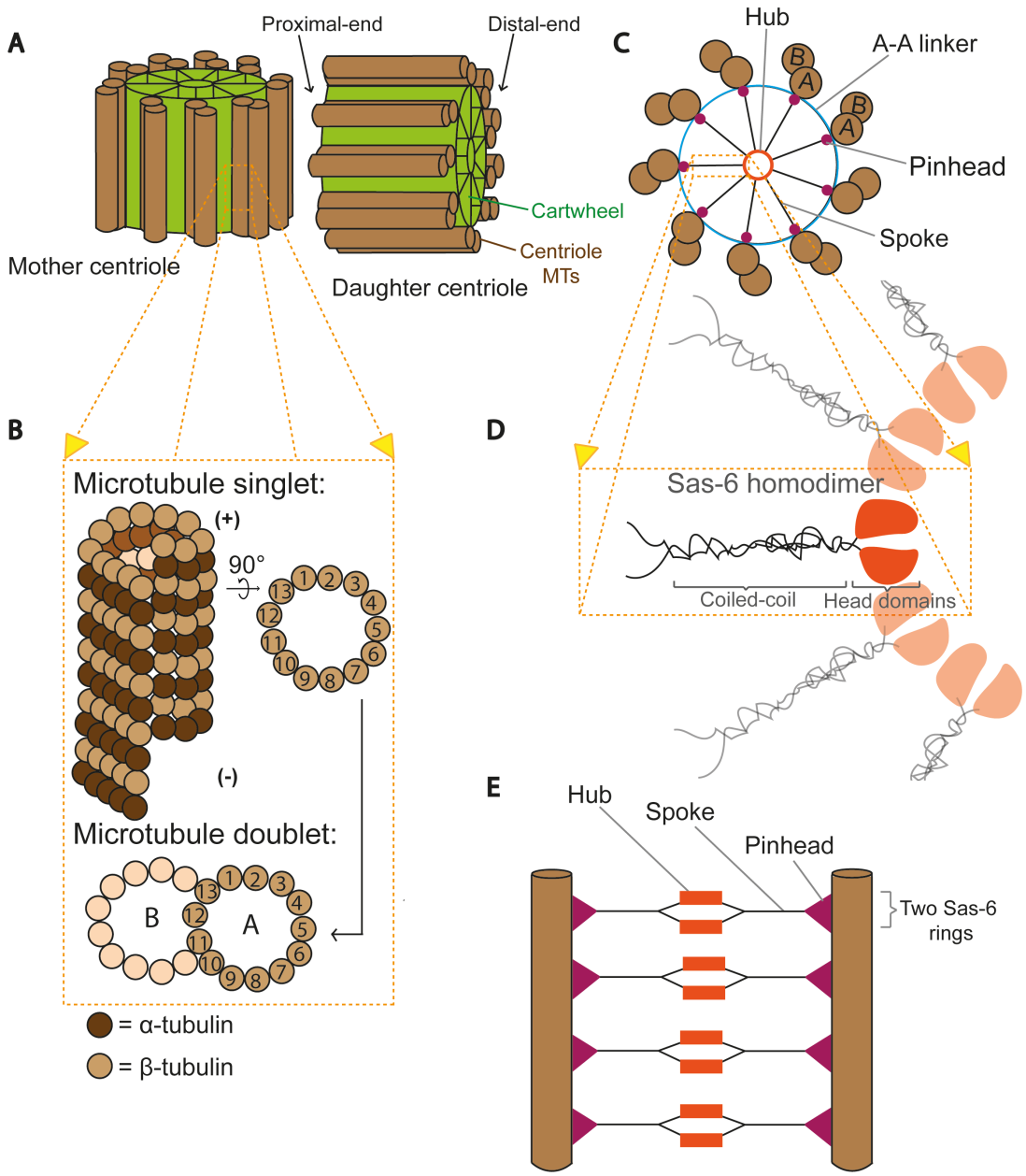


Figure 1.1 – Centriole structure in *Drosophila*.

(A) Schematic showing a typical pair of *Drosophila* centrioles, comprising a central cartwheel (green) and surrounding centriole doublet MTs (brown). The older centriole ('mother') nucleates a new centriole ('daughter') during S-phase. The proximal- and distal-ends of the daughter are indicated with arrows. **(B)** Schematic showing the typical structure of a MT singlet comprising 13 protofilaments of α- and β-tubulin (*above*) and a MT doublet (*below*). **(C)** Schematic showing an end-on view of a typical *Drosophila* centriole. The A- and B-tubules (brown) are indicated and the linkers (blue) between the A-tubules are shown. The cartwheel's spokes (black), central hub (orange) and the pinhead that links the cartwheel to the MTs (purple) are also indicated. **(D)** Schematic showing that the central cartwheel is composed of nine Sas-6 homodimers. Their N-terminal head domains form the central hub (orange) and their coiled-coil domains form the spokes (black). The head domains interact with other Sas-6 homodimers (pale orange/grey) to facilitate the formation of a ring of nine dimers. **(E)** Schematic showing a longitudinal cross-section of a cartwheel to illustrate that the rings of Sas-6 homodimers stack in pairs; colour scheme is same as in (C).

1.1.1.2. The surrounding centriole microtubules

Microtubules (MTs) are well-known to be hollow cylindrical structures, typically composed of 13 protofilaments that assemble via lateral interactions (reviewed in van der Vaart *et al.*, 2009). These protofilaments are composed of dimers of α- and β-tubulin, and have an intrinsic polarity due to the orientation of these dimers, with β-tubulin at the plus-end and α- tubulin at the minus-end (**Figure 1.1 B**). MTs exhibit dynamic instability, with periods of growth and shrinkage (Kirschner and Mitchison, 1986, Mitchison and Kirschner, 1984). In the cytoplasm, they are perhaps best known for their role in forming the mitotic spindle, which facilitates chromosome segregation during mitosis, and their dynamic instability is important for this function (reviewed in Inoue and Salmon, 1995).

MTs (shown in beige, **Figure 1.1 A**) also form part of the structure of the centriole, where they assemble around the central cartwheel. Centriole MTs are very stable compared with the dynamic MTs of the cytoplasm; this stability is partly due to tubulin posttranslational modifications, including glutamylation (Bobinnec *et al.*, 1998, Janke and Magiera, 2020, Kochanski and Borisy, 1990). In fly cells, centriole MTs are formed as doublets, comprising a complete tubule of 13 protofilaments (the A-tubule, closest to the cartwheel) and an incomplete tubule (the B-tubule) (**Figure 1.1 B, C**; Callaini *et al.*, 1997, Gottardo *et al.*, 2015, Greenan *et al.*, 2018, Winey and O'Toole, 2014). These doublets are sometimes termed ‘MT blades’ and there are linkers between the A-tubules (**Figure 1.1 C**; Greenan *et al.*, 2018). In mammalian centrioles (**Figure 1.2**) and the centrioles in fly testes, the MT blades are present as triplets (Gottardo *et al.*, 2015, Winey and O'Toole, 2014).

The nine sets of MT blades are linked to the cartwheel via the pinhead structures at the ends of the cartwheel spokes (**Figure 1.1 C, E**; Guichard *et al.*, 2013). It is thought that this attachment is mediated by Cep135, a protein that localises to the pinhead region (Hiraki *et al.*, 2007, Lin *et al.*, 2013a). In *Drosophila,* loss of Cep135 increases the diameter of the ring formed by the MT blades as well as leading to destabilised centrioles (Roque *et al.*, 2012). Depletion of Cep135 in the unicellular eukaryote *Tetrahymena* also destabilises already formed centrioles (Bayless *et al.*, 2012).

It seems likely that centriole MTs grow from the distal-end of the daughter centriole, since proteins thought to inhibit inappropriate centriole MT extension are located here, and their removal from this site can promote centriole MT elongation or ciliogenesis (see **Section 1.4** for further discussion and references). These proteins are called CP110 and Cep97, and they form the main subject of this thesis. Furthermore, an electron microscopy (EM) study on human lymphoblast centrioles postulated that the A-tubule could be nucleated by the γ-tublin ring complex (γ-TuRC) to ensure that it only grows in one direction (Guichard *et al.*, 2010). However, little is known about the specific mechanisms that regulate centriole MT growth.

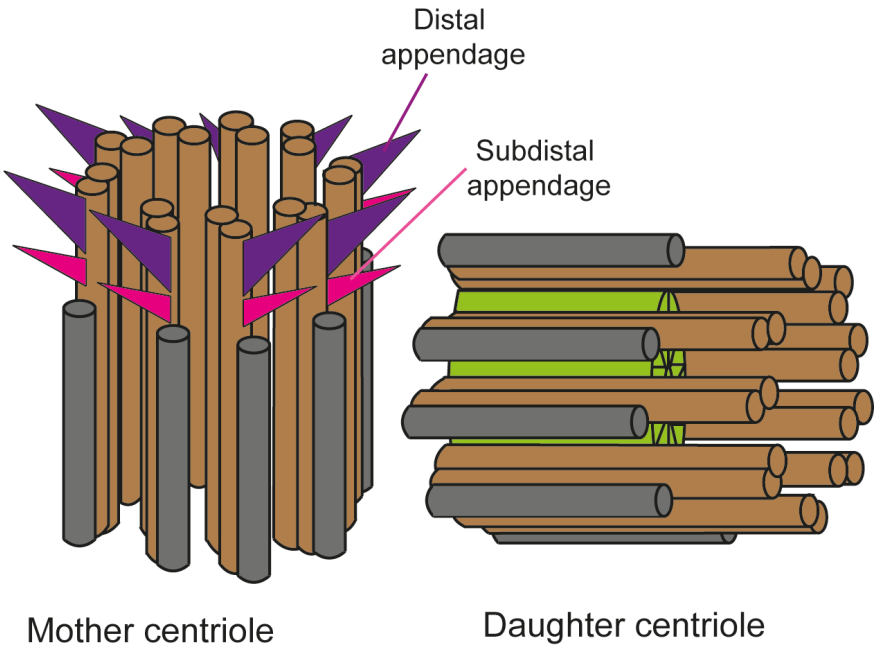


Figure 1.2 – Centriole structure in vertebrates.

Schematic showing a typical pair of vertebrate centrioles. Note that, unlike *Drosophila* centrioles, the mother centriole lacks a cartwheel, and the daughter's cartwheel (green) is only present at its proximal-end, where the surrounding MTs form triplets (C-tubule shown in grey). The mother centriole also has distal and subdistal appendages.

1.1.2. Centrosomes

Centrosome formation begins at the onset of mitosis, when the proteins surrounding the centriole pair (termed the pericentriolar material, or PCM) dramatically expand in order to nucleate the MTs of the mitotic spindle (**Figure 1.3**; reviewed in Woodruff *et al.*, 2014). Spindle-defective protein 2 (Spd-2) and Centrosomin (Cnn), two proteins that are important for this process in flies, are initially recruited to centrioles with the help of the centriolar protein Asterless (Asl) and expand outwards to form a scaffold (Conduit *et al.*, 2010, Conduit *et al.*, 2014a, Conduit *et al.*, 2014b, Dix and Raff, 2007, Dobbelaere *et al.*,

2008, Feng *et al.*, 2017, Goshima *et al.*, 2007). A recent structural study from our laboratory has shown that this scaffold is partly formed via defined molecular interactions; Cnn dimers are phosphorylated by Polo kinase, causing them to undergo a conformational change that allows them to interact and assemble into a scaffold via conserved domains (Feng *et al.*, 2017). Polo is a mitotic kinase, first discovered in *Drosophila*, that associates with centrosomes throughout the cell cycle in syncytial fly embryos (Llamazares *et al.*, 1991, Moutinho-Santos *et al.*, 1999). Polo also plays other important roles in the centrosome cycle, discussed in **Section 1.2**.

Polo has long been known to be important for mitotic centrosome formation (Conduit *et al.*, 2014a, Dobbelaere *et al.*, 2008, Feng *et al.*, 2017, Goshima *et al.*, 2007, Lane and Nigg, 1996), and its recruitment to centrosomes by Spd-2 was recently shown to be required to drive PCM expansion (Alvarez-Rodrigo *et al.*, 2019). A model for PCM expansion in *Drosophila* embryos proposed by Alvarez Rodrigo *et al.* 2019 is illustrated in **Figure 1.3**.

While centrosomes in flies appear to assemble the Cnn scaffold via specific molecular interactions (Feng *et al.*, 2017), there is evidence that the PCM in *C. elegans* may form via liquid-liquid phase separation; *in vitro* studies of SPD-5 (the functional equivalent of Cnn in worms) have shown that this protein can self-assemble into droplets that are capable of recruiting other PCM proteins, concentrating tubulin and, ultimately, assembling MT asters (Woodruff *et al.*, 2017, Woodruff *et al.*, 2015). However, it has been noted that there are key differences between the composition of these droplets and that of centrosomes (reviewed in Raff, 2019).

Maintaining proper centriole size appears to be important for centrosome function, since changes to a centriole's size may be reflected in the amount of PCM it recruits (Kirkham *et al.*, 2003, Marteil *et al.*, 2018). This affects centrosome size, which can in turn impact the size of the mitotic spindle (Greenan *et al.*, 2010, Kirkham *et al.*, 2003, Marteil *et al.*, 2018). This idea is further explored in **Section 1.3.3**.

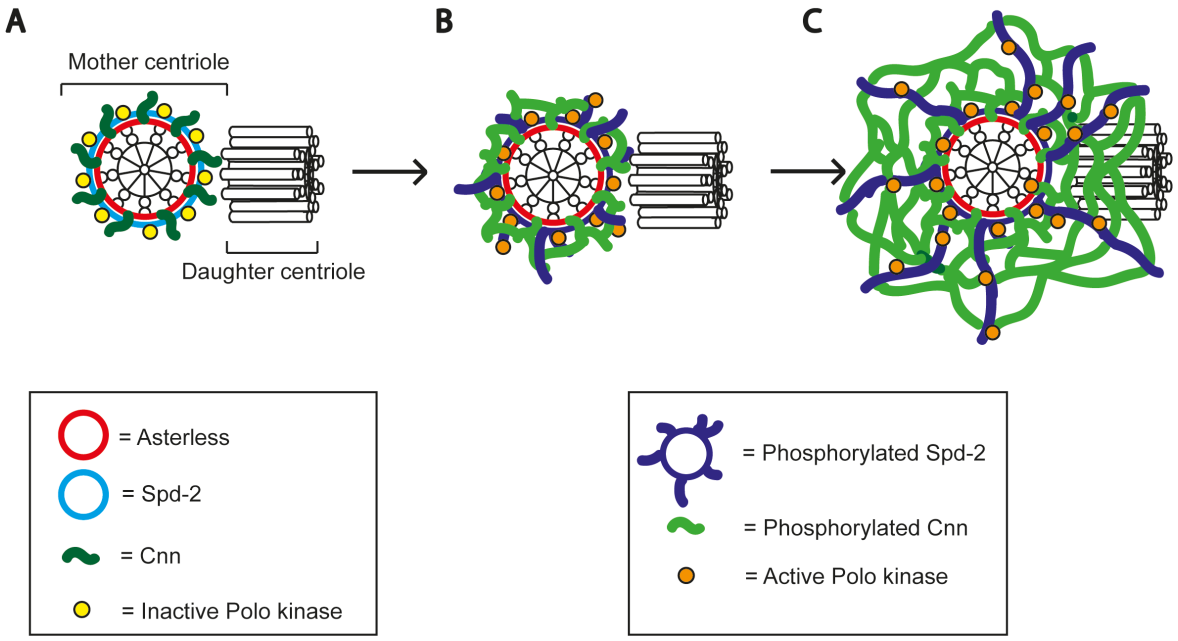


Figure 1.3 – Centrosome structure in *Drosophila* embryos.

Schematic (adapted from Alvarez Rodrigo *et al.* 2019) illustrating a recently published model of how centrosome structure changes during the cell cycle in *Drosophila* embryos. **(A)** During S-phase, the mother centriole can be distinguished from the daughter by the presence of Asl, a protein that is recruited as centrioles mature (Novak *et al.*, 2014). The mother has also recruited a small amount of Spd-2 and Cnn, as well as Polo kinase (Fu and Glover, 2012). **(B)** At the onset of mitosis, Polo becomes active and phosphorylates Spd-2 and Cnn, which initiates the process of scaffold assembly (Alvarez-Rodrigo *et al.*, 2019, Conduit *et al.*, 2014a, Feng *et al.*, 2017). This phosphorylation allows Spd-2 to spread outwards and recruit more Cnn and Spd-2 (Alvarez-Rodrigo *et al.*, 2019, Conduit *et al.*, 2014b). **(C)** Cnn stabilises Spd-2. Spd-2, in turn, recruits more Cnn; this positive feedback loop leads to further expansion of the scaffold (Alvarez-Rodrigo *et al.*, 2019).

1.1.3. Cilia

Cilia are formed in quiescent cells, when the older centriole of the pair docks at the plasma membrane and converts to a basal body (Vorobjev and Chentsov Yu, 1982). In mammalian cilia, this basal body has triplet MTs, from which the A- and B-tubule extend to produce a structure called the ciliary axoneme (Czarnecki and Shah, 2012). The transition zone (TZ) lies between the basal body and the axoneme and is important for regulating transport between the cytoplasm and the cilium; this regulation appears to be partly mediated by proteins involved in the tethering of the TZ's doublet MTs to the ciliary membrane (Craige *et al.*, 2010, Gilula and Satir, 1972, Reiter *et al.*, 2012, Ringo, 1967). **Figure 1.4** shows a schematic of a typical mammalian cilium. In vertebrates, cilia can be motile or perform signalling functions; the latter type are known as primary cilia, and defects in both classes are associated with a range of diseases, including epilepsy, retinitis pigmentosa, Joubert syndrome, and polycystic kidney disease (reviewed in Reiter and Leroux, 2017, Valente *et al.*, 2014).

In mammalian cells, the older centriole can often be identified by the presence of distal appendages (**Figure 1.2**) which are important for the docking of the centriole at the plasma membrane during ciliogenesis (Anderson, 1972, Tanos *et al.*, 2013, Vorobjev and Chentsov Yu, 1982). This process begins with the centriole attaching to the ciliary vesicle and then moving to the plasma membrane; the ciliary vesicle is then able to fuse with the membrane, thus facilitating the relocation of the centriole to the cell surface (Sorokin, 1962). *Drosophila* appendage structures have not been observed (Callaini *et al.*, 1997), though homologues of key proteins involved in their formation (such as Cep164) have been identified in flies (Graser *et al.*, 2007, Hodges *et al.*, 2010).

Two *Drosophila* cell types are ciliated:

(1) Type I mechanosensory neurons, such as those in the chordotonal organs which function as stress receptors (Kernan, 2007). Flies without centrioles cannot form cilia in these neurons, and are thus uncoordinated (Basto *et al.*, 2006).

(2) Spermatocytes. During spermatogenesis, extensions described as the 'cilium-like region' (CLR) form at the ends of all four centrioles (Gottardo *et al.*, 2013, Riparbelli *et al.*, 2012, Tates, 1971). These persist to the spermatid stage of development when they reorganise to form the sperm flagellum (Gottardo *et al.*, 2013). The axonemes of both the CLR and flagellum have doublet outer MTs attached to incomplete C-tubules (Gottardo *et al.*, 2013). The presence of these C-tubules in the CLR is essential for the flagellum formation (Gottardo *et al.*, 2018).

Again, improper centriole length can impinge on cilium formation. For example, mother centrioles with extended centriole MTs might experience difficulties in attaching to the ciliary vesicle, thus inhibiting ciliogenesis (Yadav *et al.*, 2016). This again emphasises the importance of centriole size regulation.

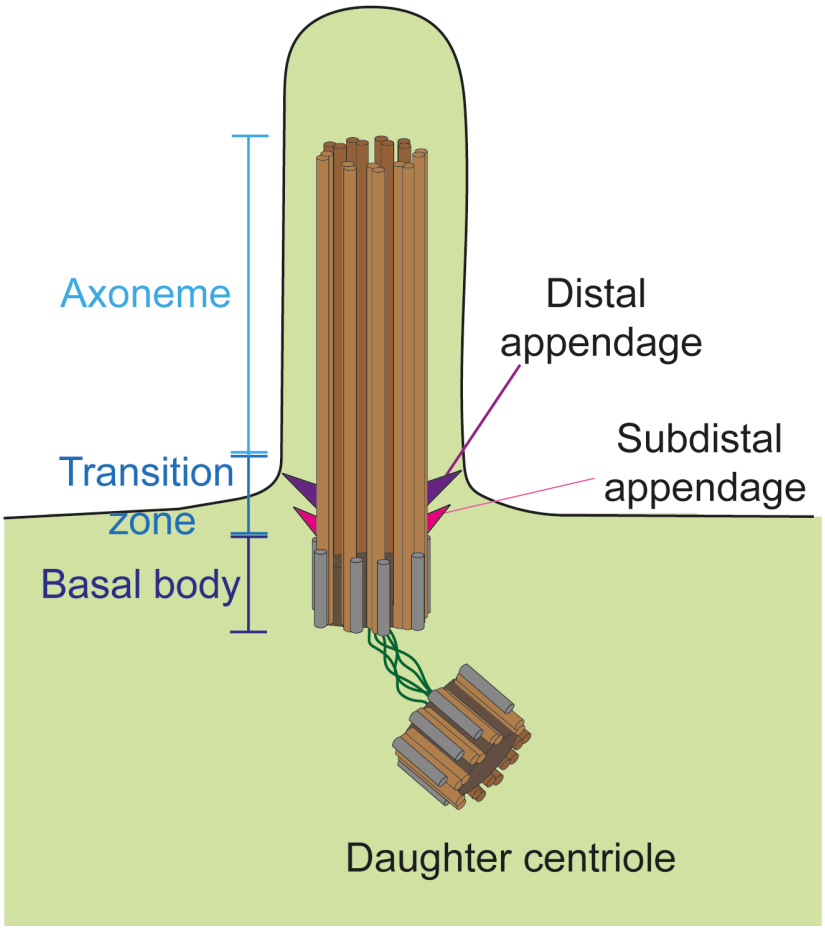


Figure 1.4 – Cilia structure.

Schematic showing a typical mammalian cilium. The basal body (the older centriole, which has docked at the plasma membrane) has triplet MTs. The doublet MTs that form the ciliary axoneme are nucleated from the A- and B-tubules of the basal body. This older centriole is decorated with distal and subdistal appendages; the former are important for the docking process. The ciliary membrane is a continuation of the plasma membrane, and the transition zone (between the basal body and axoneme) is important for regulating transport between the cilium and cytoplasm.

1.2. The centrosome cycle

Centrioles must duplicate once, and only once, per cell cycle, with daughter cells inheriting one centrosome following mitosis; this is particularly important given the

centriole's role in centrosome formation discussed in Section **1.1.2**, since an incorrect number of centrosomes can result in mitotic spindle defects and issues with cell cycle progression (reviewed in Nigg and Holland, 2018). Centrosome amplification is often associated with human tumours, and can increase the invasiveness of cancer cells (Godinho *et al.*, 2014). In addition, each mother centriole must produce just one daughter centriole per round of duplication (Nigg and Holland, 2018).

The centrosome cycle is coupled to the cell cycle, which contributes to the proper regulation of centriole duplication and also ensures that it aligns with the duplication of the cell's genetic material (Nigg and Holland, 2018, Vorobjev and Chentsov Yu, 1982). At the start of S-phase, cells have two mother centrioles connected by a proteinaceous linker. These centrioles have recruited a small amount of PCM. At this stage of the cell cycle, both mothers begin to nucleate their own daughters, which grow orthogonally to the original centriole (**Figure 1.5 A**). As the cells enter G_2, a second phase of daughter elongation commences, in which the centriole MTs grow to extend beyond the cartwheel structure (**Figure 1.5 B**). In human cells, PLK1, the homologue of *Drosophila* Polo, appears to be required for this second phase of growth in G_2 (Kong *et al.*, 2020). In most fly cells, growth in G_2 appears to be minimal (explaining why the cartwheel and centriole MTs are of similar size), but it is more extensive in spermatocyte centrioles and also seems to rely on Polo (Alvarez-Rodrigo *et al.*, 2021, Tates, 1971).

G_2-phase also sees the mother centrioles expand their PCM dramatically in preparation for mitotic onset and the centriole pairs move apart to opposite sides of the nucleus (in a process called 'centrosome separation') (**Figure 1.5 C**). During mitosis, the expanded

PCM recruits spindle MTs which attach to the sister chromatids (**Figure 1.5 D**). The daughter centriole also loses its orthogonal positioning relative to the mother and the two centrioles begin to move apart from each other (in a process called 'disengagement', indicated in **Figure 1.5 D**). Polo/PLK1 has also been shown to be important for this process, since it promotes the separation of daughters and mothers, thus licensing the mother for another round of centriole duplication (Kim *et al.*, 2015a, Shukla *et al.*, 2015, Tsou *et al.*, 2009). Daughter maturation also occurs during mitosis. This process, which also involves Polo/PLK1, is called 'centriole conversion' and will allow the daughter to produce its own daughter and to recruit its own PCM (Wang *et al.*, 2011). During centriole conversion in *Drosophila*, Sas-4 is phosphorylated by CDK1 to promote the recruitment of Polo which, in turn, facilitates the recruitment of Asl (Novak *et al.*, 2014, Novak *et al.*, 2016). In human cells, it seems that Cep152/Asl can also be recruited to, or maintained at, maturing centrioles via Cep135 and Cep295, though the role of Ana1 (the *Drosophila* Cep295 homologue) in Asl recruitment is less clear in flies (Fu *et al.*, 2016, Izquierdo *et al.*, 2014, Saurya *et al.*, 2016, Tsuchiya *et al.*, 2016). The recruitment of Asl is essential to allow new mother centrioles to form their own daughters, since Asl recruits Plk4, the kinase that facilitates centriole biogenesis (Dzhindzhev *et al.*, 2010, Novak *et al.*, 2014). As described in **Section 1.1.2**, Asl is also important for PCM recruitment (Conduit *et al.*, 2014b).

In anaphase, sister chromatids are pulled towards opposite spindle poles. When the cells divide to produce two daughter cells, each new cell inherits a set of sister chromatids as well as one centrosome. This means that, at the start of G_1, each cell has two mother centrioles (one is the old mother from the previous cycle, and the other is the new mother which has matured from the daughter formed in the last cycle – shown in **Figure 1.5 E**).

From this point, the cell might enter a new cycle (**Figure 1.5 A**) or become quiescent. The quiescent phase is known as G_0, and in some cell types involves the formation of a cilium from the older mother (**Figure 1.5 F**).

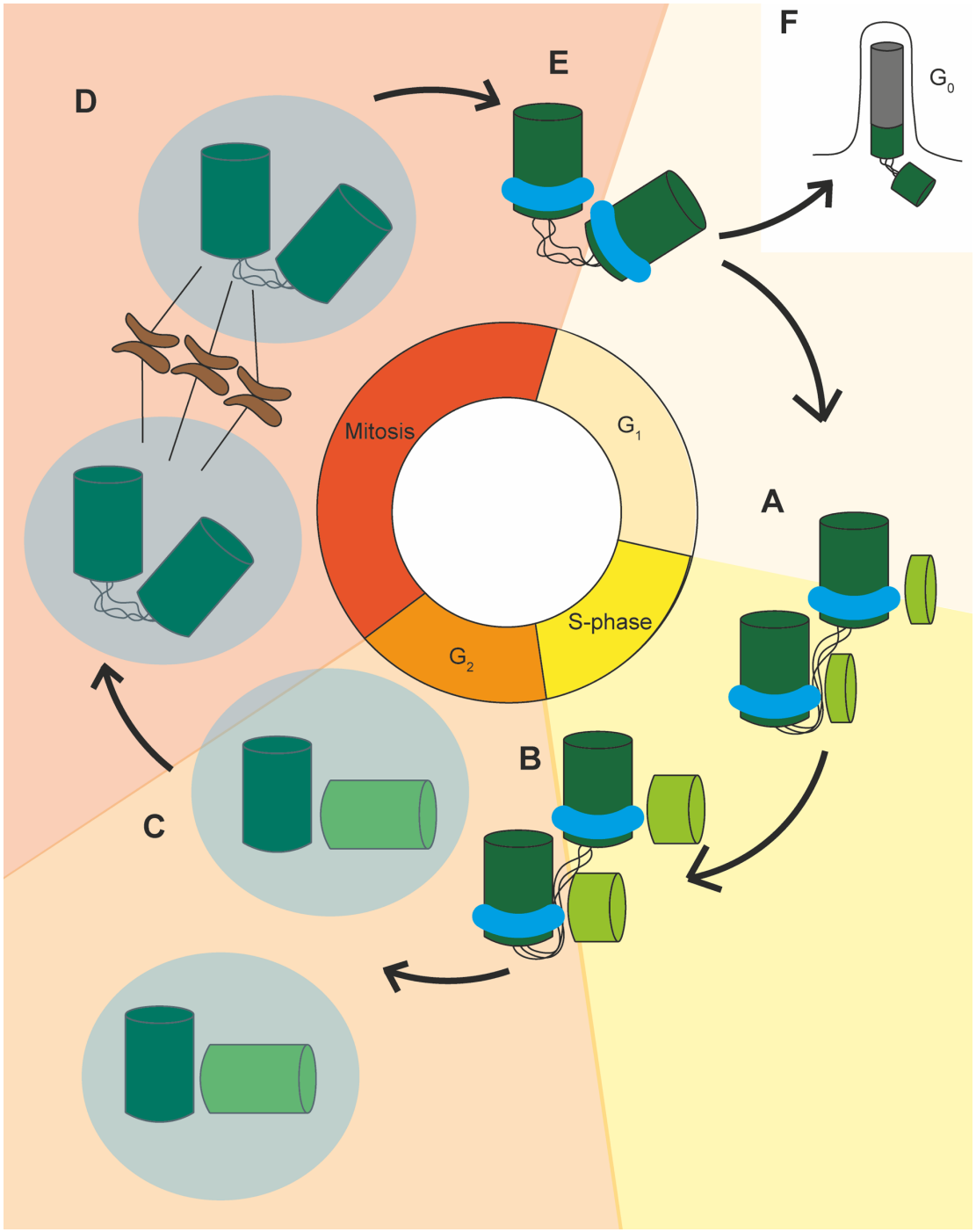


Figure 1.5 – The centrosome cycle.

Schematic summarising the centrosome cycle in vertebrate cells. **(A)** In S-phase, mother centrioles (dark green) begin to nucleate daughter centrioles (pale green) which grow orthogonally to the original mothers. **(B)** Daughters enter a second growth phase in G_2. **(C)** G_2 also sees the dramatic expansion of PCM (blue) around the mother centrioles and movement of the centrosomes to opposite sides of the nucleus (this is called 'centrosome separation'). **(D)** In mitosis, centrosomes nucleate spindle MTs which participate in the segregation of sister chromatids. During mitosis, daughters disengage from their mothers and mature. **(E)** Following mitosis, each daughter cell inherits a pair of mother centrioles, each with a small amount of surrounding PCM (blue). The cell may re-enter the cell cycle, or exit the cycle and become quiescent. **(F)** In the latter case, some centrioles go on to form cilia.

1.3. The centriole duplication pathway

As described in **Section 1.2**, centrioles duplicate once per cell cycle, with a new daughter centriole growing perpendicularly to the original mother centriole. Centrioles may also arise *de novo,* although this process is inhibited by the presence of existing centrioles (La Terra *et al.*, 2005) and the conditions in early *Drosophila* embryos do not appear to favour this pathway (Nabais *et al.*, 2021). Below, I describe the canonical centriole duplication pathway that occurs during interphase (summarised in **Figure 1.6**) and discuss how this is well-conserved across metazoans.

To understand how centriole size is regulated, it is important to consider the molecular details of centriole assembly. A series of studies identified proteins that are required for canonical centriole biogenesis in *C. elegans* (pathway summarised in **Figure 1.6 C**) by taking advantage of the worm embryo's amenability to RNAi experiments (Dammermann *et al.*, 2004, Delattre *et al.*, 2006, Delattre *et al.*, 2004, Kemp *et al.*, 2004, Kirkham *et al.*, 2003, Leidel *et al.*, 2005, Leidel and Gonczy, 2003, O'Connell *et al.*, 2001, Pelletier *et al.*, 2006, Pelletier *et al.*, 2004). Functional homologues for these proteins appear well-conserved in flies, and through to vertebrates (Breslow and Holland, 2019, Carvalho-Santos *et al.*, 2010, Conduit *et al.*, 2015b, Nabais *et al.*, 2020). Those homologues that are essential for centriole duplication in *Drosophila* are summarised in **Table 1.1.** Other proteins implicated in centriole biology that are referenced in this thesis are listed in **Table 1.2**.

In addition to the proteins identified in nematodes, centriole duplication in flies and vertebrates requires Asl/Cep152; this protein is also important for PCM expansion, as explained in **Section 1.1.2** (Blachon *et al.*, 2008, Cizmecioglu *et al.*, 2010, Dobbelaere *et al.*, 2008, Dzhindzhev *et al.*, 2010, Hatch *et al.*, 2010).

Table 1.1 – Key proteins required for centriole duplication in flies, listed alongside their functional homologues in other species.

Drosophila melanogaster	*Homo sapiens*	*Caenorhabditis elegans*
Plk4 (or Sak)	PLK4	ZYG-1
Ana2	STIL	SAS-5
Sas-6	HsSas-6	SAS-6
Sas-4	CPAP	SAS-4
Asl	Cep152	*No known homologue*

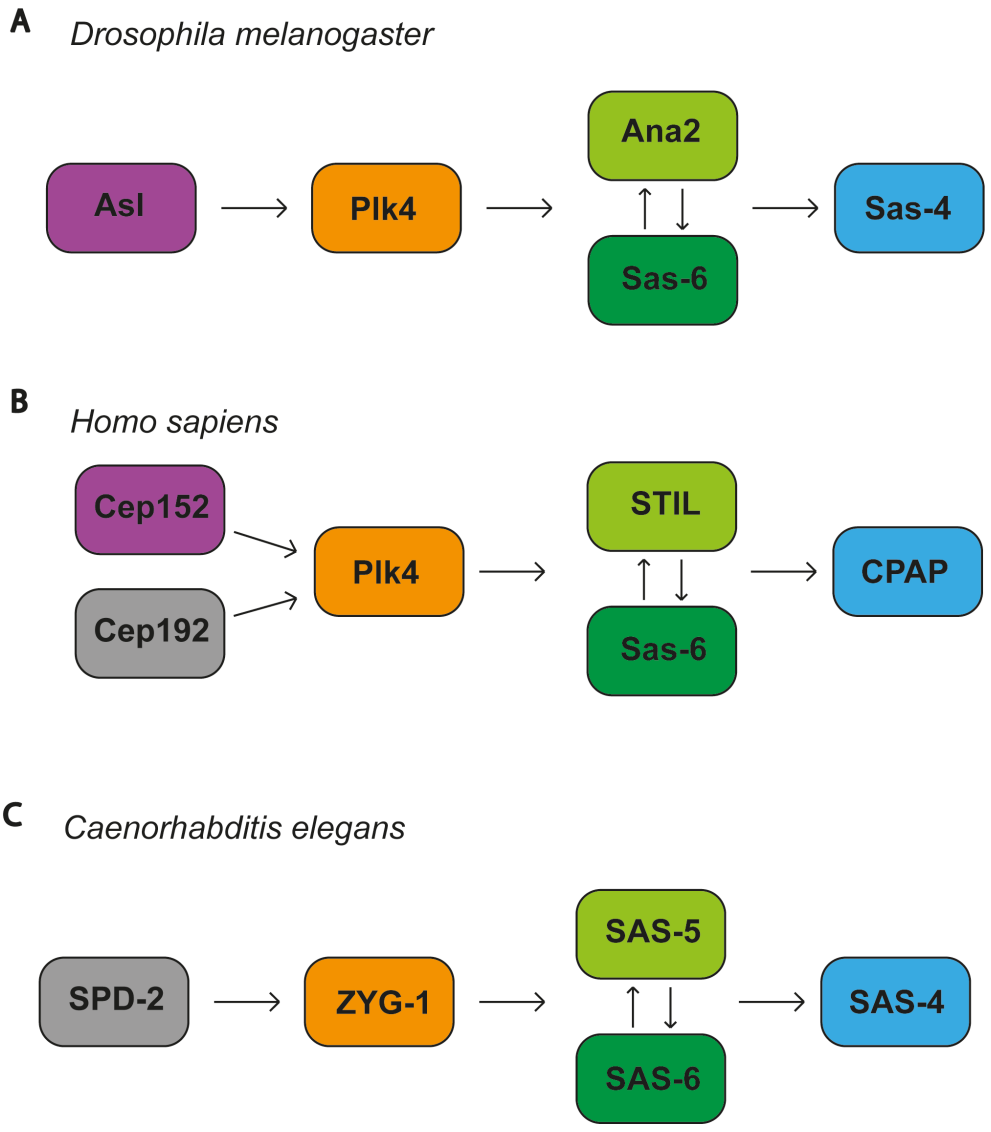


Figure 1.6 – The centriole duplication pathway is conserved across model organisms.

(A-C) Schematic showing the centriole duplication pathway in *Drosophila melanogaster* (A), *Homo sapiens* (B) and *Caenorhabditis elegans* (C). Functional homologues are shown in the same colour. As illustrated, the centriole biogenesis pathway is well-conserved from nematodes to humans (Breslow and Holland, 2019, Carvalho-Santos *et al.*, 2010, Conduit *et al.*, 2015b, Nabais *et al.*, 2020).

1.3.1. Cartwheel assembly

The first stage of centriole biogenesis is the recruitment of Polo-like kinase 4 (Plk4) (Bettencourt-Dias *et al.*, 2005, Habedanck *et al.*, 2005, Kleylein-Sohn *et al.*, 2007). Like Polo, Plk4 is a member of a serine/threonine kinase family partly defined by the presence of Polo-box (PB) domains that bind substrates via a common Ser-pSer/pThr motif (Elia *et al.*, 2003a, Elia *et al.*, 2003b, Fode *et al.*, 1994). However, Plk4 is unique amongst Polo-like kinases in that it possesses just one canonical PB as well as a 'cryptic' PB with a unique structure (Leung *et al.*, 2002, Slevin *et al.*, 2012, Swallow *et al.*, 2005).

In flies, the recruitment of Plk4 to mother centrioles is mediated by the interaction of its cryptic PB with Asl (Dzhindzhev *et al.*, 2010). In humans, an additional protein (Cep192, the human homologue of Spd-2) mediates this process, alongside the human homologue of Asl (Cep152) (Cizmecioglu *et al.*, 2010, Dzhindzhev *et al.*, 2010, Hatch *et al.*, 2010, Kim *et al.*, 2013, Sonnen *et al.*, 2013). Initially, Plk4 localises around the mother centriole as a ring, where its interaction with Ana2/STIL promotes activation of Plk4 via *trans*-autophosphorylation; this likely contributes to the resolution of the Plk4 distribution from a ring to a dot, which marks the site where the new daughter centriole will assemble (**Figure 1.7**; Arquint *et al.*, 2015, Dzhindzhev *et al.*, 2017, Kim *et al.*, 2013, Lopes *et al.*, 2015, Moyer *et al.*, 2015, Ohta *et al.*, 2014, Ohta *et al.*, 2018, Sonnen *et al.*, 2012). Once activated, Plk4 can phosphorylate Ana2/STIL in two separate domains:

(1) Plk4 phosphorylates the N-terminus of Ana2/STIL, which promotes Ana2/STIL's recruitment to the centriole (Dzhindzhev *et al.*, 2017, McLamarrah *et al.*, 2018, McLamarrah *et al.*, 2020).

(2) Plk4 phosphorylates the STIL/Ana2 (STAN) Domain of Ana2/STIL, thus promoting Sas-6 recruitment and incorporation to start building the cartwheel structure (Dzhindzhev *et al.*, 2014, Kratz *et al.*, 2015, Moyer *et al.*, 2015, Ohta *et al.*, 2014, Stevens *et al.*, 2010a). By contrast, in *C. elegans,* ZYG-1 binds SAS-6 directly to promote its recruitment (Lettman *et al.*, 2013). SAS-5, the functional homologue of Ana2/STIL in *C. elegans,* shares the conserved STAN domain (Stevens *et al.*, 2010a), though less is known about its interaction with the worm kinase ZYG-1.

Recently, the mechanisms controlling cartwheel size have been investigated, with one study demonstrating that the growth rate of the Sas-6 cartwheel is determined by the activity of Plk4 (Aydogan *et al.*, 2018). Interestingly, Plk4 levels at the centriole oscillate over the course of the cell cycle (**Figure 1.7**), and the period of this oscillation also influences the growth period of the cartwheel, at least in fly embryos (Aydogan *et al.*, 2020, Aydogan *et al.*, 2018, Takao *et al.*, 2019). As a result of this control, cartwheel growth in flies appears to be homeostatically regulated; cartwheels that grow more slowly than others also grow for a longer period, thus ensuring that centrioles grow to a consistent size (Aydogan *et al.*, 2020, Aydogan *et al.*, 2018). It has previously been shown that Plk4 can phosphorylate Asl *in vitro* and in fly cell culture (Boese *et al.*, 2018, Hatch *et al.*, 2010), and this interaction has been proposed to drive Plk4's oscillatory behaviour: mathematical models have been developed in which Asl recruits Plk4 to centrioles, where Plk4 phosphorylates itself and Asl, potentially promoting Plk4's subsequent dissociation from centrioles (Aydogan *et al.*, 2020).

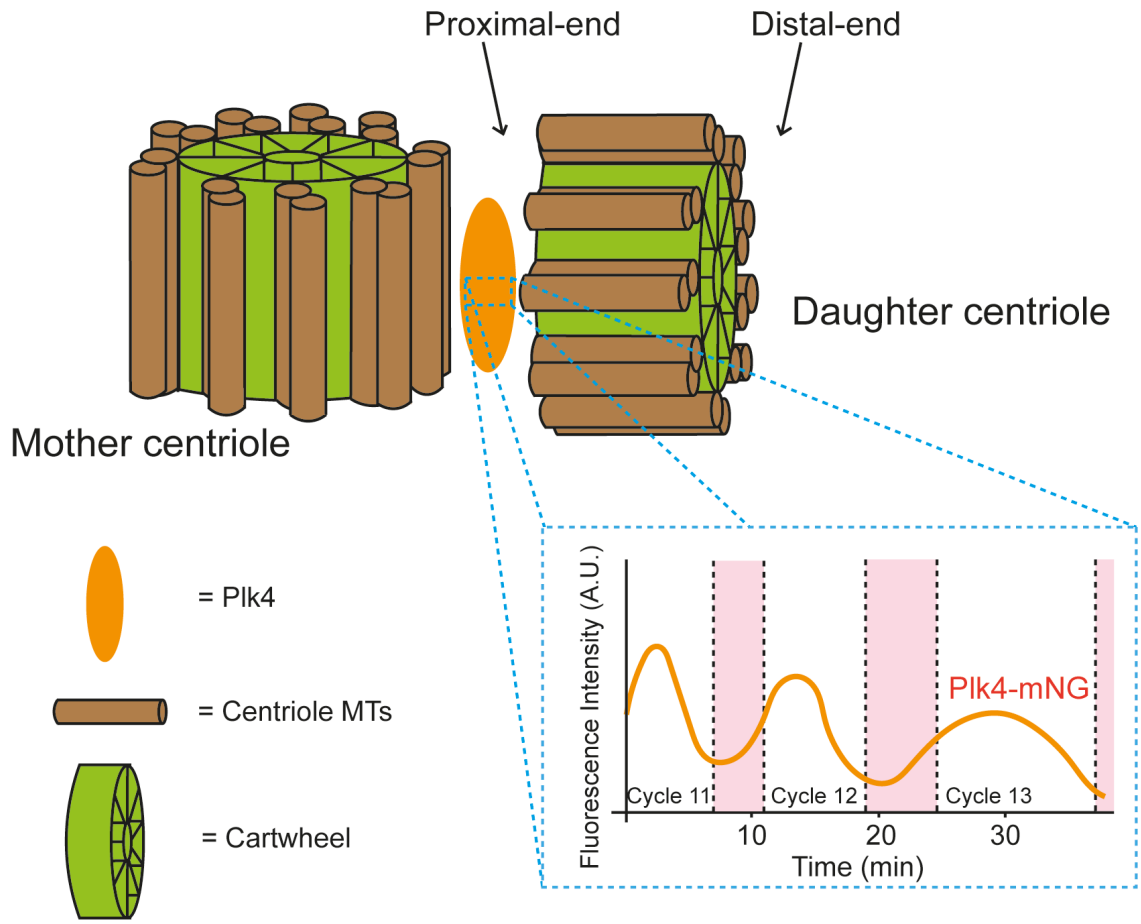


Figure 1.7 – Plk4 oscillates at the centriole.

Schematic showing that Plk4 localises to a single site on the side of the mother centriole (orange ellipse) marking the site of daughter centriole assembly. The daughter centriole then assembles orthogonally to the original mother. Graph illustrates the oscillatory recruitment dynamics of Plk4-mNeonGreen over three sequential cycles in a single syncytial early *Drosophila* embryo (adapted from Aydogan *et al.* 2020). Mitosis is indicated by the pink boxes.

Overexpression of some of these core centriole duplication factors, including Plk4 and Sas-6, can drive centriole over-duplication (Bettencourt-Dias *et al.*, 2005, Dzhindzhev *et al.*, 2010, Habedanck *et al.*, 2005, Kleylein-Sohn *et al.*, 2007, Peel *et al.*, 2007, Tang *et al.*, 2011). When activated, Plk4 phosphorylates its own phosphodegron motif, which again occurs via *trans*-autophosphorylation; this targets Plk4 for degradation by the

$SCF^{Slimb/\beta TrCP}$ ubiquitin ligase complex and seems to play an important role in controlling centriole number in cells (Cunha-Ferreira *et al.*, 2013, Cunha-Ferreira *et al.*, 2009, Guderian *et al.*, 2010, Holland *et al.*, 2012, Holland *et al.*, 2010, Klebba *et al.*, 2013, Lopes *et al.*, 2015, Rogers *et al.*, 2009, Sillibourne *et al.*, 2010). As introduced above, this degradation of Plk4 is thought to contribute to its change in localisation pattern from ring to dot at the onset of centriole biogenesis (Arquint *et al.*, 2015, Dzhindzhev *et al.*, 2017, Ohta *et al.*, 2014, Ohta *et al.*, 2018).

1.3.2. Centriole MT assembly

While the pathway of events leading to cartwheel assembly is relatively well-defined, less is known about how the centriole MTs are assembled. Sas-4/CPAP/SAS-4, one of the proteins required for centriole duplication to proceed (**Table 1.1**), is implicated in this process (Kohlmaier *et al.*, 2009, Pelletier *et al.*, 2006, Schmidt *et al.*, 2009, Tang *et al.*, 2009) and has been shown to interact with tubulin (Cormier *et al.*, 2009, Hsu *et al.*, 2008, Hung *et al.*, 2004, Sharma *et al.*, 2016). In nematodes, SAS-4 is the last of the core centriole duplication factors to load, in a process that is dependent upon SAS-5 and SAS-6 (Delattre *et al.*, 2006). Sas-4/CPAP/SAS-4 interacts with Ana2/STIL/SAS-5 in nematodes, flies, zebrafish and humans (Cottee *et al.*, 2013, Hatzopoulos *et al.*, 2013, Tang *et al.*, 2011, Vulprecht *et al.*, 2012); in the latter case, CPAP also seems dependent on Sas-6 for its recruitment to centrioles (Kohlmaier *et al.*, 2009). Together, these data suggest that the position of Sas-4/CPAP downstream of Ana2/STIL and Sas-6 in the centriole duplication pathway (**Figure 1.6**) is conserved.

Depletion of Cep135, the protein thought to form the pinhead structure linking the cartwheel to the centriole MTs, also reduces the recruitment of CPAP to centrioles and results in shorter centrioles (Lin *et al.*, 2013a). Recently, the interaction between Sas-4/CPAP and Ana2/STIL in both humans and flies was proposed to be important for connecting the centriole MTs of the growing daughter centriole to its Sas-6 cartwheel, in a process that requires the phosphorylation of Ana2/STIL by Plk4 (Moyer and Holland, 2019). These data support the idea that CPAP might form part of the pinhead (Guichard *et al.*, 2013, Hatzopoulos *et al.*, 2013).

Sas-4/CPAP appears to promote centriole MT growth, since its overexpression in human cells drives centriole over-elongation (Kohlmaier *et al.*, 2009, Schmidt *et al.*, 2009, Tang *et al.*, 2009). Furthermore, centriole over-elongation also occurs in cells where CPAP levels have risen due to depletion of Tankyrase 1, an enzyme that targets CPAP for degradation (Kim *et al.*, 2012). However, subsequent studies have revealed that this is a complex picture; in flies, Sas-4 overexpression alone does not promote centriole over-elongation (Franz *et al.*, 2013) and CPAP has even been suggested to cap centriole MTs via its LID domain, causing them to grow more slowly than their cytoplasmic counterparts (Sharma *et al.*, 2016). This latter finding could be explained by the fact that CPAP also stabilises MTs via its MT-binding domain (MBD); CPAP usually tracks MT plus-ends, but might be localising to the entire length of the centriole MTs when overexpressed, which could make its stabilisation of the MT lattice more significant than its capping function and thus promote over-elongation (Sharma *et al.*, 2016).

1.3.3. Factors involved in centriole size control

As discussed above, Sas-4/CPAP is implicated in centriole length control. The typical length of the centriole varies between difference species and cell types. For example, the centrioles of *Drosophila* are usually 180-200 nm in length (Gonzalez *et al.*, 1998, Greenan *et al.*, 2018, Moritz *et al.*, 1995). These are shorter than the centrioles of most other animals, which are typically around 400-500 nm long (Gonzalez *et al.*, 1998, Winey and O'Toole, 2014). By contrast, extremely long centrioles may be found in *Drosophila* spermatocytes, which reach around 2.5 µm (Gonzalez *et al.*, 1998, Winey and O'Toole, 2014), as well as in the unicellular flagellate *Trichonympha,* where centrioles are typically 1.5 µm long though can grow longer (Gibbons and Grimstone, 1960, Guichard *et al.*, 2012). The relative lengths of the cartwheel and centriole MTs also vary: in fly embryos, the cartwheel and the MTs are approximately the same length (**Figure 1.1 A**; Callaini *et al.*, 1997, Gonzalez *et al.*, 1998), though this is not the case in many other species. In the centrioles of the green alga *Chlamydomonas* (Guichard *et al.*, 2013) and in immature mammalian centrioles (Hirono, 2014, Vorobjev and Chentsov Yu, 1982), the cartwheel is shorter than the centriole MTs and sits at the proximal end of the overall structure. Maturing mammalian centrioles even go on to lose their cartwheels during mitosis, which suggests the cartwheel is not important for maintaining centriole stability (Arquint and Nigg, 2014, Hirono, 2014, Winey and O'Toole, 2014). By contrast, mature *Drosophila* centrioles retain their cartwheel (Gonzalez *et al.*, 1998, Moritz *et al.*, 1995). Within a given cell type, however, centriole length is remarkably consistent, with the daughter growing to reach the same size as the mother (Aydogan *et al.*, 2018, Breslow and Holland, 2019). For example, as discussed above, the Plk4 oscillation has recently been shown to regulate the growth rate and period of the cartwheel in *Drosophila* embryos to ensure it reaches a consistent size (Aydogan *et al.*, 2020).

Proper organelle sizing in general is an important physiological process; for example, the phenomenon of nuclei scaling with cell size is well-known (Cantwell and Nurse, 2019, Jorgensen *et al.*, 2007, Levy and Heald, 2010, Neumann and Nurse, 2007). Cilia, flagella, and centrioles all represent tractable models for studying this process since they can be thought of as growing in just one dimension: length (Aydogan *et al.*, 2018, Marshall, 2015, Wemmer and Marshall, 2007). Defects in centriole length, as well as number, have been implicated in some cancers (Godinho and Pellman, 2014, Marteil *et al.*, 2018). This is likely due to the important role centrioles play in centrosome formation; in human cells, over-elongated centrioles can form larger centrosomes and may also contribute to centrosome amplification by forming more than one daughter and/or fragmenting to produce extra centriole segments (Kohlmaier *et al.*, 2009, Lin *et al.*, 2013b, Marteil *et al.*, 2018). Over-elongated spermatocyte centrioles in *Drosophila* have also been reported to exhibit multi-polar spindles (Shoda *et al.*, 2021).

The over-elongation induced by Sas-4/CPAP overexpression appears to be counteracted by another protein called CP110 (Franz *et al.*, 2013, Schmidt *et al.*, 2009). Indeed, studies have suggested that CP110 (and its interaction partner Cep97) are important for the inhibition of centriole MT over-elongation (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013). By contrast, CEP120 and SPICE1 cooperate with CPAP to promote centriole elongation (Comartin *et al.*, 2013, Lin *et al.*, 2013b, Sharma *et al.*, 2018). Other proteins have also been implicated in centriole length regulation. These include Asl, whose loss may promote cartwheel elongation (Galletta *et al.*, 2016b) and Ana1/Cep295, which promotes centriole elongation, possibly by recruiting Polo kinase to mother centrioles (Alvarez-Rodrigo *et al.*, 2021, Chang *et al.*, 2016, Saurya *et al.*, 2016). Cep295 and Polo/PLK1, alongside Proteome of Centriole 5 (POC5), Anastral spindle 3 (Ana3; RTTN in humans) and

Reduction in Cnn Dots 4 (Rcd4; PPP1R35 in humans), have all been implicated in promoting the second stage of daughter centriole elongation during G_2-phase (Alvarez-Rodrigo *et al.*, 2021, Azimzadeh *et al.*, 2009, Chang *et al.*, 2016, Chen *et al.*, 2017, Kong *et al.*, 2020, Sydor *et al.*, 2018). Overexpression of POC1 also promotes centriole over-elongation, perhaps via the stabilisation of triplet MTs which makes this protein particularly important for maintaining the robustness of basal bodies forming motile cilia/flagella (Keller *et al.*, 2009, Meehl *et al.*, 2016, Pearson *et al.*, 2009). Proteins implicated in centriole length control that are referenced in this thesis are included in **Table 1.2**.

Table 1.2 – Further proteins implicated in centriole biology that are referenced in this thesis.

Protein *Where the human homologue has a different name, this is shown in brackets*	**Relevant function(s)**
Cep135	Thought to form part of the pinhead structure linking the cartwheel to the centriole MTs.
Spd-2 (Cep192)	Important for PCM scaffold formation during centrosome maturation.
Cnn (CDK5RAP2)	Important for PCM scaffold formation during centrosome maturation.
Polo kinase (PLK1)	Mitotic kinase important for centrosome maturation, centriole elongation in G_2-phase, centriole disengagement, and centriole conversion.
rha (Cep120)	Interacts with CPAP and SPICE1 in human cells to promote centriole elongation.
(SPICE1) *No known homologue in Drosophila*	Interacts with CPAP and Cep120 in human cells to promote centriole elongation.
Ana1 (Cep295)	Implicated in centriole conversion, and in centriole elongation in G_2-phase.
POC1	Implicated in centriole elongation.
(POC5) *No known homologue in Drosophila* (Heydeck, 2016)	Implicated in centriole elongation in G_2-phase.
Ana3 (RTTN)	Implicated in centriole elongation in G_2-phase.
Rcd4 (PPP1R35)	Implicated in centriole elongation in G_2-phase.
CP110	Localises to the distal-ends of centrioles where it interacts with Cep97 and inhibits centriole MT elongation and ciliogenesis.
Cep97	Localises to the distal-ends of centrioles where it interacts with CP110 and inhibits centriole MT elongation and ciliogenesis.
Cep164	A distal appendage protein implicated in regulating CP110 levels at mother centrioles.
Cep104	Interacts with CP110/Cep97 and contains a TOG domain which has been implicated in promoting MT polymerisation.
Klp10A	An MT-depolymerase of the kinesin-13 family implicated in inhibiting centriole MT over-elongation.
(Kif24) *Considered by some to be the human homologue of Klp10A* (Inaba *et al.*, 2015) *although Mitotic Centromere-Associated Kinesin (MCAK) also has some functional similarities to Klp10A and comparisons between different kinesin-13s are not always clear-cut* (Moores and Milligan, 2006).	An MT-depolymerase of the kinesin-13 family implicated in centriole MT remodelling and CP110 recruitment/maintenance.

CP110 and Cep97 and their interaction network are of particular interest, since they have been linked specifically to limiting centriole MT growth. Currently, little is known about the mechanism through which they might inhibit MT elongation. CP110 and Cep97 are the main subject of this thesis, and are described in detail below.

1.4. CP110/Cep97

CP110/Cep97 bind at the distal-ends of centrioles, where CP110 has been shown to localise as a ring in human cells (Kleylein-Sohn *et al.*, 2007, Sonnen *et al.*, 2012, Spektor *et al.*, 2007). However, little is known about how they are recruited here. Studies in human RPE-1 cells (which form cilia upon serum starvation) have demonstrated that overexpression of CP110/Cep97 can inhibit ciliogenesis, whereas depletion of these proteins promotes ciliogenesis (Spektor *et al.*, 2007, Tsang *et al.*, 2008). It was therefore suggested that CP110 and Cep97 form a 'capping complex' at the distal-ends of centrioles that prevents inappropriate ciliogenesis. Phylogenetic studies have shown that CP110 and Cep97 are often found in the same lineages, further supporting the idea that they function as a complex (Carvalho-Santos *et al.*, 2010). CP110 was postulated to inhibit the movement or attachment of the mother centriole at the plasma membrane, thus inhibiting the onset of ciliogenesis (Bettencourt-Dias and Carvalho-Santos, 2008, Tsang *et al.*, 2008). However, later studies revealed that, somewhat counterintuitively, CP110 seems to be important for attachment of the mother centriole to the ciliary vesicle, with ablation of CP110 causing the docking of the centriole at the plasma membrane to fail in the majority of cases (Song *et al.*, 2014, Walentek *et al.*, 2016, Yadav *et al.*, 2016).

In cells that usually do not form cilia, however, depletion of CP110 does not trigger ciliogenesis; rather, it appears to promote over-elongation of the centrioles (Bangs *et al.*, 2015, Kohlmaier *et al.*, 2009, Schmidt *et al.*, 2009, Tang *et al.*, 2009). Indeed, the elongated structures observed in HeLa and U2OS cells by Spektor *et al.* 2007 were subsequently reported to actually represent over-elongated centrioles, rather than true cilia (Kohlmaier *et al.*, 2009). CP110/Cep97 removal from the basal body therefore appears to be important to allow normal ciliogenesis to proceed, although it is not sufficient to trigger ectopic cilia formation in all cell types. Instead, it seems that CP110/Cep97 more generally cap centriole length; centrioles grow longer when these proteins are absent, without necessarily forming cilia. EM has allowed the ultrastructure of these elongated centrioles to be studied in *Drosophila*, and has revealed that centrioles in *cp110* null flies have slightly elongated cartwheels but dramatically longer centriole MTs (Franz *et al.*, 2013). Overall, this suggests that CP110's main role in flies is the capping of centriole MTs to prevent their over-elongation.

Intriguingly, centriole MT extensions have been observed in mouse embryonic fibroblasts lacking CP110, which resemble those observed by Franz *et al.* 2013; it has been suggested that these extensions contribute to the difficulty in the mother centriole attaching to the ciliary vesicle in these mutants (Yadav *et al.*, 2016). *cp110* null flies do not seem to have severe issues with ciliogenesis, since they are fertile and coordinated (Franz *et al.*, 2013), but this difference could be due to the fact that *Drosophila* centrioles lack the appendage structures found in mammals, which play a role in mother centriole docking (Callaini *et al.*, 1997, Franz *et al.*, 2013, Nigg and Raff, 2009).

CP110 does not seem to be essential for centriole biogenesis, since knock-outs in flies and mice still have centrioles (Franz *et al.*, 2013, Yadav *et al.*, 2016). However, CP110 depletion in human cells can inhibit centriole over-duplication normally triggered by S-phase arrest (Chen *et al.*, 2002) or Plk4 overexpression (Kleylein-Sohn *et al.*, 2007). By contrast, in flies, overexpressed centriole duplication factors were more efficient at inducing centriole over-duplication in the absence of CP110 (Franz *et al.*, 2013). These data indicate a potential complex relationship between CP110 and the core centriole duplication machinery. Indeed, CP110/Cep97 roles appear to be somewhat context-dependent; this has been suggested to be a result of CP110 forming distinct complexes that enable it to perform different functions (reviewed in Tsang and Dynlacht, 2013).

1.4.1. Regulation of CP110/Cep97 levels

Since CP110 plays such complex roles in ciliogenesis and centriole length, its levels in the cell, and at the centriole, are thought to be carefully controlled. In human cells, CP110 levels are low during G_1 and begin increasing at the onset of S-phase (Chen *et al.*, 2002). CP110 is a target of the $SCF^{CyclinF}$ ubiquitin ligase complex, which ubiquitylates CP110 in G_2 and thus targets it for degradation (D'Angiolella *et al.*, 2010). This effect is counteracted by Ubiquitin Specific Peptidase 33 (USP33), a deubiquitylase that associates with centrosomes during S- and G_2-phase to remove ubiquitin groups from CP110, thus protecting CP110 from the proteasome degradation pathway (Li *et al.*, 2013). Cep97 regulation is less well-studied, though this protein is also the target of ubiquitin ligases (Nagai *et al.*, 2018).

cp110 mRNA is also known to be the target of miRNAs, and this interaction downregulates *cp110* to promote ciliogenesis in G_0 (Bijnsdorp *et al.*, 2016, Cao *et al.*, 2012, Song *et al.*, 2014, Walentek *et al.*, 2016, Zhang *et al.*, 2015). Specifically, miR-129 miRNAs regulate *cp110* mRNA levels in cells that form a single cilium, whereas the miR-34/449 family is more important in multiciliated cells (Walentek *et al.*, 2014). The *cp110* locus is also subject to transcriptional regulation (Walentek *et al.*, 2016). Ciliogenesis can also be promoted by direct removal of CP110 protein from centrioles. M-phase phosphoprotein 9 (MPP9) co-localises with the CP110/Cep97 ring in human cells and is also removed from centrioles at the onset of ciliogenesis (Huang *et al.*, 2018, Jakobsen *et al.*, 2011). Tau tubulin kinase 2 (TTBK2) promotes the removal of CP110 from the basal body by phosphorylating MPP9, targeting it for degradation (Huang *et al.*, 2018). This appears to be crucial for promoting ciliogenesis, since *ttbk2* mutant mice fail to form cilia (Goetz *et al.*, 2012). In U2OS cells, CP110 levels at the mother centriole may already begin decreasing before docking at the plasma membrane, since centrioles bearing Cep164, a distal appendage protein, exhibit reduced levels of CP110 compared to the other three centrioles present in the cell (Sonnen *et al.*, 2012). This is likely due to the role Cep164 plays in recruiting TTBK2 (Cajanek and Nigg, 2014, Oda *et al.*, 2014). However, this network does not appear to influence centriole length, since depletion of MPP9 promotes ciliogenesis in RPE-1 cells but has no effect on centriole length in U2OS cells (Huang *et al.*, 2018).

1.4.2. CP110/Cep97 interaction partners at the distal-end

CP110 and Cep97 interact with each other (Spektor *et al.*, 2007). They have also been shown to interact with other proteins that bind at the distal-end of centrioles, including

Cep104 and Klp10A (Al-Jassar *et al.*, 2017, Delgehyr *et al.*, 2012, Jiang *et al.*, 2012, Rezabkova *et al.*, 2016).

1.4.2.1. Cep104

Six different mutations (from four patients) in the human *cep104* gene have been linked with Joubert syndrome, a rare autosomal recessive condition whose major characteristics include developmental defects of the brain (Luo *et al.*, 2019, Srour *et al.*, 2015, Valente *et al.*, 2014). Joubert syndrome has been defined as a ciliopathy since many of its known associated mutations have been identified in genes encoding centriole/primary cilia proteins (Reiter and Leroux, 2017, Valente *et al.*, 2014). Indeed, Cep104 is implicated in ciliogenesis; it tracks the growing ciliary tip during axoneme extension, and depletion of Cep104 can result in reduced cilia formation (Frikstad *et al.*, 2019, Jiang *et al.*, 2012, Louka *et al.*, 2018, Satish Tammana *et al.*, 2013). Those cilia that do form in *cep104* mutant RPE-1 cells are usually shorter than WT cilia (Frikstad *et al.*, 2019). In *Chlamydomonas,* flagella formation is also inhibited by mutation of the Cep104 homologue Flagellar-Associated Protein 256 (FAP256); again, those flagella that do form are shorter than WT, and EM has revealed that they exhibit structural defects at their tips (Satish Tammana *et al.*, 2013). Similar defects are observed in the cilia of *fap256 Tetrahymena* mutants (Louka *et al.*, 2018). It has been postulated that these changes to the structure of the cilium's distal segment upon Cep104 depletion may cause defective primary cilium signalling (Louka *et al.*, 2018). Support for this idea came from Frikstad *et al.* 2019 and Yamazoe *et al.* 2020, who showed that Cep104 is crucial for proper Hedgehog signalling at the ciliary tip, thus providing a possible explanation for the developmental defects observed in individuals with Joubert syndrome.

Cep104 is not required for the removal of CP110 from the basal body at the onset of ciliogenesis, since *cep104* mutant RPE-1 cells that failed to form cilia were still able to remove CP110 from the mother centriole (Frikstad *et al.*, 2019). Similar results have also been reported by Yamazoe *et al.* 2020. Instead, Cep104's role in promoting ciliogenesis seems to primarily involve axoneme extension; FAP256 has been shown to bind both the A-tubules and the central pair MTs of cilia, where, in *Tetrahymena,* it appears to specifically promote A-tubule extension (Louka *et al.*, 2018, Satish Tammana *et al.*, 2013). This role in axoneme extension is likely mediated via Cep104's TOG domain, which can bind tubulin and may promote its incorporation at the growing tip of the cilium in an analogous mechanism to that of the XMAP215 and CLASP proteins (Akhmanova and Steinmetz, 2015, Al-Bassam and Chang, 2011, Al-Jassar *et al.*, 2017, Louka *et al.*, 2018, Rezabkova *et al.*, 2016, Yamazoe *et al.*, 2020).

In non-ciliated cells, Cep104 binds at the distal-ends of both mother and daughter centrioles, where it co-localises with the CP110 ring (Frikstad *et al.*, 2019, Gilani, 2018, Jakobsen *et al.*, 2011, Jiang *et al.*, 2012, Satish Tammana *et al.*, 2013). Human Cep104 has been shown to directly interact with both CP110 and Cep97 (Al-Jassar *et al.*, 2017, Jiang *et al.*, 2012, Rezabkova *et al.*, 2016). These interactions are summarised in **Figure 1.8**.

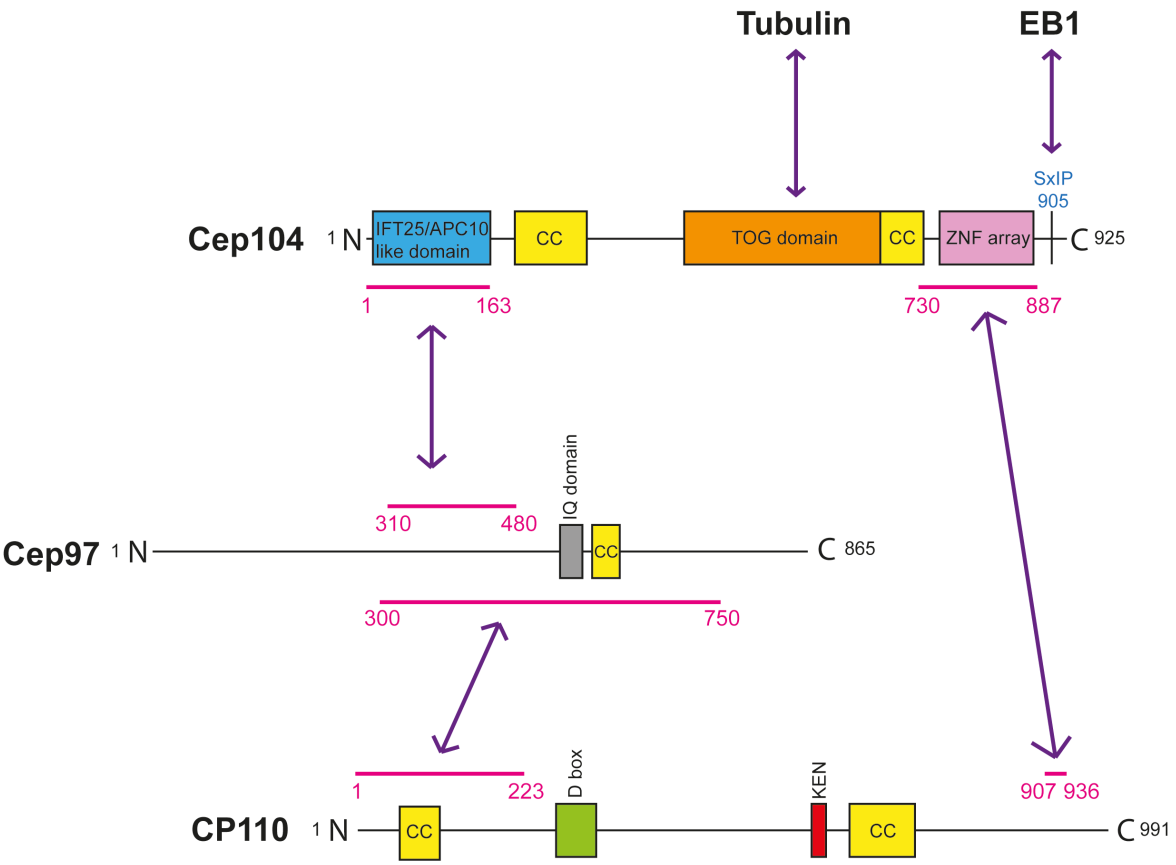


Figure 1.8 – Schematic summary of the Cep104 interaction network.

Schematic shows the interactions between the human homologues of Cep104, CP110, Cep97 and EB1 as described in the main text. Interactions are indicated with purple arrows, and the borders of the interaction regions are delineated in pink. Briefly, Cep104's SxIP motif directly interacts with EB1 (Rezabkova *et al.*, 2016). The Cep104 TOG domain also interacts with tubulin (Al-Jassar *et al.*, 2017, Rezabkova *et al.*, 2016). Cep104's ZNF array interacts with the C-terminus of CP110, while Cep104's IFT25 domain can interact with the mid-region of Cep97 (Al-Jassar *et al.*, 2017, Jiang *et al.*, 2012, Rezabkova *et al.*, 2016). Finally, CP110 and Cep97 also directly interact with each other (Spektor *et al.*, 2007)

Since CP110/Cep97 have not been shown to interact with MTs directly, it has been postulated that Cep104 might facilitate CP110/Cep97's binding at the growing ends of centriole MTs (Jiang *et al.*, 2012, Rezabkova *et al.*, 2016, Satish Tammana *et al.*, 2013). This might be mediated via Cep104's SxIP domain, which interacts with EB1, a protein that tracks the growing ends of MTs (Jiang *et al.*, 2012, Maurer *et al.*, 2011, Rezabkova *et*

al., 2016, Tirnauer and Bierer, 2000). Indeed, Jiang *et al.* 2012 showed that overexpression of Cep104 could facilitate the recruitment of Cep97 to MTs. However, the SxIP motif is not present in flies (Al-Jassar *et al.*, 2017), and immunofluorescence experiments have not identified EB1 at the distal tip of centriole MTs in *Drosophila* (Saroj Saurya, personal communication). In addition, EB1 binding does not appear necessary for proper Cep104 recruitment, since depletion of EB1 does not perturb Cep104 localisation in either ciliated or non-ciliated human RPE-1 cells (Yamazoe *et al.*, 2020). Consequently, it seems unlikely that an interaction between Cep104 and EB mediates the recruitment of Cep104, and by extension CP110/Cep97, to centriole MTs in *Drosophila.* The precise role of the interaction between Cep104 and CP110/Cep97 therefore remains to be discovered.

1.4.2.2. Klp10A

Kinesin-like protein at 10A (Klp10A) is a member of the kinesin-13 family, a group of MT depolymerases whose ATPase motor domains have been co-opted for this depolymerase activity (reviewed in Moores and Milligan, 2006). It was first discovered as an important player at the mitotic spindle, localising to the minus-ends of MTs in anaphase and promoting their depolymerisation in order to facilitate the segregation of sister chromatids (Goshima and Vale, 2003, Rogers *et al.*, 2004). Depletion of Klp10A results in over-elongated, denser mitotic spindles (Goshima and Vale, 2003, Goshima and Vale, 2005, Rogers *et al.*, 2004) and similar phenotypes have also been observed in the meiotic spindle of *Drosophila* oocytes, where no centrosomes are present (Radford *et al.*, 2012).

Klp10A's localisation changes in interphase, when Klp10A associates with the plus-ends of MTs; this recruitment is promoted by an interaction with EB1 (Goshima and Vale,

2005, Mennella *et al.*, 2005). Once recruited, Klp10A seems to bind to MTs directly, since loss of EB1 at this stage does not result in loss of Klp10A from the plus-ends (Mennella *et al.*, 2005). It was initially suggested that kinesin-13s bind more favourably to bent protofilaments than to straight ones (Ogawa *et al.*, 2004). Klp10A does appear to be able to bind straight protofilaments, but can only close its nucleotide binding pocket (thus facilitating ATP hydrolysis) when bound to bent protofilaments (Benoit *et al.*, 2018). EB1 binding does not promote Klp10A's recruitment to the minus-ends of the spindle during mitosis, suggesting that the dynamic localisation of Klp10A might be driven by its affinity for EB1 varying in a cell cycle-dependent manner (Mennella *et al.*, 2005).

In addition to its roles at the mitotic spindle, Klp10A appears to interact with CP110 and has been shown to influence centriole length, with depletion of Klp10A promoting centriole over-elongation (Delgehyr *et al.*, 2012, Franz *et al.*, 2013). Shorter, fragmented centrioles or centrioles with ragged ends are also observed in *klp10A* mutant spermatocytes and Klp10A-depleted Dmel cells, suggesting that depletion of Klp10A reduces centriole stability; this appears to be due to the over-elongated centrioles being more susceptible to fragmentation from cellular forces, since co-depletion of dynein abolished the formation of these shorter centrioles (Delgehyr *et al.*, 2012). Indeed, a later study observed similar defects in *klp10A* mutant spermatocytes and proposed that the longer centrioles were mothers, whereas the shorter centrioles were daughters (Gottardo *et al.*, 2016). Gottardo *et al.* 2016 also observed that the shorter daughters were displaced from the proximal-end of their elongated mothers, as well as noting signs of centriole instability which again suggests that daughters are shorter in these mutants due to reduced stability and thus fragmentation. In *Drosophila* male germline stem cells (GSCs), depletion of Klp10A also leads to centriole elongation but this is restricted to the centrioles of the older centrosome

and thus results in asymmetric spindle formation in mitosis, which produces daughter cells of different sizes (Chen *et al.*, 2016).

In *klp10A* mutant wing discs, the centriole MTs are dramatically extended compared with WT, which is reminiscent of the *cp110* null phenotype described above (Franz *et al.*, 2013). Similar MT elongations have also been reported on mother centrioles in *klp10A* mutant spermatocytes (Gottardo *et al.*, 2016). Interestingly, Franz *et al.* 2013 noted that the MT extensions observed in *klp10A* mutant wing discs appear to be extended doublets, whereas the MT extensions in *cp110* null wing discs represent singlets. Overall, these data suggest that CP110 and Klp10A cooperate to inhibit the over-elongation of centriole MTs.

1.4.2.3. Kif24

In mammals, another kinesin-13 called Kif24 localises to the mother centriole where it remodels centriole MTs to inhibit cilia formation (Kim *et al.*, 2015b, Kobayashi *et al.*, 2011). Kif24 interacts with CP110/Cep97, and its depletion in RPE-1 cells results in the loss of CP110 from the mother and promotes ciliogenesis (Kobayashi *et al.*, 2011). This suggests that Kif24 cooperates with CP110 to inhibit ciliogenesis but, unlike Klp10A, Kif24 does not influence centriole length in non-ciliated cells (Kobayashi *et al.*, 2011).

Interestingly, Klp10A has also been implicated in ciliogenesis; in addition to elongated centrioles, *klp10A* mutant spermatocytes have disrupted transition zone organisation (Gottardo *et al.*, 2013), and a recent study has postulated that this kinesin could form part of the ciliary transition zone in *Drosophila* (Persico *et al.*, 2019).

1.5. Tracking centriole protein dynamics in syncytial *Drosophila* embryos

1.5.1. The centrosome cycle in *Drosophila* blastoderms

The *Drosophila* blastoderm stage constitutes nuclear cycles 10-13 and, as development progresses, the length of S-phase in these cycles increases (Foe and Alberts, 1983). These early cell cycles lack G_1- or G_2-phases (comprising only S-phase and mitosis) and the embryo at this stage is syncytial, with no cell membranes forming between individual nuclei (Foe and Alberts, 1983). The first G_2-phase occurs in cycle 14, when cellularisation begins, and, by cycle 17, there is also a G_1-phase (Edgar and O'Farrell, 1990). This means that the centrosome cycle of the syncytial embryo, while similar to the typical cycle presented above, takes place over a more rapid timescale, with mother centrioles already moving apart at the end of mitosis (**Figure 1.9 C**; Callaini and Riparbelli, 1990). This process (termed 'centrosome separation') normally occurs in G_2-phase (**Figure 1.5 C**). Daughters are nucleated from these separated mothers as soon as mitotic exit is achieved, at S-phase onset (**Figure 1.9 A**; Callaini and Riparbelli, 1990).

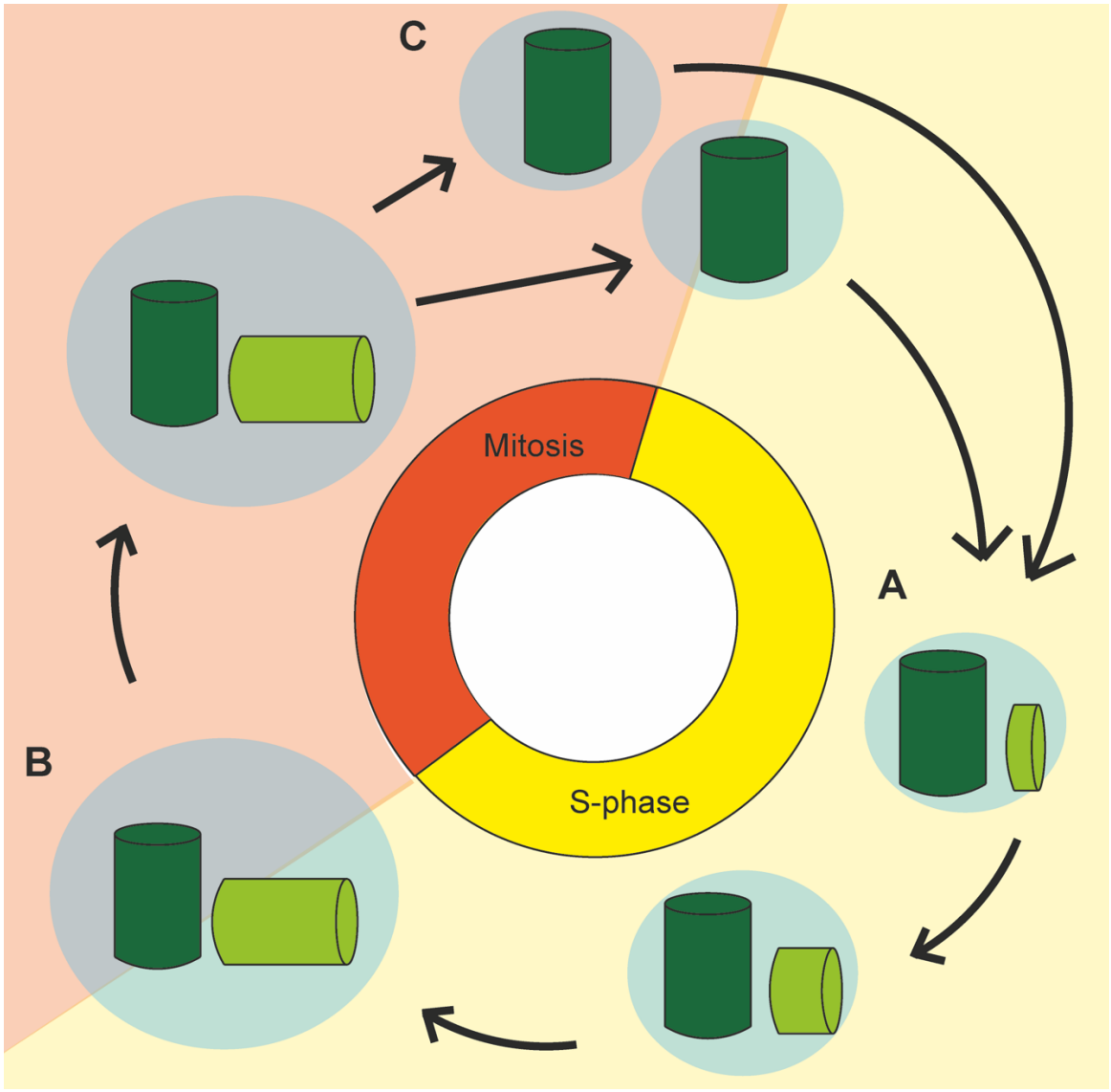


Figure 1.9 – The centrosome cycle in early *Drosophila* embryos.

Schematic summarising the centrosome cycle in the *Drosophila* early embryo. **(A)** In S-phase, mother centrioles (dark green) begin to nucleate daughter centrioles (pale green) which grow orthogonally to the original mothers. **(B)** In preparation for mitosis, the PCM (blue) expands. **(C)** Centrosome separation occurs at the end of mitosis, with each mother centriole going on to immediately nucleate its own daughter as they re-enter S-phase.

1.5.2. An assay for tracking centriole cartwheel growth

The syncytial nature of the early *Drosophila* embryo is advantageous for imaging experiments, since cell cycles (and thus centriole duplication events) within this common cytoplasm are approximately synchronised (Foe and Alberts, 1983). In cycles 11-13, these centrioles and their nuclei are also positioned near the cortex of the embryo, (Foe and Alberts, 1983) which makes it possible to image hundreds of centriole pairs within a single plane. This allows us to take a quantitative approach, following the dynamic recruitment of tagged proteins over the course of one/multiple cycles. As described above, centrosome separation (CS) occurs at the end of mitosis in fly embryos, while the start of mitosis is indicated by nuclear envelope breakdown (NEB). We can therefore define S-phase length in our movies as the time from CS to NEB.

This technique has proved particularly useful for our laboratory in the development of a cartwheel growth assay. As described in **Section 1.1.1.1**, Sas-6 is a key structural component of the centriole's cartwheel, and is stably and irreversibly incorporated into the daughter centriole as it grows in S-phase; this means that the increase in Sas-6-GFP fluorescence intensity over the course of S-phase can be measured as a proxy for centriole cartwheel size (Aydogan *et al.*, 2018). An illustration of how this assay works is provided in **Figure 1.10**. This approach can also be generalised to study the recruitment dynamics of other centriole proteins, provided that these proteins are engineered with a fluorescent tag.

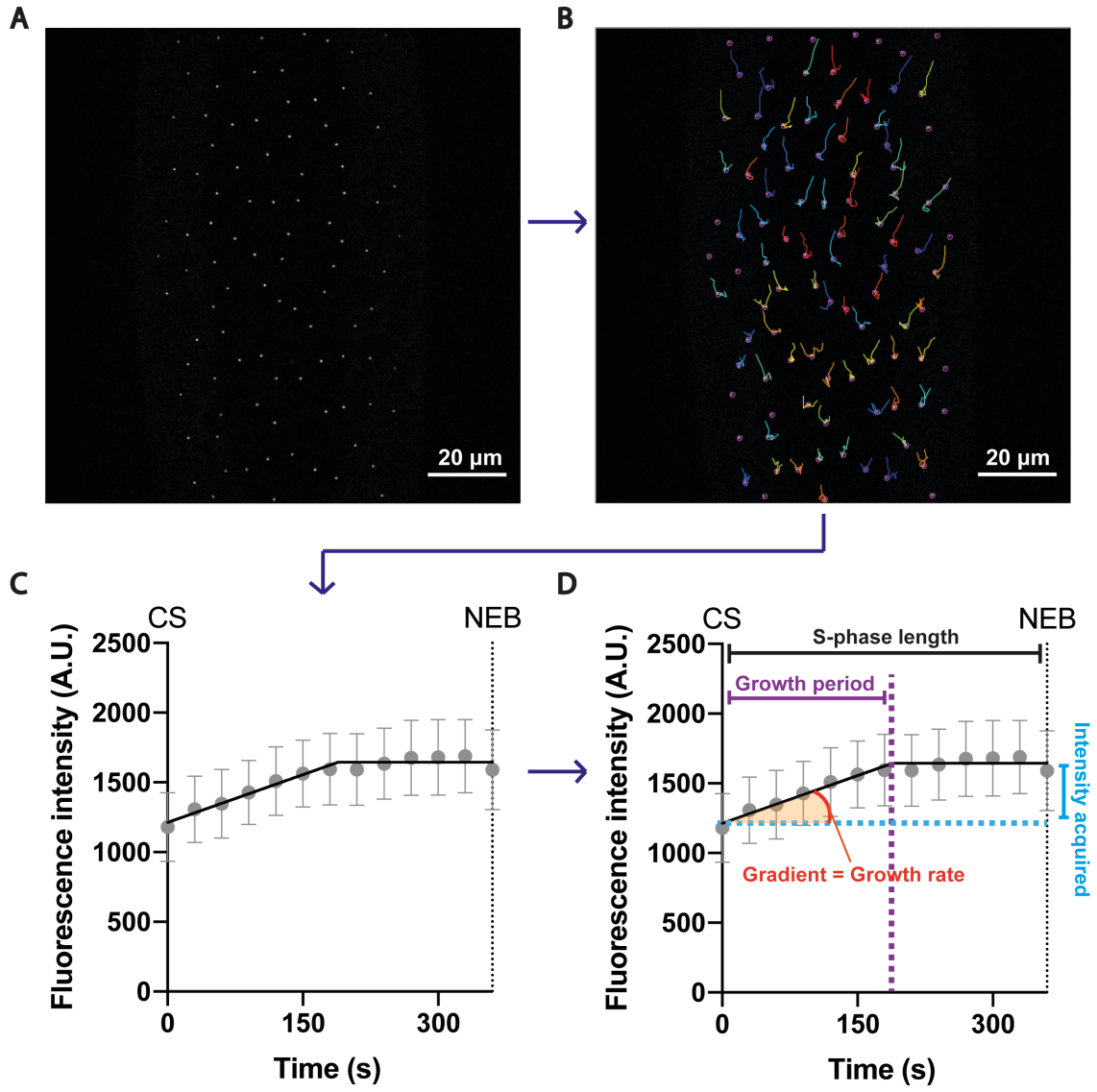


Figure 1.10 – An assay to monitor the cartwheel growth dynamics of daughter centrioles.

Figure showing the methodology behind the Sas-6-GFP incorporation assay, where the increase in Sas-6-GFP fluorescence intensity over the course of S-phase can be measured as a proxy for centriole cartwheel size (see **Chapter 2** for additional details). **(A)** Micrograph showing an embryo expressing Sas-6-GFP in S-phase of cycle 12. Puncta correspond to pairs of centrioles. The image is shown as a maximum intensity projection with bleach correction and background subtraction. **(B)** Trackmate (Tinevez *et al.*, 2017) is used to track each centriole pair and to quantify how the fluorescence intensity of each pair increases over time. Pink circles indicate detected centriole pairs. Rainbow-coloured lines represent the tracks these pairs make over the analysis period. **(C)** Graph showing the mean change in centrosomal Sas-6-GFP fluorescence intensity in a single embryo over S-phase of cycle 12, calculated from the fluorescence intensity of the pairs shown in (B). The data can be modelled with a *Linear increase followed by a plateau* function (Aydogan *et al.*, 2018). This regression is plotted here as a black line. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SD. **(D)** Once the *Linear increase followed by a plateau* function is fitted, it can be used to extract the following cartwheel growth parameters: S-phase length (time from CS to NEB, indicated by the black bracket), growth period (time to plateau, indicated by the purple bracket), growth rate (gradient of the line, indicated in orange), and final cartwheel size (intensity acquired, indicated by the blue bracket).

This assay is a key asset for studies investigating the regulation of cartwheel growth. However, no equivalent assay exists for the study of centriole MT growth, with prior studies into this limited to EM approaches (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013).

1.6. Thesis outline

In this thesis, I explore how centriole size is regulated by CP110/Cep97 and their interaction partners. Specifically, I began my studies by focussing on the regulation of centriole MT size. A key aim of this study was to investigate the dynamic growth of these centriole MTs, which has not previously been possible; in **Chapter 3**, my collaborators and I developed an assay to achieve this in living *Drosophila* embryos. This revealed that centriole MTs grow slower and reach a shorter length when CP110/Cep97 are overexpressed, and may also grow for a longer period. We also showed that CP110/Cep97 oscillate at the centriole; however, these dynamics are not important for directly regulating centriole MT growth. Instead, they appear to play a surprising role in influencing cartwheel growth, possibly via an interaction with Plk4. This raised the question of how such an interaction might be mediated, and how CP110/Cep97 are recruited to the centriole MTs in the first place. In the Chapters that follow, I present my studies of how CP110/Cep97 recruitment might be mediated by their interaction network, and the overall role of this network in regulating centriole size: in **Chapter 4**, I investigated the function of Cep104 (a protein postulated to recruit CP110/Cep97 to centriole MTs) by generating a *cep104* mutant fly line. My results suggest that Cep104 actually acts downstream of CP110/Cep97. In **Chapter 5**, I present preliminary work investigating the role of Klp10A in CP110/Cep97 recruitment, and in centriole growth. In **Chapter 6**, I found that CP110/Cep97 mutants unable to bind Polo could still influence centriole size, and that the

CP110/Cep97 oscillation can still occur even when the cell cycle machinery is perturbed. Finally, in **Chapter 7**, I summarise my findings and propose an overall model of centriole growth control, in which the cartwheel drives centriole MT growth and CP110/Cep97 recruitment dynamics are partly mediated by their affinity for growing MTs. I propose that, once daughter centrioles reach their final size, subsequent roles of CP110/Cep97 and their interaction network may come into play.

Chapter 2: Materials and Methods

2.1. Fly husbandry

2.1.1. Fly handling

Flies were kept at 25°C/18°C in plastic vials/bottles containing *Drosophila* culture medium (0.77% agar, 6.9% maize, 6.9% malt, 1.9% molasses, 1.4% yeast, 0.8% soya, 0.5% propionic acid, 0.3% nipagin, 0.03% orthophosphoric acid). Standard fly handling techniques were used (Roberts, 1998).

2.1.2. Embryo handling for live imaging

Embryos were collected for one hour on fruit juice plates (25% cranberry-raspberry juice, 2.5% sucrose, 1.8% agar) then incubated at 25°C for 45 min. Embryos were dechorionated by hand using double-sided tape and mounted onto a strip of glue on 35 mm glass-bottomed Petri dishes with 14 mm micro-wells (MatTek). After desiccation for 30 s at 25°C, embryos were covered in Voltalef H10S oil (ARKEMA).

For mRNA injection, embryos were collected for 20 min, then dechorionated and mounted as described above. For drug injection, embryos were collected for 1 h on fruit juice plates then incubated for 30 min before being dechorionated and mounted as described above. For RNA interference, embryos were collected for 20 min and incubated at 25°C until they

were 70 min old, then dechorionated and mounted as described above. For all injection experiments, the desiccation time was increased to 3-5 min.

2.1.3. *Drosophila melanogaster* stocks

Oregon-R flies were used as the wild type (WT) stock in all experiments. Balancer chromosomes and phenotypic markers used here have been previously described (Lindsley and Zimm, 1992). All fly stocks used in this study are listed in **Table 2.1**.

Table 2.1 – *Drosophila* lines used in this thesis.

Chapter	Genotype	Description	Reference
3/4/6	Ubq-CP110-GFP	Line expressing CP110-GFP under the ubiquitin (Lee *et al.*, 1988) promoter	(Dobbelaere *et al.*, 2008)
3/4/5	Ubq-GFP-Cep97	Line expressing GFP-Cep97 under the ubiquitin promoter	(Dobbelaere *et al.*, 2008)
3/4/6	Ubq-RFP-Cep97	Line expressing RFP-Cep97 under the ubiquitin promoter	(Dobbelaere *et al.*, 2008)
3	eCP110-GFP	Line expressing CP110-GFP under the endogenous *cp110* promoter	This thesis
3	eCep97-GFP	Line expressing Cep97-GFP under the endogenous *cep97* promoter	(Dobbelaere *et al.*, 2020)
3/4/6	*cp110* null	Line where the entire *cp110* coding region has been deleted via FLP-induced recombination of two P-element insertions	(Franz *et al.*, 2013)
3/4	*cep97* null	Line where the entire *cep97* coding region has been replaced with the white marker gene via ends-out gene targeting	(Dobbelaere *et al.*, 2020)
3	Ubq-CP110	Line expressing untagged CP110 under the ubiquitin promoter	Generated by Saroj Saurya (Raff Lab, University of Oxford)

3	Ubq-Cep97	Line expressing untagged Cep97 under the ubiquitin promoter	Generated by Saroj Saurya
3/6	eAsl-mCherry	Line expressing Asl-mCherry under the endogenous *asl* promoter	(Conduit *et al.*, 2015a)
3/6	ePlk4-GFP	Line expressing Plk4-GFP under the endogenous *plk4* promoter	(Aydogan *et al.*, 2018)
3	ePlk4-mNeonGreen	Line expressing Plk4-mNeonGreen under the endogenous *plk4* promoter	(Aydogan *et al.*, 2020)
3	$plk4^{\Delta a74}$	Line expressing a mutated version of the *plk4* gene where the first third of the coding sequence has been deleted via imprecise P-element excision	(Aydogan *et al.*, 2018)
3	$plk4^{RKA}$	Line expressing a mutated version of the *plk4* gene with reduced kinase activity, generated by introducing the L89G substitution (Holland *et al.*, 2010, Zitouni *et al.*, 2016)	(Aydogan *et al.*, 2018)
3	ePlk4	Line expressing untagged Plk4 under the endogenous *plk4* promoter	(Aydogan *et al.*, 2018)
3/4/5	eSas-6-GFP	Line expressing Sas-6-GFP under the endogenous *sas-6* promoter	(Aydogan *et al.*, 2018)
3	Ubq-GFP-Cep97, *cp110* null	Recombinant on Chromosome I of the *cp110* null mutation and GFP-Cep97 expressed under the ubiquitin promoter	Generated by Mustafa Aydogan (Raff Lab, University of Oxford)
3	eAslmCherry, *cep97* null	Recombinant on Chromosome II of the *cep97* null mutation and Asl-mCherry expressed under the endogenous *asl* promoter	Generated by Mustafa Aydogan
3	Ubq-CP110-GFP, *cep97* null	Recombinant on Chromosome II of the *cep97* null mutation and CP110-GFP expressed under the ubiquitin promoter	Generated by Mustafa Aydogan
3	eSas-6-GFP, *cep97* null	Recombinant on Chromosome II of the *cep97* null mutation and Sas-6-GFP expressed under the endogenous *sas-6* promoter	Generated by Mustafa Aydogan

3	ePlk4-GFP, *cep97* null	Recombinant on Chromosome II of the *cep97* null mutation and Plk4-GFP expressed under the endogenous *plk4* promoter	Generated by Mustafa Aydogan
4	eSas-6-mNeonGreen	Line expressing a CRISPR/Cas9 knock-in of the mNeonGreen tag at the *sas-6* gene locus	Generated by Thomas Steinacker (Raff Lab, University of Oxford)
4	nos-Cas9	Line expressing Cas9 under the nos promoter	(Port *et al.*, 2014)
4	Cep104 gRNA	Line expressing the 3' and 5' gRNAs (designed for *cep104* deletion generation) under U6 promoters	This thesis
4	*cep104* mutant	Line homozygous for a CRISPR/Cas9 deletion in the C-terminus of the *cep104* gene	This thesis
4	Ubq-Cep104-GFP	Line expressing Cep104-GFP under the ubiquitin promoter	Generated by Metta Pratt and Helio Roque (Raff Lab, University of Oxford)
4	Ubq-Cep104	Line expressing untagged Cep104 under the ubiquitin promoter	Generated by Saroj Saurya
4	$ana1^{mecB}$	Line expressing a mutated version of the *ana1* gene identified as part of a genetic screen	(Avidor-Reiss *et al.*, 2004)
4	*ana3* mutant	Ana3 CRISPR/Cas9 knockout line where the entire *ana3* coding region has been deleted	Generated by Saroj Saurya
4	*rcd4* mutant	Rcd4 CRISPR/Cas9 knockout line where the majority of the *rcd4* coding region has been deleted; the remaining sequence codes for a small peptide of 22 amino acids	Generated by Saroj Saurya
4	$poc1^{W87}$	Line expressing a mutated version of the *poc1* gene, where W87 has been replaced by a stop codon; this is predicted to cause loss of function	(Khire *et al.*, 2016)

5	*klp10A* mutant	Line with reduced Klp10A expression due to disruption by P-element insertion	(Delgehyr *et al.*, 2012)
5	Ubq-Klp10A	Line expressing untagged Klp10A under the ubiquitin promoter	Generated by Saroj Saurya
5	Ubq-Klp10A-GFP	Line expressing Klp10A-GFP under the ubiquitin promoter	Generated by Saroj Saurya
6	Ubq-Cep97-WT-GFP	Line expressing WT Cep97-GFP under the ubiquitin promoter	This thesis, in collaboration with Ines Alvarez Rodrigo
6	Ubq-CP110-PBM-GFP	Line expressing CP110-GFP under the ubiquitin promoter, where putative Polo-binding sites have been mutated to inhibit binding	This thesis, in collaboration with Ines Alvarez Rodrigo
6	Ubq-Cep97-PBM-GFP	Line expressing Cep97-GFP under the ubiquitin promoter, where putative Polo-binding sites have been mutated to inhibit binding	This thesis, in collaboration with Ines Alvarez Rodrigo

2.1.4. Transgenic line generation

Transgenic integration of gRNAs for generation of a *cep104* CRISPR knockout (**Section 2.1.6**) was achieved via *phiC31* integrase-mediated transgenesis in *attP40* embryos. All other transgenic fly lines were generated via random *P*-element insertion of the chosen construct into a w^{1118} background. This construct contained a w^{+} gene for selection purposes. Plasmid generation is described in **Section 2.2.2**. Injection, selection of transformants, establishment of the transgenic lines and mapping of the insertion were performed by the Cambridge Fly Facility (CFF; Department of Genetics, University of Cambridge). Transgenic lines produced for this study are included in **Table 2.1**.

2.1.5. Recombinant line generation

Flies for recombination were crossed as required, with the resulting female double heterozygous larvae being kept at 30°C for 48 h. This treatment aimed to increase the probability of a recombination event occurring (Plough, 1921). Any resulting recombination events were scored via eye colour. Recombinant lines used in this study are included in **Table 2.1**.

2.1.6. CRISPR knockout line generation

The last 5 exons of *cep104* were targeted, since this region does not overlap with other genes (see **Chapter 4** for more details on selecting the target region). This region was sequenced in nos-Cas9 flies (into which guide RNAs (gRNAs) were later introduced) to check for single nucleotide polymorphisms (SNPs).

An online tool (Gratz *et al.*, 2014) was used to design two gRNAs (**Table 2.2**): one *5'gRNA* (upstream of the target exons) and one *3'gRNA* (downstream). gRNAs were designed to bind as close to the target exons as possible, with 0 or less than 2 predicted off-target effects in other parts of the *Drosophila* genome (Gratz *et al.*, 2014).

Table 2.2 – gRNAs used in this thesis.

Construct name	Target Sequence
5'gRNA	GTGCCAAGGCAACGCGCTAC
3'gRNA	GGTCTGCAATATAATACCGT

Primers that bind either side of the targeted exons were designed to screen for deletions. In WT flies, these primers produce a ~3 kb PCR product. They produce a smaller (~2 kb) PCR product in the event of a deletion. Primers used in this thesis are listed in **Table 2.3**.

gRNAs were cloned into a pCFD4 plasmid (see **Section 2.2** for details) which drives expression of two gRNAs under U6 promoters (Port *et al.*, 2014). Constructs were sent to the CFF, where two strategies were attempted:

2.1.6.1. Strategy 1: transient expression of gRNAs in nos-Cas9 flies

At the CFF, constructs were injected into nos-Cas9 larvae (Port *et al.*, 2014) for transient expression. Resulting larvae were placed in quarantine until flies began to eclose. Males and virgin females were collected and placed in separate vials for crossing with SM5/Gla (**Figure 2.1**).

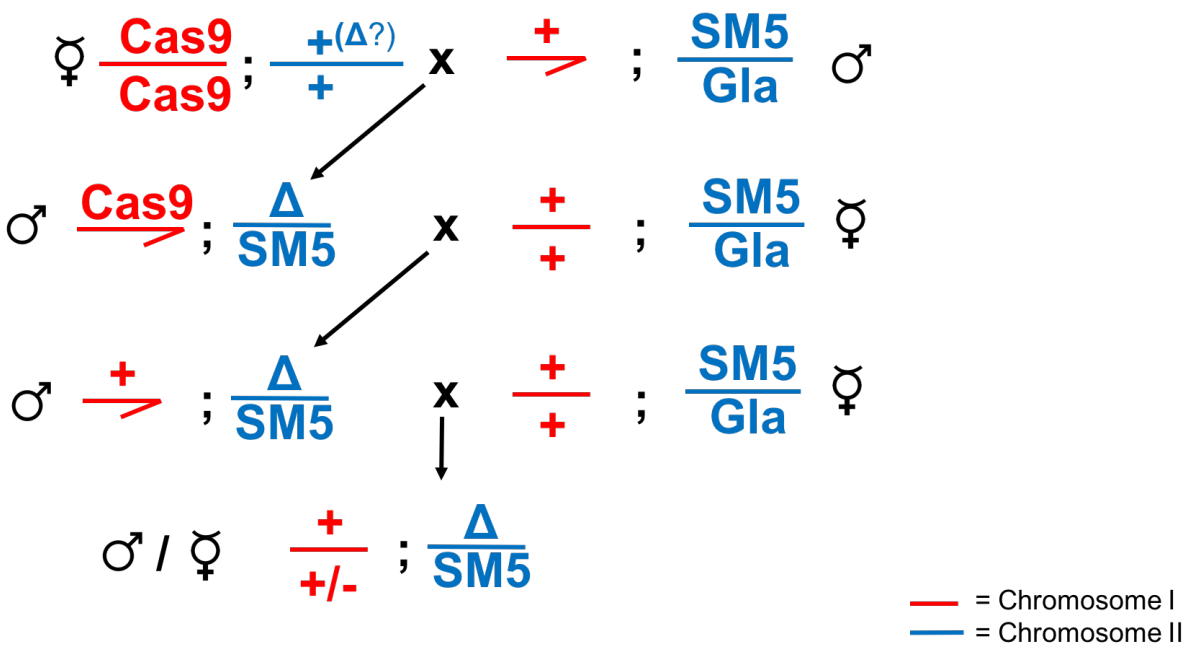


Figure 2.1 – Crossing scheme for transient gRNA expression strategy.

Crossing scheme for generating a balanced *cep104* deletion stock starting with a female founder with transiently expressed gRNAs. The crossing scheme for starting with a male founder was the same, but had one fewer crossing step when crossing with the SM5/Gla flies to eliminate the Cas9. Δ indicates a suspected deletion.

Once these crosses produced larvae, founders were lysed and screened for deletions by PCR, using the screening primers (see Section **2.2.1.2**). No PCR products corresponding to deletion events were observed.

2.1.6.2. Strategy 2: transgenic expression of gRNAs

At the CFF, constructs were injected into BL25709 flies for transgenic integration (**Section 2.1.4**) of gRNAs on Chromosome II. gRNAs were balanced over Curly O. These flies were put through quarantine before crossing with nos-Cas9 flies (**Figure 2.2**).

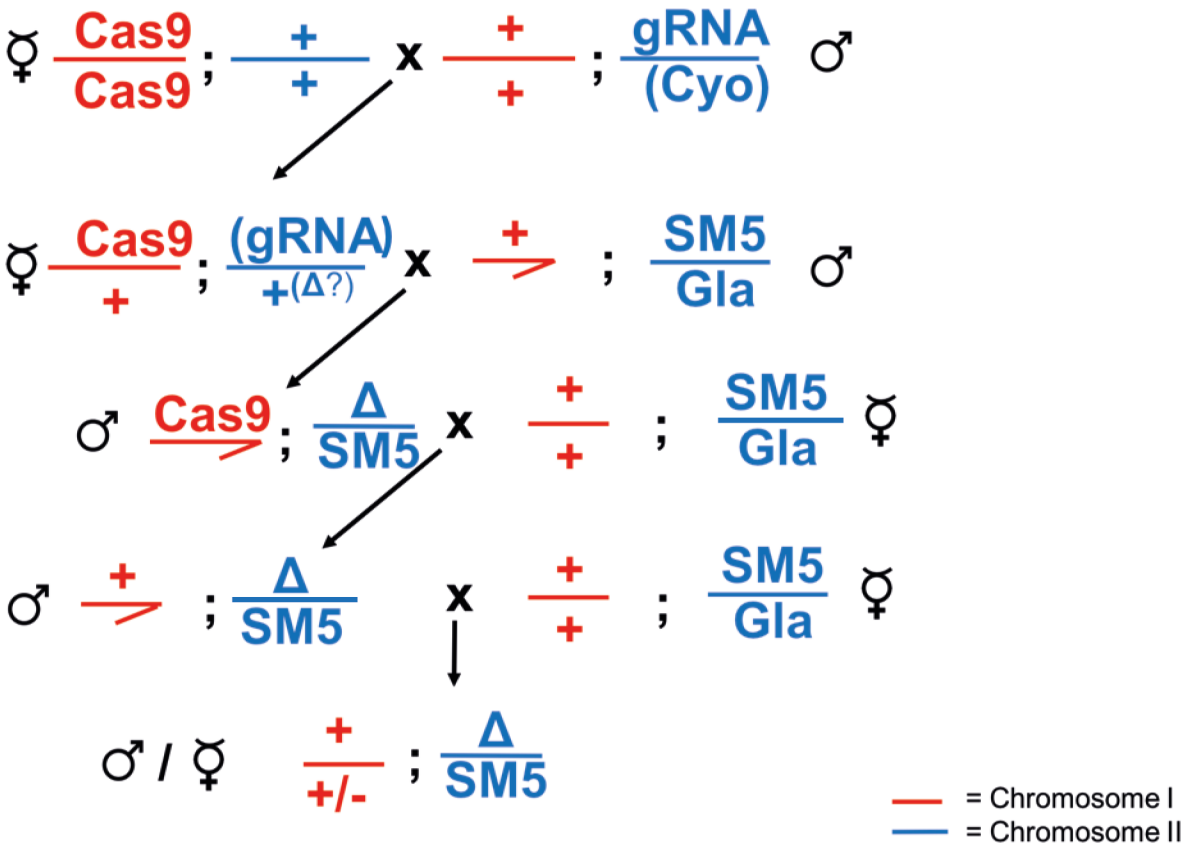


Figure 2.2 – Crossing scheme for transgenic gRNA integration strategy.

Crossing scheme for generating a balanced *cep104* deletion stock starting with Cas9 flies and flies stably expressing the gRNAs. These two fly strains were crossed and their female progeny were used as founders to set up crosses with SM5/Gla, since any deletion events would occur in these females. Δ indicates a suspected deletion.

Virgin female progeny were collected from this initial cross and placed in separate vials for crossing with SM5/Gla. The founder females were then lysed and screened as before. Deletion events were observed with an efficiency of ~1 in 50. Of three recovered deletions, one female founder transmitted the deletion to its male offspring, which were selected to set up the balanced deletion stock. This final stock was also screened by PCR to confirm the absence of Cas9 and gRNA expression.

2.1.7. Climbing assays

10 adult males (1-3 days old) of each genotype were placed into separate 10 mL plastic cylinders. Males were used (rather than a mixture of males and females) in order to control for any neurological differences or gene dosage effects between male and female flies. They were allowed to recover from CO_2 for 40 min at RT. Flies from both samples were simultaneously knocked to the bottom of their cylinders. Movies of their recovery were recorded (iPhone 6). 4 technical repeats were carried out. Movies were analysed to measure (1) the mean distance climbed in 5 s after knock down (2) the number of flies that climbed at least 9.6 cm in 10 s.

2.2. Molecular Biology

2.2.1. PCR reactions

All sequencing was performed by SourceBioscience. Sequencing analysis and primer design was carried out using MacVector (MacVector, Inc) and primers were synthesised by Invitrogen (Thermo Fisher). Primers used in this thesis are listed in **Table 2.3**.

2.2.1.1. DNA amplification via PCR

Primers were designed to introduce the gRNA sequences (**Table 2.2**) into a pCFD4 plasmid; PCR reactions using the pCFD4 plasmid as a template were performed with Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB M0531L) to generate gRNA-integrated pCFD4 fragments. Amplification of the promoter region ~2kb upstream

of the *cp110* coding sequence (to generate the eCP110-GFP fly line) was performed with Q5 Hot Start High-Fidelity 2x Master Mix (NEB M0494L), using 10 ng/µl w^{67} genomic DNA as template. Amplification of the pDONR containing the *cp110* cDNA sequence (Raff Lab ID #34), and addition of attB sites to CP110 and Cep97 Polo-binding mutant constructs was also performed with Q5 Hot Start High-Fidelity 2x Master Mix.

Resulting PCR products were run on a gel (1xTAE, 1% agarose) with SYBR Safe DNA Gel Stain (Thermo Fisher S33102) and imaged using ultraviolet light. Where necessary, bands were extracted from the gel and DNA purified using the Monarch DNA Gel Extraction Kit (NEB T1020L).

Table 2.3 – Primers used in this thesis.

Gene	Primer name	Sequence	Use in this thesis
CP110	78_Lau_CP110loop1del_sense	CAAGCGGGTGAACCGTGTGG AGCGTATAATAG	To delete conserved phosphorylation sites in CP110 sequence
CP110	79_Lau_CP110loop1del_anti	CTATTATACGCTCCACACGGT TCACCCGCTTG	To delete conserved phosphorylation sites in CP110 sequence
CP110	81_Lau_CP110loop2del_anti	GGTGAGGCGCTTTTTGGGACA ACTTGGACACTTC	To delete conserved phosphorylation sites in CP110 sequence
CP110	88_Lau_CP110loop1del_sense2	TGAACCGTGTGGAGCGTATAA TAGAACGG	To delete conserved phosphorylation sites in CP110 sequence

CP110	91_Lau_CP110_promoter_F	TGTACAAAAAAGCAGGCTTCG TTCCCTTTCGCTGTCAAG	To clone the promoter region ~2kb upstream of *cp110*
CP110	92_Lau_CP110_promoter_R	ATTGCCCACGTCGCATCCATT GGTGTTTTGCTACTGGG	To clone the promoter region ~2kb upstream of *cp110*
CP110	89_Lau_pDonr-CP110no_stop_R	GAAGCCTGCTTTTTTGTAC	To amplify vector #34 for assembly with the ~2kb promoter region of *cp110*
CP110	90_Lau_pDonr-CP110no_stop_F	ATGGATGCGACGTGGGCA	To amplify vector #34 for assembly with the ~2kb promoter region of *cp110*
CP110	72_LauCP110_Plk1s_m_attb1F1	GGGGACAAGTTTGTACAAAA AAGCAGGCTTCATGGATGCGA CGTGGGCAATG	To introduce attB sites at each end of the CP110-PBM construct
CP110	73_LauCP110_Plk1s_m_attb2R1	GGGGACCACTTTGTACAAGAA AGCTGGGTCATCCAATCGGCG ATGTTTGAGGAGTAG	To introduce attB sites at each end of the CP110-PBM construct
Cep97	74_ LauCep97_Plk1s_m_attb1F1	GGGGACAAGTTTGTACAAAA AAGCAGGCTTCATGAGTGGCG ACGAAAGTGG	To introduce attB sites at each end of the Cep97-PBM construct
Cep97	75_ LauCep97_Plk1s_m_attb2R1	GGGGACCACTTTGTACAAGAA AGCTGGGTCTGGATCTTTATC AAGATTTTCCGAGTCTTC	To introduce attB sites at each end of the Cep97-PBM construct
Cep104	Cep104_pCFD4_FW	TATATAGGAAAGATATCCGGG TGAACTTCGTGCCAAGGCAAC GCGCTACGTTTTAGAGCTAGA AATAGCAAG	To integrate the 5'gRNA into the pCFD4 plasmid
Cep104	Cep104_pCFD4_3'up_RV	ATTTTAACTTGCTATTTCTAGC TCTAAAACGGTCTGCAATATA ATACCGTCGACGTTAAATTGA AAATAGGTC	To integrate the 3'gRNA into the pCFD4 plasmid
Cep104	Cep104_pair2_FW	CCTTTAATCGTACTTACCTTTC TACTCCTA	To screen for Cep104 deletion events

Cep104	Cep104_pair2_RV	AGAACATGATGACAGAAGTC TAACG	To screen for Cep104 deletion events

2.2.1.2. Diagnostic PCRs for *cep104* deletion events

CO_2-anaesthetised flies were placed on ice in individual PCR tubes. 10 µl of microLYSIS-PLUS (Clent Life Science Ltd. 2MLP-1000) was added to each tube. Tubes were incubated in a PCR machine using the programme in **Table 2.4**. Microlysates were screened for deletions via PCR, using the Cep104 screening primers and OneTaq Quick-Load 2x Master Mix with Standard Buffer (NEB M0486L). Following DNA amplification, the resulting PCR products were run on a gel (1xTAE, 1% agarose) with SYBR Safe DNA Gel Stain and imaged using ultraviolet light.

Table 2.4 – The microlysis PCR programme.

Temperature	Incubation time
65°C	15 min
96°C	2 min
65 °C	4 min
96°C	1 min
65°C	1 min
96°C	30 s
20 °C	∞

2.2.1.3. Diagnostic PCRs to sequence the *cep104* deletion

Diagnostics were also performed via PCR amplification, clean-up and sequencing. Flies with the *cep104* deletion were lysed as described in **Section 2.2.1.2**. PCR reactions using these microlysates as a template were performed using Phusion High-Fidelity PCR Master

Mix with HF Buffer (NEB M0531L) and the two Cep104 screening primers. 0.5 µl of ExoProStar (Illustra US77702V) was added to 6 µl of PCR product and incubated at 37°C for 30 min, followed by 15 min at 80°C. This removed unincorporated primers/nucleotides. These cleaned-up samples were then sent for sequencing.

2.2.2. Construction of plasmids

All plasmids described below were isolated using Qiagen miniprep or midiprep kits (QiaPrep Spin MiniPrep Kit 27106, QiaPrep MidiPrep Kit 12243).

2.2.2.1. Construction of plasmids for generation of transgenic *D. melanogaster* lines

The promoter region of *cp110* and a pDONR containing the *cp110* cDNA sequence (without stop codon) were amplified via PCR and purified as described in **Section 2.2.1.1**. These were assembled using NEBuilder HiFi assembly (NEB E2621L) to produce a new plasmid where the endogenous *cp110* promoter sequence is incorporated upstream of the *cp110* cDNA sequence. This entry clone was recombined via the Gateway LR reaction into a destination vector (Raff Lab ID #313, kindly generated by S. Saurya) that lacks a promoter sequence and incorporates a C-terminal mGFP tag. These plasmids were sent to the CFF for injection into BL25709 flies.

CP110 and Cep97 Polo-binding mutant (PBM) sequences were designed by Ines Alvarez Rodrigo as previously described (Alvarez-Rodrigo *et al.*, 2019). These constructs were

synthesised by GENEWIZ Co. Ltd. (Suzhou, China) alongside a WT Cep97 sequence which would serve as a control. attB sites were added to these constructs via PCR (see **Section 2.2.1.1**) thus allowing them to be cloned into the pDONR/Zeo vector (Gateway, Thermo Fisher) via the Gateway BP reaction. The resulting entry clones were recombined via the Gateway LR reaction into a destination vector that incorporates a C-terminal mGFP tag and a ubiquitin promoter (Raff Lab ID #69). These plasmids were sent to the CFF for injection into BL25709 flies.

2.2.2.2. Construction of plasmids for CRISPR gRNA expression

As described in **Section 2.2.1.1**, the pCFD4 plasmid was used as a template in a PCR reaction to produce a fragment in which the gRNA sequences were integrated. A fresh stock of pCDF4 plasmid was digested using BbsI enzyme (NEB R0539S) in NEBuffer 2.1 (NEB B7202S) at 37°C for 2 h. The PCR fragment and digested vector were assembled using NEBuilder HiFi assembly (NEB E2621L). The resulting plasmids were sent to the CFF for injection into nos-Cas9 or BL25709 flies as required (described in **Section 2.1.6**).

2.2.2.3. Construction of plasmids for *in vitro* mRNA generation

To generate constructs where the two clusters of conserved phosphorylation sites in CP110 were deleted, the Quikchange II XL Site-directed Mutagenesis Kit (Agilent 200522) was used as per manufacturer's instructions, with plasmid #34 serving as the template. The resulting entry clone (pDONRZeo-CP110del) was recombined with the pRNA destination vector (Raff Lab ID #103) via Gateway LR reaction. Plasmid #34 was also recombined with #103 in order to generate a control (pRNA-CP110-GFP-CT-DEST) containing the WT *cp110* sequence.

2.2.2.3.1. Generation of mRNA for injection into *D. melanogaster* embryos

pRNA-CP110del-GFP-CT-DEST and pRNA-CP110-GFP-CT-DEST were taken forward for mRNA production. 4 µg of each vector was incubated in with 4 µl AscI enzyme (NEB R0558S) and 8 µl CutSmart buffer (NEB B7204S), in a total reaction volume of 80 µl, at 37°C for 2 h. AscI was then inactivated by incubating at 80°C for 15 min. Linearised DNA was precipitated overnight at -20°C by adding 4 µl of 0.5 M EDTA, 8 µl of 3 M Na-acetate, and 180 µl 100% ethanol. Samples were spun at 14000 rpm, 4°C for 1 hour, to pellet DNA. Pellets were washed 2x with 70% RNAse-free ethanol before air-drying and resuspension of DNA in 10 µl DEPC-treated water. The mMESSAGE mMACHINE T3 Transcription Kit (Thermo Fisher AM1348) was then used as per manufacturer's instructions to synthesise RNA, using 2 µg DNA as a template and with incubation at 37°C for 2 h. Finally, the RNeasy MinElute Cleanup Kit (Qiagen 74204) was used to purify the RNA. The resulting RNA was diluted to produce 2 µg/µl aliquots for use in experiments and stored at -80°C.

Table 2.5 – Plasmids used in this thesis.

Plasmid	**Use in this thesis**	**Source**
pDONR/Zeo	Production of entry clones.	Gateway, Thermo Fisher
Ubq-mGFP-CT-DEST	Destination vector for generation of tagged CP110/Cep97 Polo-binding mutant (CP110/Cep97-PBM-GFP) and Cep97-WT-GFP fly lines.	Raff Lab (#69)
Ubq-Cep97-WT-mGFP	Final plasmid sent to the CFF to generate Ubq-Cep97-WT-GFP flies.	This thesis
Ubq-CP110-PBM-mGFP	Final plasmid sent to the CFF to generate Ubq-CP110-PBM-GFP flies.	This thesis
Ubq-Cep97-PBM-mGFP	Final plasmid sent to the CFF to generate Ubq-Cep97-PBM-GFP flies.	This thesis

pDONRZeo-CP110 cDNA - no stop (RE33938)	Amplified and combined (using NEBuilder HiFi assembly) with the *cp110* promoter region to produce an entry clone where *cp110* expression will be driven by its endogenous promoter. This plasmid was also used as a template for site-directed mutagenesis to generate pDONRZeo-CP110del, an entry vector in which the two *cp110* regions coding for conserved phosphorylation sites have been deleted.	Raff Lab (#34)
pNoP-mGFP-CT-DEST	Destination vector for generation of eCP110-GFP flies.	Raff Lab (#313), kindly generated by Saroj Saurya via Q5 mutagenesis of plasmid #69
eCP110-mGFP	Final plasmid sent to the CFF to generate eCP110-GFP flies.	This thesis
pCFD4	Used as a vector to drive expression of the two gRNAs targeting the Cep104 C-terminus for deletion.	(Port *et al.*, 2014)
pCFD4-Cep104KO_gRNAs	pCFD4 plasmid containing the gRNAs targeting the Cep104 C-terminus for deletion. These plasmids were sent to the CFF for injection into nos-Cas9 flies (Strategy 1) or BL25709 flies (Strategy 2).	This thesis
pDONRZeo-CP110del	Entry vector where the two *cp110* regions coding for conserved phosphorylation sites have been deleted.	This thesis
pRNA-GFP-CT-DEST	Destination vector with pRNA promoter for *in vitro* mRNA production.	Raff Lab (#103)
pRNA-CP110del-GFP-CT-DEST	Final plasmid used to generate mRNA of *cp110,* where the two *cp110* regions coding for conserved phosphorylation sites have been deleted. This mRNA was used for injection into embryos.	This thesis
pRNA-CP110-GFP-CT-DEST	Final plasmid used to generate WT *cp110* mRNA for injection into embryos.	This thesis

2.3. Biochemistry and Cell Biology

2.3.1. Fixation of *D. melanogaster* embryos for western blotting

Embryos were collected on fruit juice plates for 3 h. They were dechorionated by incubation in 60% bleach for 2 min, before being washed (in PBS + 0.0005% Tween-20) and transferred into a glass bottle containing 2 mL 100% methanol and 2 mL 100% heptane. The bottle was shaken for 10 s before allowing embryos to sink to the bottom. Excess liquid and embryos which did not sink were discarded. The remaining embryos were washed 3x with 100% methanol before storage in methanol at 4°C for at least 24 h.

2.3.2. Preparation of *D. melanogaster* brains for western blotting

Third-instar larvae were placed in 1x PBS and brains were dissected out using tweezers. Six brains per sample were transferred into 18 µl of 2x SDS loading dye. Samples were boiled at 100°C for 10 min. A pipette was used to measure the resultant sample volume; an equivalent volume of distilled water was then added and samples were stored at -20 °C.

2.3.3. SDS-PAGE and western blotting

Embryos fixed in 100% methanol were washed 3x for 15 min (in PBS + 0.2% Tween-20) at RT, while rotating. They were transferred to a watch glass to enable selection of embryos in the syncytial blastoderm stage. The required number of embryos were counted

into a pipette with 20 µl PBS + 0.2% Tween-20. This was mixed with 20 µl of 2x SDS loading dye and boiled at 100 °C for 10 min. Samples were stored at -20 °C overnight.

For western blotting, samples were defrosted and loaded onto 3-8% NuPAGE Tris-Acetate Protein Gels (Thermo Fisher EA03785BOX) in NuPAGE Tris-Acetate SDS running buffer (Thermo Fisher LA0041). Samples were run on the gel at 150 V for 75 min, in an XCell SureLock Mini-Cell (Thermo Fisher). Samples were then transferred onto a nitrocellulose membrane (Bio-Rad 1620112, 0.2 µm) via semi-wet transfer using in-house transfer buffer (25 mM Tris pH 7.4, 192 mM glycine, 25% methanol, 1% SDS). A Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad) was used for this purpose, running at 100 V for 90 min.

Nitrocellulose membranes were incubated with Ponceau (Sigma-Aldrich P7170-1L) for 30 s, rinsed with distilled water and cut to appropriate sizes for antibody incubation. The membrane sections were blocked for 1 h at RT (with 1x PBS, 4% milk powder, 0.1% Tween-20) then incubated for 1 h at RT with primary antibody solution. Primary antibodies used in this thesis are included in **Table 2.7**. Membranes were washed 3x for 15 min (with 1x TBS, 0.1% Tween-20) and incubated for 45 min at RT with HRP conjugated secondary antibody solution. Secondary antibodies used in this thesis are included in **Table 2.8**. Membranes were washed 3x for 15 min (with 1x TBS, 0.1% Tween-20) then incubated for 1 min in the required dilution of SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher 34095). Finally, film was exposed to the membranes and then developed.

Where necessary, membranes were stripped with Restore Western Blot Stripping Buffer (Thermo Fisher 21059) to allow for re-probing. These membranes were washed in 1x PBS, then incubated in stripping buffer for 15 min at RT.

2.3.3.1. Quantification from western blot to assess global CP110/Cep97 levels under different Plk4 conditions

Four technical repeats of each blot were carried out. Films were scanned to produce TIFF files (600 pixels per inch). These were opened in ImageJ (NIH) and the α-Cep97 or α-CP110 signal in all lanes was selected by using the rectangle tool to create a region of interest (ROI). The 'gels' ('select first lane') option was then selected from the 'analyse' menu. 'Plot lanes' from the 'gels' selection was used to plot the detected signals for each separate lane; area under the curve for each lane was then calculated and exported to Microsoft Excel. The same size ROI was used to repeat the process for the α-actin signal (which served as the loading control) in each lane. The α-Cep97 or α-CP110 signals were normalized to their respective loading controls, then to the mean value of all WT α-Cep97 or α-CP110 signals respectively.

2.3.4. Fixation of *D. melanogaster* embryos for immunoprecipitation

ePlk4-GFP embryos in a Plk4 mutant background and WT embryos were collected on fruit juice plates for 3 h. They were dechorionated by incubation in 60% bleach for 2 min, then washed (in PBS + 0.0005% Tween-20) and flash frozen in liquid nitrogen for storage at -80°C. The mass of each sample was recorded.

2.3.5. Immunoprecipitation

2.3.5.1. Coupling the α-GFP antibody to Dynabeads

The Dynabead Coupling Kit (Thermo Fisher 14311D) was used as per manufacturer's instructions. Briefly, 6.2 mg magnetic Dynabeads were weighed out and washed with 1 mL of solution C1, with mixing via flicking and vortexing. Beads were immobilised on a magnetic rack and the supernatant was removed. 264 µl of C1 was added to the beads, with flicking to mix. 46.5 µl of α-GFP antibody (#118) was then added to the beads. Again, mixing was carried out via flicking and gentle vortexing before 310 µl of solution C2 was added. This mixture was left rotating at 37°C overnight. The next day, beads were immobilised as before and the supernatant was removed. Beads were washed in 800 µl HB wash buffer with vortexing. This supernatant was removed, and 800 µl LB wash buffer was added, with vortexing to mix. This supernatant was removed and beads were washed 2x in 800 µl SB buffer. Finally, beads were washed in 800 µl SB buffer for 15 min, rotating at RT. The supernatant was removed and beads were resuspended in 620 µl SB and stored overnight at 4°C.

2.3.5.2. Preparing the embryo extract

Buffers were prepared on the day of the IP (**Table 2.6**). Samples of ePlk4-GFP or WT embryos prepared as described in **Section 2.3.4** were combined in 1.5 mL Eppendorf tubes at 4°C to form a total mass of 250 mg. One volume (250 µl) of homogenising buffer was added to each tube with homogenisation. DTT was added to each tube at a concentration of 1 in 100, before tubes were centrifuged until pelleted at 13000 rpm, 4°C. 90 µl of supernatant from each tube was transferred to separate tubes for dilution with 45 µl

homogenising buffer. A 10µl aliquot of this embryo extract was diluted in 10 µl 2x SDS loading dye, boiled for 10 min and stored at -20°C as the '**input**' fraction, leaving 125 µl extract to use in the IP.

2.3.5.3. Starting the IP

Antibody-coupled beads were divided equally into two 0.5 mL LoBind (Eppendorf) tubes and immobilised as before. The supernatant was removed from the beads as before and replaced with 1x PBS + 0.1% BSA before beads were left rotating for 15 min at RT. The supernatant was removed and replaced with 125 µl embryo extract. Beads were incubated in embryo extract overnight, rotating at 4°C.

The next day, beads were transferred into fresh 1.5 mL LoBind tubes. Beads were then immobilised and the supernatant was removed. 10 µl of this supernatant was diluted in 10 µl 2x SDS loading dye, boiled for 10 min and stored at -20°C as the '**unbound**' fraction. The rest of the supernatant was discarded. Beads went through 3x quick 130 µl PBST (1x PBS + 0.1% Triton) washes at RT. Beads were then washed 5x with 130 µl PBST for 10 min per wash, rotating at 4°C. Beads were resuspended in 50 µl 2x SDS loading dye, mixed with vortexing, then boiled for 10 min. This supernatant was removed and stored at -20°C as the '**bound**' fraction.

Table 2.6 – Buffers used for immunoprecipitation.

Buffer	Components	Stock Concentration	Volume (µl)	Final Concentration
Homogenising buffer	HEPES pH 7.6	1 M	75	50 mM
	$MgCl_2$	1 M	1.5	1 mM
	EGTA	0.5 M	3	1 mM
	Phosphatase inhibitor II (Sigma P5726)	100 x	15	1x
	Phosphatase inhibitor III (Sigma P0044)	100 x	15	1x
	Protease inhibitor cocktail	100 x	15	1x
	dH_2O		1374	
Protease inhibitor cocktail	Benzamidine (Sigma B6506)	300 mM	165	100 mM
	Phenanthroline (Sigma P9375)	200 mM	125	50 mM
	Aprotinin (Sigma A6103)	10 mg/µl	50	1 mg/µl
	Leupeptin (Sigma L2884)	50 mg/µl	10	1 mg/µl
	Pepstatin (Sigma P5318)	1 mg/µl	150	0.3 mg/µl
PBST 0.1%	PBS	1 x	50000	1x
	Triton	1 M	50	1 mM

2.3.5.4. Western blot

'**Input**', '**unbound**', and '**bound**' samples were thawed and spun down. Aliquots from the '**input**' and '**unbound**' samples were diluted 3/10 using 1x SDS loading dye. The '**bound**' sample did not require dilution. Western blotting was performed as described in **Section 2.3.3**, with 10 µl of sample loaded per lane. The membrane was probed for Cnn and GFP before being stripped and re-probed with α-CP110, then α-Cep97 primary antibodies.

2.3.6. Fixation and staining of *D. melanogaster* wing discs for imaging

Wing discs were prepared and stained as previously described (Gartenmann *et al.*, 2017). Wing discs were dissected in 1x PBS from third-instar larvae. Following a 20 min fixation in PBS with 3.7% formaldehyde, the wing discs were placed in an 8 µl drop of PBS on a high precision #1.5H (170 ± 5 µm thickness) coverslip (Marienfeld Superior) which had been washed in 100% ethanol and air-dried. The coverslip was lowered onto a #1 thickness cover slide (VWR International) which had been washed in silica suspension and air-dried. The coverslip and slide were then compressed together between two pieces of Whatman filter paper for 10 s to form a thin layer of cells, before being plunged into liquid nitrogen for 10 min. A razor blade was used to separate the coverslip from the slide. The sample remained attached to the high precision coverslip, which was placed in 100% ethanol at -20°C for 15 min. The sample was then placed in PBS + 0.1% TritonX for 10 min before being washed in 1x PBS for 5 min. Coverslips were dried so that 20 µl primary antibody solution could be added to each sample. The samples were then incubated overnight at 4°C in a humid box. The next day, coverslips were washed 3x 5 min in 1x PBS, then incubated with secondary antibody solution in a humid box for 2 h at RT. After 2x 5 min washes in 1x PBS, the coverslips were mounted in 8 µl Mowiol 4-88 and stored at -20 °C. Slides were imaged as described in **Section 2.4.2.1**.

2.3.7. Antibody purification

Purification of the Cep104 antibody (generated by Metta Pratt; Raff Lab, University of Oxford) was carried out with the assistance of Saroj Saurya as described previously

(Huang and Raff, 1999). Briefly, a column was set up containing Affigel-15 beads (Bio-Rad Laboratories) covalently coupled to MBP fusion proteins. Sera from the final rabbit bleed was passed 3-4x through this column until no MBP-specific antibodies remained. A second column containing Affigel-15 beads coupled to MBP-Cep104 was used to purify Cep104-specific antibodies from the supernatant. 1x PBS + 0.5 M KCl was used to wash the column until no more protein came off. Cep104 antibody was then eluted with 0.1 M glycine (pH 2.3) and neutralised with 1 M Tris (pH 9.5). A Bradford Assay (Bio-Rad Laboratories) was used to identify the two peak eluted fractions; these were pooled and an equivalent volume of 100% glycerol was added. This sample was stored at -20°C. Side fractions were flash frozen in liquid nitrogen and stored at -80°C.

2.3.8. Primary antibodies used in this study

Table 2.7 lists the primary antibodies used for western blotting (WB), immunofluorescence (IF), and immunoprecipitation (IP).

Table 2.7 – Primary antibodies used in this thesis.

Lab ID number	Protein	Species raised in	Use in this thesis	Dilution	Source and relevant publications
#15	CP110	Rabbit	WB	1:500	Eurogentec (animal #SK4824) (Franz *et al.*, 2013)
#250	Cep97	Rabbit	WB	1:500	Generated by Jeroen Dobbelaere (Dobbelaere *et al.*, 2020)
#186	GFP	Mouse	WB	1:500	Roche (11814460001)
#166	GFP	Llama	IF (Alexa Fluor 488)	1:500	Chromotek (gba488)

#118	GFP	Rabbit	IP	N/A	Eurogentec (animal #SK495) (Huang and Raff, 1999)
#37	Cnn	Rabbit	WB	1:500	Eurogentec (animal #SK3516)
#221	Actin	Mouse	WB	1:500	Sigma (A3853)
#248	Cep104 (sera)	Rabbit	WB	1:500	Eurogentec (animal #1251)
#199	Asl	Guinea pig	IF	1:1000	Eurogentec (animal #SKC123) (Roque *et al.*, 2012)

2.3.9. Secondary antibodies used in this study

Table 2.8 lists the secondary antibodies used for western blotting (WB), immunofluorescence (IF), and immunoprecipitation (IP).

Table 2.8 – Secondary antibodies used in this thesis.

Lab ID number	Reacts with	Species raised in	Use in this thesis	Dilution	Source and relevant publications
#13	Mouse	Sheep	WB (HRP)	1:3000	(na931V)
#39	Rabbit	Swine	WB (HRP)	1:3000	Agilent (P039901-2)
#46	Guinea pig	Goat	IF (Alexa Fluor 594)	1:500	Invitrogen (A11076)

2.3.10. Injection experiments in living embryos

2.3.10.1. Injection of mRNA into living embryos

Embryos were prepared as described in **Section 2.1.2**. mRNA (produced *in vitro,* see **Section 2.2.2.3**) was pipetted into glass needles for injection, which was carried out on a Nikon TMS inverted microscope mounted with a manual injection apparatus (Micro Instruments LTD). Embryos were then incubated at 25°C for 120 min before imaging, so that the syncytial stage of development would be captured.

2.3.10.2. Drug injection into living embryos

Colchicine in DMSO was diluted in Schneider's medium to 100 µg/mL and pipetted into glass needles for injection. Embryos were prepared as described in **Section 2.1.2**. Injection was carried out on a spinning disc microscope (see below); embryos in S-phase of nuclear cycle 11 were identified and injected just after nuclear envelope breakdown (NEB), with imaging commencing immediately.

2.3.10.3. RNA interference in living embryos

Double-stranded RNAs (dsRNAs) against *Drosophila* mitotic cyclins A, B and B3 were generated by Mohammad Mofatteh (Raff Lab, University of Oxford) as described previously (Aydogan *et al.*, 2020, McCleland and O'Farrell, 2008). Primers used are listed in **Table 2.9**.

Table 2.9 – Primers used for amplification of cyclin genes.

Primer	Sequence
Primer to amplify cyclin A: Forward	TACAATTGCAAGCAGAAAGAGG
Primer to amplify cyclin A: Reverse	GGAGTTCACAGAACATTTACGG
Primer to amplify cyclin B: Forward	TAATTCTTGGAGAGACAACGCC
Primer to amplify cyclin B: Reverse	CTACAACAAGTACCAGGGCAGC
Primer to amplify cyclin B3: Forward	GAACGAAATATTGGCGTAATCC
Primer to amplify cyclin B3: Reverse	CGCCATGGATATATTCAACTACC

A further primer (TAATACGACTCACTATAGGGAGACCACGGGCGGGT) was also used to add the linker sequence GGGCGGGT to the 5' end of each primer. This would enable a T7 promoter to be added in second round of PCR.

Embryos were prepared as described in **Section 2.1.2**. They were then injected with 0.6-0.8 mg/mL dsRNA using glass needles on a Nikon TMS inverted microscope mounted with a manual injection apparatus (Micro Instruments LTD). After injection, embryos were incubated at 25°C until they were 105 min old (nuclear cycle 12) for imaging on a spinning disc confocal microscope.

2.4. Microscopy

2.4.1. Live imaging of *D. melanogaster* embryos

Embryos were prepared as described in **Section 2.1.2**.

2.4.1.1. Spinning disc confocal microscopy

Images for Figure 4.10, Figure 4.11, Figure 4.14, Figure 4.15, Figure 6.2, Figure 6.4 and Figure 6.5 were acquired at RT on an Andor Dragonfly 505 spinning disc system (40 µm pinholes) mounted on a Leica DMi8 stand. An HC PL APO 63x/1.40 oil immersion objective was used. 488 nm solid state diode and 561 nm OBIS LS solid state smart lasers were used to excite GFP/NeonGreen and RFP, respectively. Images were acquired as a stack of 17x 0.5 µm slices using an Andor iXon Ultra 888 EMCCD camera.

Images for **Figure 3.16**, **Figure 4.12**, and for the colchicine injection experiment (see **Section 2.3.10.2**), were acquired on a Perkin Elmer ERS spinning disc mounted on a Zeiss Axiovert 200M microscope with injection apparatus. A Plan-Apochromat 63x/1.4 NA oil immersion objective was used. A 488 nm Argon gas laser was used to excite GFP. Images were acquired as a stack of 13x 0.5 µm slices with 1x1 binning using an Orca ER CCD camera (Hamamatsu Photonics, Japan). Focus adjustments sometimes needed to be made during imaging, to account for movement of centrosomes in the *z* plane. Since only GFP was being imaged, the UltraVIEW ERS 'fast sequential' mode was employed.

All other spinning disc images in this thesis were acquired using a Yokogawa CSU-X1 spinning disc mounted on a Nikon Eclipse TE200-E microscope. A Plan-Apochromat 60x/1.42 NA oil immersion objective was used. 488 nm solid state and 561 nm lasers were used to excite GFP and RFP/mCherry, respectively. Images were acquired as a stack of 17x 0.5 µm slices using an EMCCD Andor iXon+ camera.

525/50 nm and 600/50 nm band pass emission filters were used for 488 nm and 561 nm excitation, respectively. For simultaneous imaging of GFP and RFP/mCherry, 'emission discrimination' mode was employed.

2.4.1.2. Zeiss 880 LSM with Airyscan detector

Embryos were imaged using an inverted Zeiss 880 Laser Scanning Microscope (LSM) fitted with Airyscan detector and a Plan-Apochromat 63x/1.4-NA oil lens. The immersion oil had a refractive index of 1.518; this closely matches the refractive index of the embryos, thus reducing spherical aberration. 488 nm argon and 561 nm diode lasers were used for excitation of GFP and mCherry, respectively. Channel alignment was ensured by switching excitation of each wavelength per line. Images were acquired as a stack of 5x 0.2 µm slices (zoom value = 24.41 pixels per µm for GFP-Cep97 distancing experiments; zoom value = 24.61 pixels per µm for assessing CP110/Cep97 levels on mother and daughter centrioles), with manual focus adjustment between acquisitions to maintain centrioles within the plane of imaging.

2.4.1.3. OMX V3 Blaze with 3D-structured illumination (3D-SIM)

Embryos were imaged in nuclear cycle 14 at 21°C using a DeltaVision OMX V3 Blaze Microscope (GE Healthcare) with a 60x/1.42-NA oil UPlanSApo objective (Olympus). The immersion oil had a refractive index of 1.514; this closely matches the refractive index of the embryos, thus reducing spherical aberration. 488 nm and 593 nm diode lasers were used to excite GFP and mCherry respectively, with Edge 5.5 sCMOS cameras (PCO) employed to capture images in each channel. 525/50 nm and 615/24 nm band pass

emission filters were used for 488 nm and 593 nm excitation, respectively. SIM images were acquired as a stack of 6x 0.125 μm slices in 5 phases, with 3 angles per slice.

2.4.2. Imaging of fixed and stained *D. melanogaster* wing discs

2.4.2.1. OMX V2 with 3D-SIM/single-molecule localisation microscopy (SMLM)

Fixed and stained wing discs were imaged on a modified DeltaVision OMX V2 microscope (Applied Precision/GE Healthcare) in 3D-SIM mode, with a 100x/1.4 NA oil objective (UPlanSApo, Olympus). The immersion oil had a refractive index of 1.514; this closely matches the refractive index of the embryos, thus reducing spherical aberration. 488 nm diode (Sapphire 488-200, Coherent) and 592 nm fibre (F-04306-01, MPB Communications) lasers were used, with EMCCD cameras running at 10 MHz (Evolve 512 Delta C, Photometrics) employed to capture images in each channel. 525/50 nm and 629/53 nm band pass emission filters were used for 488 nm and 592 nm excitation, respectively. Images were acquired as a stack of 15x 0.125 μm slices in 5 phases, with 3 angles per slice.

2.4.3. Transmission electron microscopy (TEM)

All TEM was carried out by Saroj Saurya. Wing discs from third-instar larvae were prepared essentially as previously described (Stevens *et al.*, 2010b) with slight adjustments. Briefly, dissected wing discs were fixed for 1-2 h at RT in 2.5% glutaraldehyde, 4% paraformaldehyde in 0.1 M PIPES buffer (pH 7.2). They were then

moved to 4°C overnight. Samples were washed 2x in 0.1 M PIPES, 1x in 0.1 M PIPES + 50 mM glycine (to quench free aldehydes) then 1x in 0.1 M PIPES. They were post-fixed for 2 h at RT in 1% OsO_4 then washed thoroughly with distilled water before being stained overnight at 4°C with 0.5% uranyl acetate. Samples were washed in distilled water then dehydrated via an ethanol series and embedded in Agar100 (Agar Scientific). Polymerisation was carried out at 50°C for 24-42 h. Semi-thin serial sections (100nm) were generated using a Leica EM UC7 ultramicrotome (Leica Microscystems, Austria). Sections were stained in lead citrate. Images were acquired using a TECNAI T12 transmission microscope (FEI, Netherlands) at 13,000x magnification. Centriole MT doublet length within the electron-dense area was measured in ImageJ using the 'line' tool.

2.5. Image processing and analysis

2.5.1. Processing spinning disc images

2.5.1.1. Centriole tracking with TrackMate to generate fluorescence intensity curves

Images acquired on the spinning disc confocal microscopes were processed using ImageJ as described previously (Aydogan *et al.*, 2018). Maximum intensity projections were processed in preparation for analysis by performing a bleach correction (exponential fit algorithm) and then using the 'subtract background function' to remove background (rolling ball radius = 10 pixels for Andor images, 5 pixels for Perkin Elmer images). Tracking of centriole pairs was performed using the TrackMate plug-in (Tinevez *et al.*, 2017). Track spot size diameter was set at 1.1 µm while quality was adjusted each time to select the centrosomes within the images. This generated an output of the mean

fluorescence intensity of each centrosome at each time point. The output was imported into GraphPad Prism 7 or 8 (GraphPad Software Inc.) to plot the change in mean centrosome intensity over time for individual embryos.

Where Plk4-GFP was imaged with Ubq-RFP-Cep97, and where Ubq-CP110-GFP was imaged with Ubq-RFP-Cep97, the RFP fluorescence intensity was lower than GFP due to differences in the properties of these tags. The highest fluorescence intensity measurements in cycle 12 were therefore scaled to 1 arbitrary unit (A.U.), with this scaling factor then being applied to cycles 11 and 13. This allowed comparison of the oscillations' phases.

In the Cyclin A-B-B3 dsRNA injection experiment (**Figure 6.10**), centrioles move away from the cortex. This caused a reduction in fluorescence signal intensity (y_{raw}). Felix Zhou (Lu Lab, University of Oxford) therefore used an asymmetric least-squares filter (Eilers and Boelens, 2005), with parameters $\lambda = 1\text{x}10^6$, p = 0.1; 10 iterations, to estimate the baseline trend of the signal ($y_{baseline}$). He found the baseline was linear ($y_{baseline} = m\,t + c$, where m =gradient offset, c = constant offset; values of c and m were extracted from least-mean squares fitting). The final de-trended signal was therefore found by subtraction of the baseline's gradient component from the raw signal: $y_{detrend} = y_{raw} - m\,t$

2.5.1.2. Non-linear regression models

Having plotted the fluorescence intensity data over time from individual embryos, the 'Nonlinear Regression' function in GraphPad Prism 7 or 8 was used to fit Ubq-GFP-Cep97 and Ubq-CP110-GFP data against four functions:

(1) Lorentzian

(2) Gaussian

(3) Linear increase, plateau, linear decrease

(4) Linear increase, linear decrease

Fitting was performed on data from individual embryos by considering the mean *y* value of each time point. R^2, absolute sum-of square values and *Runs* tests were used to assess goodness-of-fit. Comparisons between fits were performed for each embryo using the sum-of-squares *F* test; the simpler model was selected unless $p<0.05$. M. Aydogan found that *(1)* fitted the data best, so all CP110/Cep97 fluorescence intensity data in this thesis were regressed with this model.

Using the same approach as described above, Ubq-Cep104-GFP data was fitted against three functions:

(1) Linear increase, linear decrease

(2) Exponential increase, linear decrease

(3) Sum of two Gaussians

(1) fitted the data best, so all Cep104 fluorescence intensity data in this thesis were regressed with this model. The following functions are described in GraphPad Prism: *Lorentzian; Gaussian; Sum of Two Gaussians.* The following functions are in-house, as described in Aydogan *et al.* 2018 and Alvarez Rodrigo *et al.* 2019: *Linear increase, plateau, linear decrease; Linear increase, linear decrease.* The *Exponential increase, linear decrease* function is also an in-house equation, defined as follows:

$$Y_1 = intercept_1 + \exp(k * X)$$

$$Y_{X0} = intercept_1 + \exp(k * X0)$$

$$Y_2 = Y_{X0} + slope_1 * (\mathrm{X} - \mathrm{X0})$$

$$Y = \mathrm{IF}(\mathrm{X} < \mathrm{X0}, \mathrm{Y1}, \mathrm{Y2})$$

The following constraints were applied to the above equations: *X0* must be greater than 0, and $slope_1$ must be greater than 0.

Sas-6-GFP/Sas-6-mNeonGreen data were fitted using either *(1) Linear increase*, or *(2) Linear increase followed by plateau* functions, as described in Aydogan *et al.* 2018.

Once regression models were fitted to each individual embryo, parameters such as *growth rate* and *time to oscillation peak* could be extracted and compared between datasets. Average curves were also created by plotting the mean parameters from a population of embryos.

2.5.1.3. Correlation analysis

To test whether the peak of the CP110/Cep97 oscillation correlated with S-phase length, GraphPad Prism 7 or 8 was used to carry out a Pearson's correlation analysis between the *time to oscillation peak* and *S-phase length* data and to fit a linear regression where a significant correlation was detected. Statistically significant correlations were defined by $p<0.05$. For all graphs, *p*-values, R^2 values, and Pearson's correlation coefficient (r) are indicated.

2.5.1.4. Correcting the CP110/Cep97 oscillation centre for S-phase length

S-phase length was measured as the time period from centrosome separation (CS) to nuclear envelope breakdown (NEB). To account for variation in S-phase length between datasets, mean S-phase lengths were calculated for each dataset. A correction factor was determined by dividing mean S-phase length of the control dataset by mean S-phase length of the experimental dataset. The *time to oscillation centre* values from the experimental dataset were then multiplied by this correction factor.

2.5.1.5. Quantification of Ubq-CP110-GFP/Ubq-GFP-Cep97 in *cep97/cp110* null embryos

Images of cycle 14 embryos were acquired using spinning disk microscopy. Maximum intensity projections were generated and analysed in ImageJ. In each embryo, Asl-mCherry was used as a marker to select the ten brightest centriole pairs 20 min into S-phase. A 5x5 oval ROI was placed around selected centriole pairs, and measurements of mean GFP intensity were taken. Background correction was achieved by subtracting a

background mean intensity measurement for each image. Bar charts of background-subtracted mean GFP intensity measurements for each embryo were produced in GraphPad Prism 7 or 8.

2.5.2. Processing Zeiss 880 LSM Airyscan images

2.5.2.1. Airyscan processing

Images acquired using Airyscan super-resolution microscopy were Airy-processed in 3D with a processing strength of Auto (~6) or 6.5 using the microscope's software (ZEN black, Zeiss International).

2.5.2.2. Quantification of Ubq-CP110-GFP/Ubq-GFP-Cep97 on mother/daughter centrioles

Quantification of Ubq-CP110-GFP or Ubq-GFP-Cep97 fluorescence intensity on mother/daughter centrioles was performed in ImageJ. An 8x8 oval ROI was used to select individual centrioles, from which mean intensity measurements in the GFP channel were taken. Mother centrioles were identified using the Asl-mCherry marker (Novak *et al.*, 2014). Background correction was achieved by subtracting a background mean intensity measurement for each image. Graphs were produced in GraphPad Prism 7 or 8.

2.5.2.3. Analysis of centriole MT growth dynamics via machine learning

Images of Ubq-GFP-Cep97 over the course of cycle 12 were analysed using the GFP-Cep97 distancing algorithm (see **Chapter 3** and **Appendix A1** for more details on

development of this algorithm). The resulting MT distancing data were fitted using either *(1) Linear increase*, or *(2) Linear increase followed by plateau* as described for modelling cartwheel growth in Aydogan *et al.* 2018.

2.5.3. Processing OMX V2 and OMX V3 Blaze images

2.5.3.1. Reconstruction and alignment

Reconstruction of the images was performed using softWoRx 6.1 (GE Healthcare). Alignment of the red and green channels was achieved by imaging 0.2 μm diameter TetraSpeck beads (Thermo Fisher Scientific) which were used to calibrate for alignment on the OMX Editor software.

2.5.3.2. Measurements of CP110/Cep97 ring radius

Alan Wainman (Raff Lab, University of Oxford) assessed the orientation of centrioles and determined the radius of Ubq-CP110-GFP, Ubq-GFP-Cep97 and Ubq-Cep97-GFP in fixed and stained *Drosophila* wing disc centrioles. Wing discs were prepared as previously described (Gartenmann *et al.*, 2017). Imaging was performed as per **Section 2.4.2.1**; reconstruction and alignment was performed as per **Section 2.5.3.1**. A Python code (SIM-COOL, https://github.com/RaffLab/SIM-centriole-ring-measurement) was used to assess the orientation of centrioles. Rings were fitted to the Asl (red) and CP110/Cep97 (green) signals using an elliptical Gaussian profile. This obtained fit parameters for the major and minor axis. Centrioles were considered well-oriented if the eccentricity (major:minor axis ratio) of both the Asl and CP110/Cep97 rings were <1.2. If these conditions were satisfied,

the radial localisation of CP110/Cep97 was measured by the script.

2.5.4. Image presentation for figures

For presentation purposes, contrast and brightness were sometimes adjusted in ImageJ to use the full range of available pixel intensities. Within an experiment, consistent adjustments were made to all control and experimental images. Micrographs are presented as maximum intensity projections unless otherwise stated.

2.6. Statistics

GraphPad Prism 7 or 8 software was used to perform statistical analysis of the data. T-tests were performed to determine if the differences between the means of two datasets were statistically significant. To achieve this, the D'Agostino-Pearson omnibus normality test was first used to assess whether data were normally distributed. If so, an unpaired t-test with Welch correction (which does not assume equal standard deviation between datasets) was applied. If the data were not normally distributed, a non-parametric unpaired Mann-Whitney U test was applied. In each case, the resulting *p*-values were inspected and significance was defined as $p<0.05$.

To assess whether the differences between the means of three or more datasets were statistically significant, one-way ANOVAs were used. Again, normality of the data was assessed. If the data was normally distributed, ordinary one-way ANOVAs were applied.

Otherwise, data were assessed using the non-parametric Kruskal-Wallis test. In each case, the resulting *p*-values were inspected and significance was defined as $p<0.05$.

For all graphs shown in this thesis, * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$, **** represents $p<0.0001$.

2.7. Bioinformatics

2.7.1. Identification of putative phosphorylation sites

Amino acid sequences of CP110 and Cep97 in various *Drosophila* species (*D. melanogaster, D. simulans, D. sechellia, D. erecta*, *D. yakuba, D. ananassae, D. pseudoobscura, D. persimilis, D. willistoni, D. virilis, D. mojavensis, D. grimshawi*) were obtained from FlyBase (Larkin *et al.*, 2020) and aligned with the Clustal Omega algorithm (Madeira *et al.*, 2019). Alignment revealed two clusters of Ser/Thr (S/T) residues in CP110 (within amino acids 270-312 and 480-553). The primary sequence was not well-conserved, but the high concentration of S/T sites suggested these regions might be regulatory. These two regions were therefore targeted for deletion.

2.7.2. Identification of cyclin-dependent kinase (CDK) sites

Amino acid sequences of CP110 and Cep97 were inspected for Ser-Pro/Thr-Pro (SP/TP) sites. These are well-known as the preferred phosphorylation sites of CDK Constructs were designed in which these sites were mutated to Ala-Pro (AP, to generate mutants that

cannot be phosphorylated by CDK) or Asp-Pro/Glu-Pro (DP/EP, to 'mimic' a constitutively phosphorylated protein).

2.7.3. Structural predictions

Amino acid sequences of Cep104 protein from *H. sapiens* and *D. melanogaster* were obtained from UniProt (UniProt-Consortium, 2021) and FlyBase (Larkin *et al.*, 2020) respectively. Alignments of these sequences were generated using NCBI BLAST's Needleman-Wunsch algorithm (NCBI-Resource-Coordinators, 2018). PSIPRED (Buchan and Jones, 2019, Jones, 1999, Jones and Cozzetto, 2015) and trRosetta (Yang *et al.*, 2020) were used to generate secondary and tertiary structural predictions of *Drosophila* Cep104, respectively.

Chapter 3: The role of CP110 and Cep97 in centriole MT length control

3.1. Introduction

3.1.1. Centriole length is regulated

The growth of the cartwheel structure during S-phase is approximately linear and its growth rate and growth period are homeostatically regulated, ensuring that daughter centrioles grow to reach the same size as their mothers (Aydogan *et al.*, 2018). Plk4, the master regulator of centriole duplication, is recruited to centrioles in an oscillatory manner and the parameters of this oscillation are important for determining both the growth rate and growth period of the cartwheel (Aydogan *et al.*, 2020, Takao *et al.*, 2019). Centriole biogenesis also involves the assembly of centriole MTs and, in fly embryos, these MTs grow to approximately the same size as the cartwheel structure (Callaini *et al.*, 1997, Gonzalez *et al.*, 1998). However, little is known about how the length of the centriole MTs is regulated.

3.1.2. CP110/Cep97 are implicated in centriole MT length control

Two proteins that have been implicated in the control of centriole MT length are CP110 and Cep97. CP110/Cep97 bind at the distal-ends of centrioles and appear to act as a

functional complex, with each protein relying on the other for its proper recruitment (Kleylein-Sohn *et al.*, 2007, Spektor *et al.*, 2007). Early studies found that both proteins 'cap' mother centrioles to inhibit ciliogenesis (Spektor *et al.*, 2007, Tsang *et al.*, 2008). Subsequent work showed that centrioles may grow longer in the absence of CP110 without necessarily forming cilia, suggesting that these proteins more generally act to cap centriole length (Kohlmaier *et al.*, 2009, Schmidt *et al.*, 2009, Tang *et al.*, 2009). Centrioles in *cp110* null *Drosophila* wing discs have slightly elongated cartwheels but dramatically elongated centriole MTs (Franz *et al.*, 2013), suggesting that, in flies, CP110's main role in centriole growth control is the capping of centriole MTs. Similar MT extensions are also apparent in *cep97* mutants, although *cep97* mutant phenotypes have been reported to be more severe than *cp110* mutant phenotypes (Dobbelaere *et al.*, 2020, Galletta *et al.*, 2016b).

CP110 does not seem to be essential for centriole biogenesis, since knock-outs in flies and mice still have centrioles (Franz *et al.*, 2013, Yadav *et al.*, 2016). However, when Plk4 is overexpressed in human cells (which usually induces centriole over-duplication), centrioles do not over-duplicate if CP110 is depleted (Kleylein-Sohn *et al.*, 2007). Meanwhile, in flies, overexpressed centriole duplication factors were more efficient at inducing centriole over-duplication in the absence of CP110 (Franz *et al.*, 2013). These data indicate a potential complex relationship between CP110 and the core centriole duplication machinery.

These previous studies into the effect of CP110/Cep97 on centriole MT length have largely relied on electron microscopy (EM) due to the challenges of tracking centriole MT

growth using live cell imaging. In this Chapter, I wanted to follow daughter centriole MT growth dynamically in living *Drosophila* embryos, in order to elucidate whether these proteins play a role in centriole biogenesis.

3.1.3. Key findings in this chapter

Some of the work in this chapter was carried out in collaboration with Mustafa Aydogan and Alan Wainman (Raff Lab, University of Oxford), together with Felix Zhou (Lu Lab, University of Oxford), whose contributions are noted throughout the text. We first show that CP110/Cep97 localise to *Drosophila* centrioles in the early embryo, and that they are required for each other's recruitment. We develop an automated technique for tracking the dynamic growth of centriole MTs over the course of S-phase, and use this to demonstrate that the dose of CP110/Cep97 may influence the growth rate and period of these centriole MTs. We find that CP110/Cep97 are recruited to centrioles in an oscillatory manner during the course of the cell cycle, but I demonstrate that the properties of this oscillation do not appear to regulate MT growth dynamics. However, we discover that CP110/Cep97 appear to play an unexpected role in regulating the growth of the central cartwheel, possibly via an interaction with Plk4, the master regulator of centriole biogenesis. These results lead me to propose a cartwheel-driven model of centriole assembly.

3.2. Results

3.2.1. CP110/Cep97 are required for each other's recruitment to *Drosophila* centrioles

Previous studies on the role of CP110/Cep97 in centriole MT growth have primarily utilised EM and have demonstrated that loss of these distal-end proteins causes a dramatic elongation of the centriole MTs (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013). However, we were interested in extending these studies to live embryos so that we could study this process dynamically. A. Wainman had previously used super-resolution microscopy on fixed and stained wing discs to demonstrate that CP110/Cep97 are recruited to both the mother and daughter centriole, where they localise as rings. The diameter of these rings is consistent with these proteins binding at/close to the centriole MTs (**Figure 3.1**). This finding is also consistent with a report that CP110 localisation overlays with the MTs of the axoneme in human RPE-1 cells (Yang *et al.*, 2018).

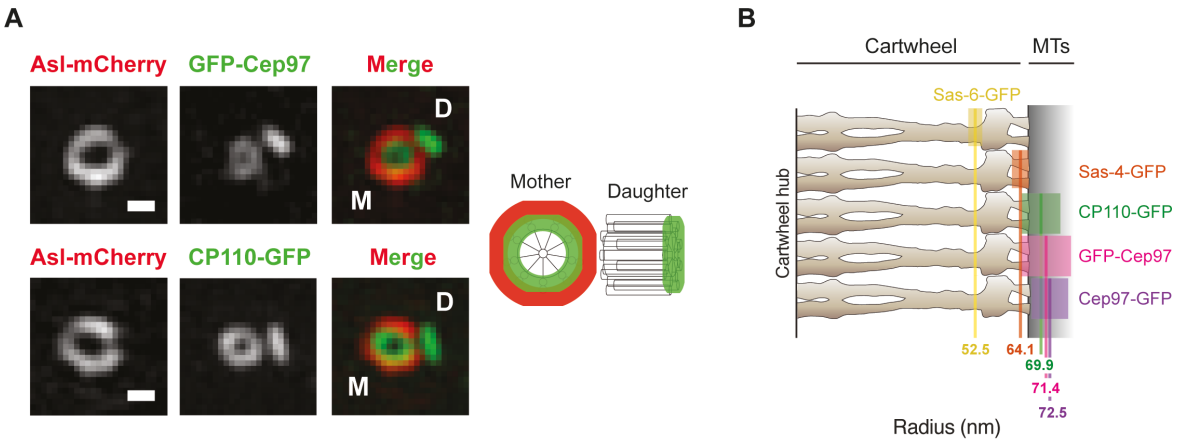


Figure 3.1 – CP110/Cep97 localisation is consistent with them binding to the centriole MTs.

(A) 3D-SIM images showing Ubq-GFP-Cep97 and Ubq-CP110-GFP localisation at the centriole of *Drosophila* wing discs. Asl-mCherry is shown as a marker of the mother centriole (Novak *et al.*, 2014), and to determine whether mother centrioles are oriented towards the microscope objective. CP110 and Cep97 localise to both mother (M) and daughter (D) centrioles. They form a ring on the mother that lies inside the Asl-mCherry ring. Scale bars represent 200 nm. **(B)** Schematic representation of radius measurements from images in (A), showing where these proteins lie relative to the underlying centriole structure. The centriole structure is adapted from Guichard *et al.*, 2013. The possible distribution of each protein relative to the structure is represented by coloured boxes, and demonstrate that CP110/Cep97 seem to localise near to the centriole MTs. The positions of Sas-6-GFP and Sas-4-GFP were previously published in Gartenmann *et al.* 2017. Figure courtesy of Alan Wainman for Hankins, Aydogan *et al.* (in preparation).

To address whether this is also the case in live embryos, I used 3D-SIM microscopy to examine embryos expressing Asl-mCherry – a marker of the mother centriole (Novak *et al.*, 2014) – together with either Ubq-GFP-Cep97 or Ubq-CP110-GFP. I confirmed that CP110/Cep97 are also recruited to both mother and daughter centrioles in live embryos (**Figure 3.2 A, B**). To test whether these proteins are required for each other's recruitment, as has been demonstrated in human U2OS cells (Spektor *et al.*, 2007), I also imaged Ubq-CP110-GFP in embryos where Cep97 has been genetically ablated (Dobbelaere *et al.*, 2020), and Ubq-GFP-Cep97 in embryos where CP110 has been genetically ablated (Franz *et al.*, 2013). I observed that CP110-GFP recruitment is dramatically reduced when Cep97

is absent; similarly, GFP-Cep97 recruitment is dramatically reduced when CP110 is absent (**Figure 3.2 A, B**). To quantify this effect, I also imaged these embryos on a spinning disc confocal microscope. 20 minutes into S-phase of nuclear cycle 14 (when the GFP signal at centrioles should be brightest – see experiments later in this Chapter), I measured the level of GFP fluorescence at ten centriole pairs in each embryo (**Figure 3.2 C, D**). I conclude that CP110 and Cep97 are required for each other's recruitment to the centriole in fly embryos.

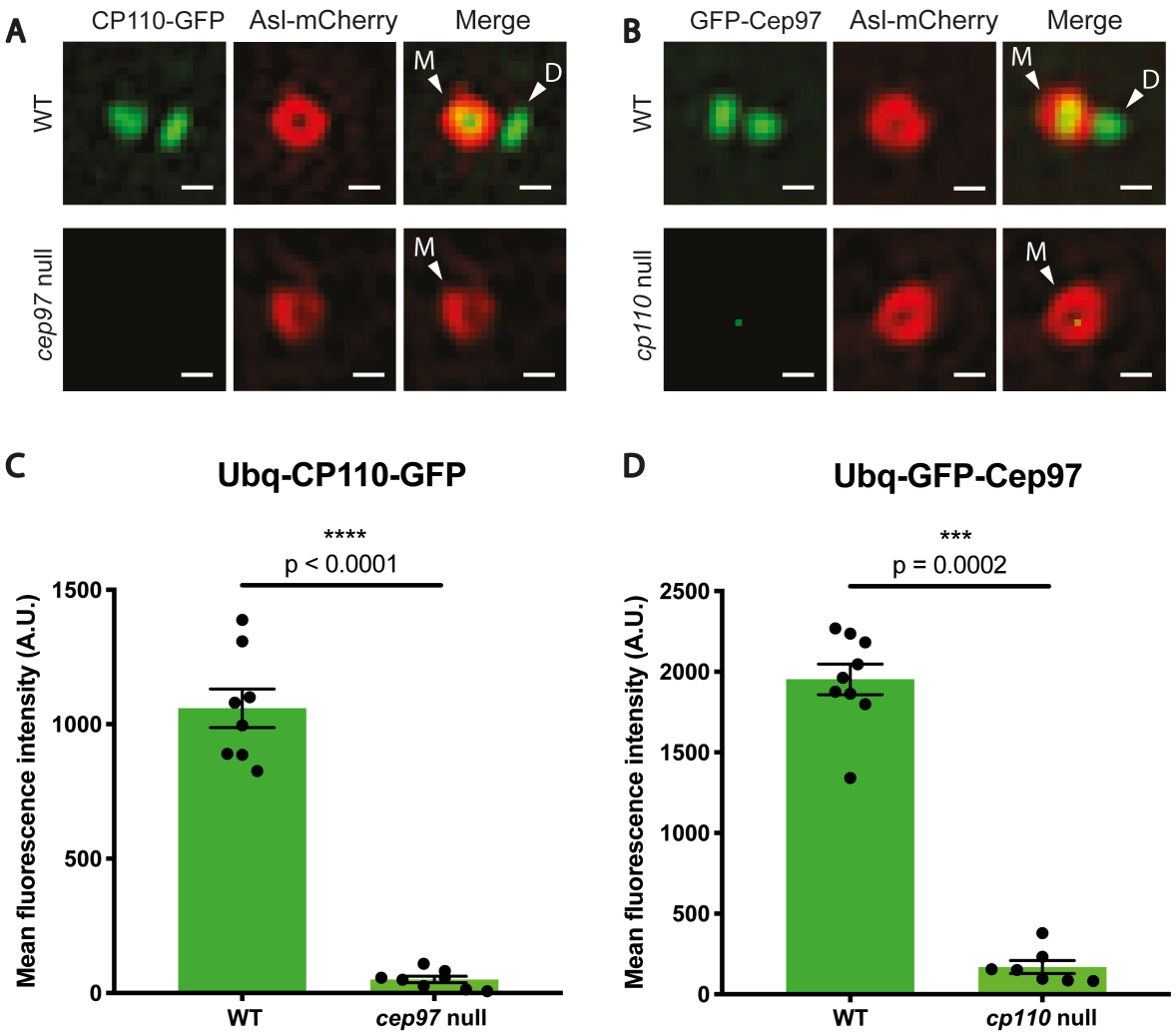


Figure 3.2 – CP110/Cep97 are required for each other's recruitment to centrioles.

(A, B) 3D-SIM images showing Ubq-CP110-GFP (A) or Ubq-GFP-Cep97 (B) localisation in WT (*above*) and either *cep97* or *cp110* null (*below*) embryos. Asl-mCherry is shown as a marker of the mother centriole. In WT embryos, CP110 and Cep97 localise to both mother (M) and daughter (D) centrioles; the CP110-GFP signal is dramatically reduced in *cep97* null embryos, and the GFP-Cep97 signal is dramatically reduced in *cp110* null embryos. Scale bars represent 200 nm. **(C, D)** Bar charts showing quantification of CP110-GFP (C) or GFP-Cep97 (D) centriolar fluorescence intensity in WT and either *cep97* or *cp110* null embryos (as indicated). Data are presented as mean ± standard error of the mean (SEM); N ≥ 7 embryos, n = 10 centriole pairs measured within each embryo, 20 min into S-phase of cycle 14. A.U. = arbitrary units. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

3.2.2. Centriole MTs grow linearly during S-phase and stop growing before the onset of mitosis

3.2.2.1. Live imaging of centriole MT growth in *Drosophila* early embryos

A key goal of this project was to assay centriole MT growth dynamics. Thus far, we had established that CP110/Cep97 bind at centrioles in the *Drosophila* early embryo, and A. Wainman's experiments had suggested that they are recruited to the tips of the centriole MTs. We therefore reasoned that it would be possible to use GFP-Cep97 as a marker of the growing ends of the centriole MTs. This is because, if CP110/Cep97 bind at the distal-ends of centriole MTs, we would expect the distance between the mother and daughter GFP-Cep97 signals to increase during centriole biogenesis as the daughter centriole MTs grow (**Figure 3.3 A**). We could therefore use this increase in distance as a proxy measure for the increase in length of the centriole MTs over time. This would allow us to follow centriole MT growth dynamically over the course of S-phase.

Fully-grown centrioles in *Drosophila* embryos are just 180-200 nm long (Gonzalez *et al.*, 1998, Moritz *et al.*, 1995), so the distance between the mother and daughter GFP-Cep97 signals, especially in early S-phase when the daughter has just started to grow, is not resolvable by conventional light microscopy. To address this, M. Aydogan acquired images of WT embryos expressing one copy of Ubq-GFP-Cep97 using a Zeiss 880 Laser Scanning Microscope (LSM) fitted with an Airyscan detector. Airyscan processing rendered the mother and daughter signals visible as separate puncta in the resulting micrographs, thus making it possible to measure the distance between each centriole pair. It is possible to make these measurements by hand, and this approach was successfully

applied to investigate the effect of reduced Plk4 expression on centriole MT growth (Aydogan *et al.*, 2018); these results are reproduced here for comparison (**Figure 3.3 B**). However, this process is very labour-intensive, taking more than a month of intensive analysis to produce the data shown in **Figure 3.3 B**.

Working with M. Aydogan, I established a collaboration with Felix Zhou to develop an automated measurement approach, which incorporates a machine learning algorithm to automatically identify mother and daughter centriole pairs within a micrograph and measure the distance between them. I will refer to this as the ***'GFP-Cep97 distancing algorithm'***. F. Zhou trained this algorithm using M. Aydogan's initial hand-annotated images as a training set. Importantly, I had recently found that there was an asymmetry between the fluorescence intensity of Ubq-GFP-Cep97 on mother and daughter centrioles, and that the degree of this asymmetry changed over the course of S-phase, with the daughter being much brighter than the mother by mid-S-phase (data discussed below; **Figure 3.10**). We therefore realised it was important to separately train the algorithm on images from early-, mid-, and late-S-phase, in order to account for this difference. More details on the development and testing of the GFP-Cep97 distancing algorithm can be found in **Appendix A1**.

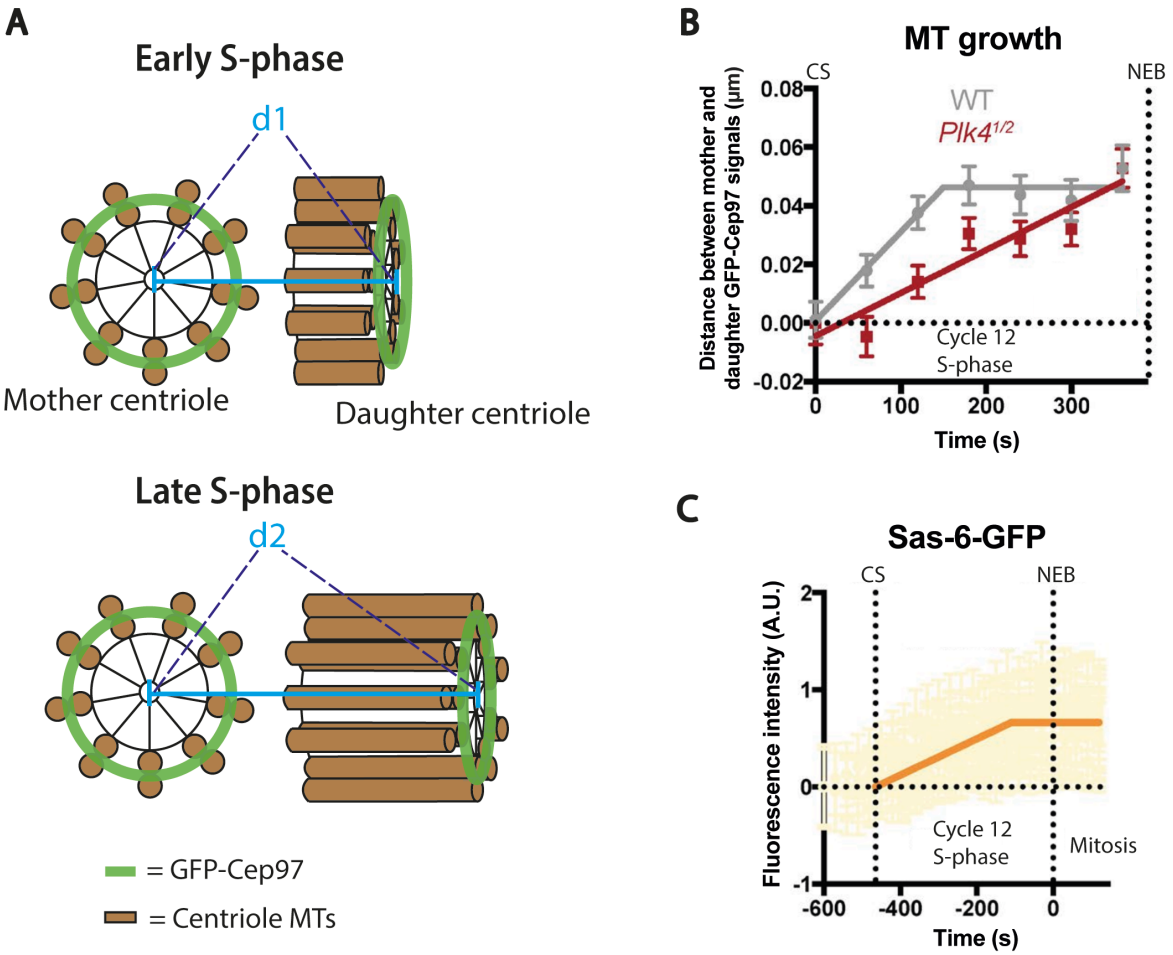


Figure 3.3 – Development of an assay to track the dynamic growth of centriole MTs.

(A) Schematic showing the localisation of GFP-Cep97 on mother and daughter centrioles. Mother centrioles are oriented towards the cortex in *Drosophila* syncytial embryos. d1 represents the initial distance between the mother and daughter GFP signals in early S-phase. As the daughter centriole MTs grow, we would expect this distance to increase (d2). The difference between d1 and d2 can be used as a proxy measurement for the increasing length of daughter centriole MTs at different time points in S-phase, thus allowing dynamic measurement of centriole MT growth. Adapted from Hankins, Aydogan *et al.* (in preparation). **(B)** Figure adapted from Aydogan *et al.* 2018 showing the increasing mean distance between the GFP-Cep97 signals on mothers and daughters over S-phase of cycle 12 in WT embryos (grey) or embryos with reduced Plk4 expression (*Plk4*$^{1/2}$, maroon). These distances can be used as a proxy measure for centriole MT length, and show that centriole MTs in WT embryos grow in an approximately linear fashion and stop growing before the onset of mitosis. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM; N ≥ 4 embryos for each dataset, n ≥ 69 centriole pairs per embryo. **(C)** Figure adapted from Aydogan *et al.* 2018 showing that the incorporation of Sas-6-GFP (regressed from hundreds of centriole pairs measurements from N ≥ 15 embryos) is approximately linear and stops before the onset of mitosis. Sas-6-GFP incorporation is used here as a proxy for centriole cartwheel growth.

3.2.2.2. Perturbing the global level of CP110/Cep97 affects the growth rate and period of centriole MTs

With this automated GFP-Cep97 distancing assay in place, we wanted to test whether altering the global levels of CP110/Cep97 might affect the growth rate and growth period of the centriole MTs. I had already shown that GFP-Cep97 does not localise to the tips of growing centrioles in embryos lacking CP110, so we decided to investigate the effect of overexpressing CP110/Cep97. M. Aydogan collected images of embryos expressing Ubq-GFP-Cep97 in WT embryos (which I shall refer to as ***'WT embryos'***) and in embryos that overexpressed an untagged version of CP110 (which I shall refer to as ***'Ubq-CP110 embryos'***), which is expressed at ~3-4x higher levels than endogenous CP110 (M. Aydogan, personal communication). Note that Cep97 is overexpressed in both sets of embryos, due to the expression of Ubq-GFP-Cep97. Therefore, in the Ubq-CP110 embryos, both CP110 and Cep97 are overexpressed. These images were then analysed using the GFP-Cep97 distancing algorithm. I took the output from the script and plotted the increase in distance between the mother and daughter GFP-Cep97 signal for each individual embryo over time.

I then attempted to analyse the growth rate and growth period of the centriole MTs in each embryo. Previous analysis of centriole MT growth via hand measurements of GFP-Cep97 distancing had shown that they best fitted a linear mode of growth, and that they stop growing before the onset of mitosis (Aydogan *et al.* 2018, reproduced in **Figure 3.3 B**). This is also strikingly reminiscent of how the cartwheel grows (Aydogan *et al.* 2018, reproduced in **Figure 3.3 C**). Based on this finding, and following extensive trials of different regression models, we reasoned that our data could also be modelled with linear

growth (**Figure 3.5 A-C**). The growth rate of the centriole MTs can therefore be determined by the gradient of the linear increase function, while their growth period is given by the time at which their growth reaches a plateau.

Since the mode of centriole MT growth seemed similar to that of the central cartwheel structure, I first attempted to analyse the data in the same way as we would the cartwheel growth data, as described in Aydogan *et al.* 2018. I therefore resolved to model the growth rate and growth period of the centriole MTs in each embryo by comparing the fits of two regressions, to determine which mode of linear increase fitted each individual embryo the best:

(1) Linear increase

(2) Linear increase followed by a plateau

Mathematically, this comparison approach is the purest way of completing this analysis, since it compares models of the two possible forms of linear growth. I also thought it biologically plausible that embryos in which the growth rate is slow might best fit a linear model of growth *(1)*, while embryos exhibiting faster growth might best fit a linear growth with plateau model *(2)*. During this comparison process, I found that 43.8% of WT embryos preferentially fitted *(2)*, compared with 23.1% of Ubq-CP110 embryos (**Figure 3.4**).

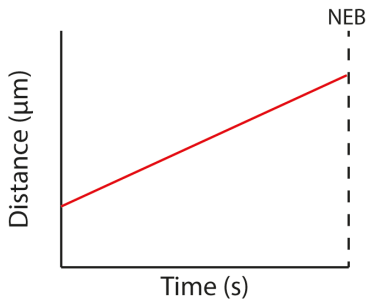
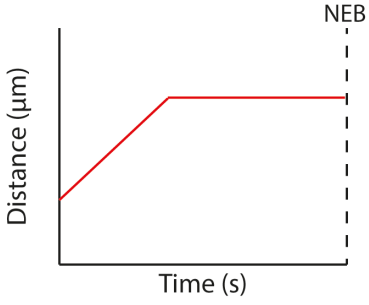
	Linear increase	Linear increase followed by a plateau
		
Genotype	**Number of embryos that fit *(1)***	**Number of embryos that fit *(2)***
WT	9	7
Ubq-CP110	10	3

Figure 3.4 – Summary of the number of embryos in each dataset that were fitted with model *(1)* or model *(2)*.

Table showing the number of embryos in each dataset where the preferred fit was either *(1) Linear increase* or *(2) Linear increase followed by a plateau.* Schematics of each model are shown above.

Having fitted each embryo with either model *(1)* or *(2)* (i.e. whichever fit the data best), I used the parameters of these best fits to extract the MT growth rate and growth period for each embryo. This allowed me to compare MT growth dynamics in the two datasets. I found that there was a significant decrease in the growth rate of the centriole MTs when CP110/Cep97 are both overexpressed (**Figure 3.5 D**). The distance acquired between the mother and daughter GFP-Cep97 signals (a proxy measure for the final length of the daughter centriole MTs) was also significantly reduced in the Ubq-CP110 embryos. These results suggest that perturbing the levels of CP110/Cep97 in the embryo influences the growth parameters of the MTs, with CP110/Cep97 overexpression resulting in a slower MT growth rate and thus shorter centrioles; there was also a hint that the centriole MTs might grow for a longer period when CP110/Cep97 are overexpressed, but this difference was not statistically significant.

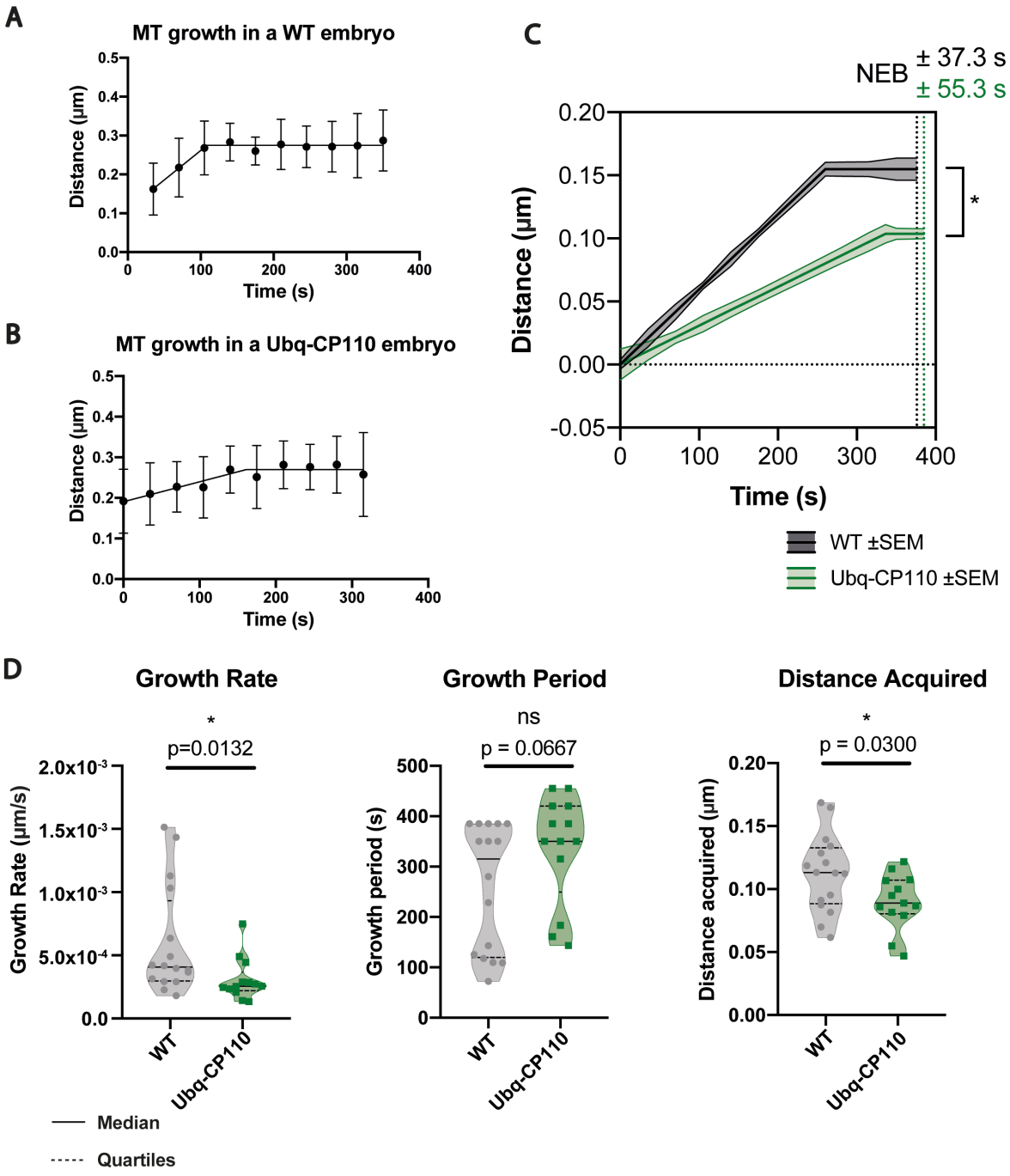


Figure 3.5 – Centriole MTs grow slower and possibly for a longer period when CP110/Cep97 are overexpressed.

(A-B) Graphs showing the increase in distance between GFP-Cep97 signals over time in a representative WT (A) and Ubq-CP110 (B) embryo. *Linear increase followed by a plateau* regressions are plotted here as black lines to show how these fit the raw data. Data are presented as mean ± standard deviation (SD). **(C)** Average curves showing how centriole MTs grow during S-phase of cycle 12 in WT and Ubq-CP110 embryos. NEB = Nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(D)** Violin plots showing quantification of centriole MT growth rate, growth period and final MT length (given by the proxy measurement of 'distance acquired') in WT and Ubq-CP110 embryos. Data are represented as median ± interquartile range; N ≥ 13 embryos. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

While these results are potentially interesting, I was concerned that these findings might be skewed by the fact that I was fitting two different models to the data, and the proportion of embryos fitted with each model was different between WT and Ubq-CP110 embryos (**Figure 3.4**). As argued above, this may be biologically justified, since if CP110/Cep97 overexpression really slows MT growth, it might be expected that more of these embryos would continue to grow throughout S-phase and not reach a plateau. I decided to check whether I could see the same trends in centriole MT growth parameters when I only compared embryos that preferentially fitted model *(2)*. Encouragingly, the trends in growth rate, growth period and distance acquired were consistent with what I had observed in the original analysis (**Figure 3.6 A**). However, the low *n* numbers in this analysis mean that these trends do not appear to be statistically significant; as **Figure 3.4** shows, just 7 WT embryos and 3 Ubq-CP110 embryos fitted model *(2)*. I therefore attempted a final analysis, where I consistently applied model *(2)* fits to all the data in order to increase my sample size. Again, plotting the growth parameters from these forced fits showed a significant decrease in the growth rate of the centriole MTs when CP110/Cep97 are overexpressed, as well as the same trend towards a longer growth period (**Figure 3.6 B**). Furthermore, the trend towards a shorter final MT length in Ubq-CP110 embryos was also captured. This gave me confidence that MT growth rate decreases, and there is a trend towards a longer MT growth period, when CP110/Cep97 are overexpressed.

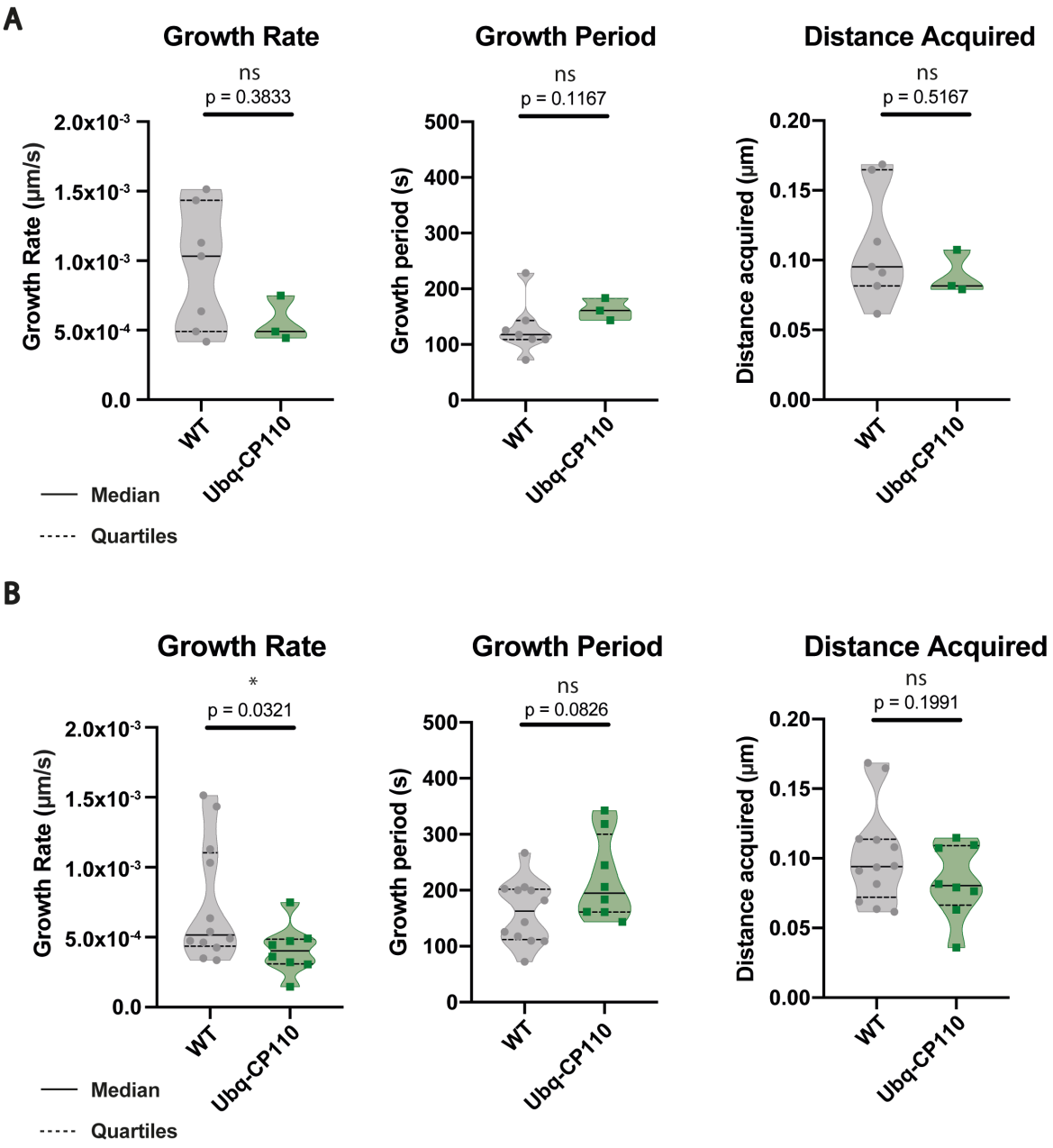


Figure 3.6 – Further examples of MT distancing analysis.

(A-B) Violin plots showing quantification of centriole MT growth rate, growth period and final MT length (given by the proxy measurement of 'distance acquired') in WT and Ubq-CP110 embryos. Results are plotted only from embryos which preferentially fitted regression model *(2)* (A), or from fitting all embryos using regression model *(2)* (B). The growth parameters shown are extracted from these fits. Data are presented as median ± interquartile range; N ≥ 3 embryos (for A), N ≥ 8 embryos (for B). T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

As explained above, it is not possible to investigate centriole MT dynamics in *cp110* null embryos, since I showed in **Figure 3.2** that our GFP-Cep97 marker is not recruited to centrioles when CP110 is absent. Therefore, in an attempt to investigate what effect decreasing the dose of these distal-end proteins has on centriole MT growth, I decided to look at a recently published fly line where Cep97-GFP expression is driven by the endogenous Cep97 promoter; this should therefore be expressed at lower levels than in the Ubq-GFP-Cep97 stock (Dobbelaere *et al.*, 2020). I shall refer to this fly line as ***'eCep97-GFP'*** from now on. Unfortunately, I found that the fluorescence intensity of this fusion protein at centrioles was too low to detect on the Zeiss 880 LSM during early- and late-S-phase. Therefore, I could not pursue this approach.

3.2.2.3. Attempts to use CP110-GFP as a marker to assay MT growth

Having acquired results from GFP-Cep97 embryos expressing different levels of CP110, we also wanted to apply this assay to embryos expressing CP110-GFP. This would allow us to investigate the effects of different doses of Cep97 on MT growth dynamics. M. Aydogan generated flies expressing Ubq-CP110-GFP in either WT embryos or embryos expressing Ubq-Cep97. He collected images of CP110-GFP distancing in these embryos during cycle 12, and these images were then analysed with the GFP-Cep97 distancing algorithm. Again, I took the script output and plotted the data for each embryo. I then compared the goodness-of-fit of regression models *(1)* and *(2)*, as above.

However, on inspection of the algorithm's detections, I noticed that the script seemed to be erroneously identifying background signal as centriole pairs, and that this effect appeared to be more dramatic than it had been in the GFP-Cep97 dataset (**Figure 3.7 A, B**). Indeed,

I had previously observed when imaging these embryos that Ubq-CP110-GFP embryos appear to have higher levels of background signal than Ubq-GFP-Cep97. Since the script had originally been trained on the GFP-Cep97 images, we decided to attempt a retraining on the CP110-GFP images. It was hoped that this retraining would improve the script's performance when presented with CP110-GFP images, since there appeared to be a difference in signal:noise ratio between these two genotypes. To do this, I hand-annotated the images of CP110-GFP to indicate the location of the mother and daughter centrioles (**Figure 3.7 C**). These images were then used as a training set for the machine learning algorithm. F. Zhou also applied a higher threshold to the script during processing in a further attempt to eliminate background signal, since the background signal should have a lower fluorescence intensity than the signal at the centrioles. I took the output of this dataset and compared the results of my analysis with the graphs generated before retraining and higher thresholding. The graphs from both treatments were very similar (**Figure 3.7 D**). This suggested that the retraining and higher thresholding had not improved the ability of the script to correctly identify centriole pairs labelled with CP110-GFP. However, it was encouraging to see that the script gives consistent results. Since the level of background signal in CP110-GFP embryos had proved a challenge, we decided to focus on the GFP-Cep97 approach.

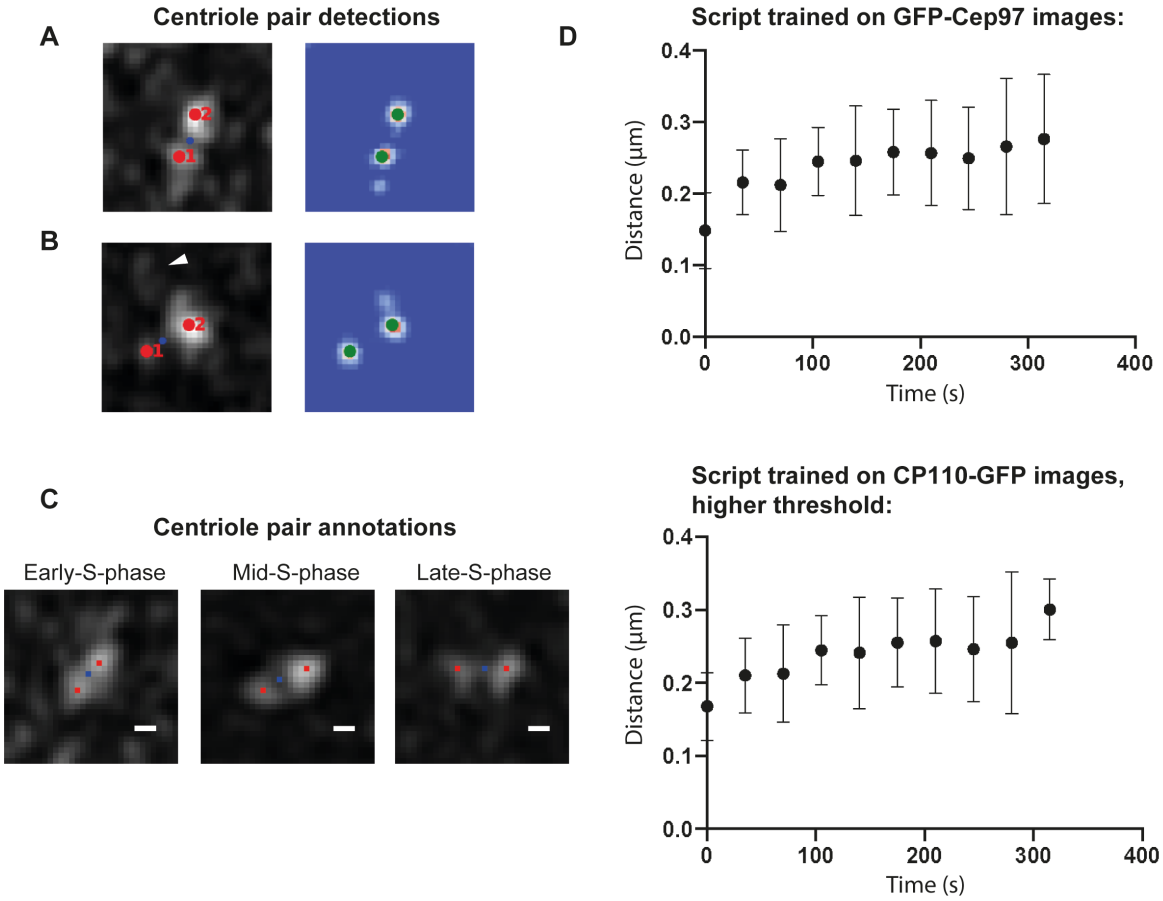


Figure 3.7 – Retraining the GFP-Cep97 distancing algorithm on CP110-GFP data.

(A-B) Zeiss 880 LSM Airyscan-processed images showing Ubq-CP110-GFP localisation at mother and daughter centrioles in WT embryos. In (A), the two centrioles have been successfully detected (indicated by the red labels) by the GFP-Cep97 distancing algorithm. In (B), one centriole has been successfully detected (indicated by the red label '2') by the GFP-Cep97 distancing algorithm. However, the algorithm has also detected a feature that seems likely to represent background signal (indicated by red label '1'). A similar feature in the vicinity of the correctly identified centriole is indicated by the arrowhead. Images represent 32x32 pixel crops from the input micrograph. **(C)** Examples of annotations that I made, by hand, to images of Ubq-CP110-GFP at different periods in S-phase of cycle 12. These annotations were used as a new training set for the distancing algorithm, in an attempt to generalise it to images of CP110-GFP. Mother and daughter centrioles are marked with red dots, and the space between them is marked with a blue dot. Scale bar represents 200 nm. **(D)** Graphs showing the increase in distance between CP110-GFP signals over time in a representative WT embryo expressing one copy of Ubq-CP110-GFP when the images are passed through either the original GFP-Cep97 distancing algorithm (*above*) or the retrained algorithm employing a higher threshold (*below*). The graphs appear strikingly similar to one another. Data are presented as mean ± SD.

3.2.3. CP110/Cep97 oscillate at the centriole

Our results thus far suggested that overexpression of CP110/Cep97 causes a reduction in centriole MT growth rate, and a trend towards a longer centriole MT growth period, which ultimately results in a reduction in centriole MT length when compared with WT centrioles. We were therefore curious as to the mechanism by which CP110/Cep97 were affecting the growth of the centriole MTs; it seems possible that, under overexpression conditions, more CP110/Cep97 is recruited to the centriole MTs and that this accumulation at the growing tips somehow hinders MT growth. This might explain how the levels of CP110/Cep97 are able to influence MT growth rate, but does not account for the potential change in growth period. We therefore decided to explore the dynamics of CP110/Cep97 at the centriole, in order to gain more insight into how these proteins are recruited.

3.2.3.1. CP110/Cep97 levels oscillate at the centriole, peaking in S-phase

I imaged Ubq-GFP-Cep97 or Ubq-CP110-GFP over nuclear cycles 11-13 in collaboration with M. Aydogan. These movies were then analysed in TrackMate (Tinevez *et al.*, 2017) to investigate how the fluorescence intensity of these proteins changes at the centriole over time. We found that CP110 and Cep97 are both recruited to the centriole in an oscillatory manner, with their levels reaching a peak in S-phase (**Figure 3.8 A, B**). This was an exciting observation, since Plk4 levels also oscillate at the centriole (Aydogan *et al.*, 2020, Takao *et al.*, 2019). We found that CP110/Cep97 appeared to oscillate in phase with each other, but out of phase with Plk4 (**Figure 3.8 C, D**). These phase differences were confirmed by formal quantification of the different phase angles of the oscillations (F. Zhou, personal communication).

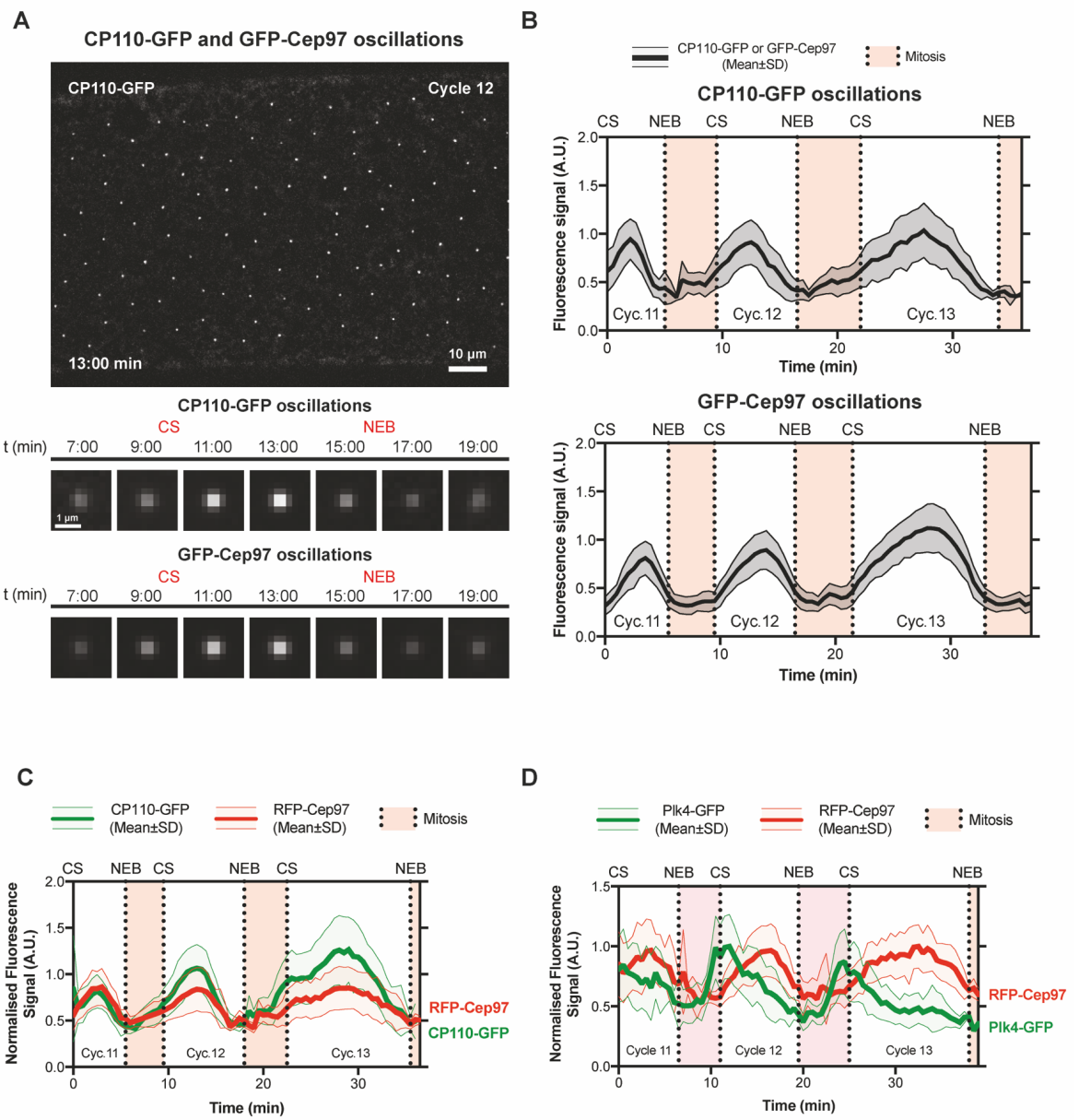


Figure 3.8 – CP110/Cep97 levels oscillate at the centriole.

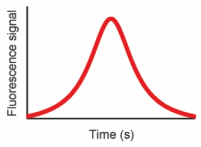
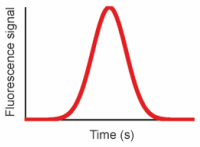
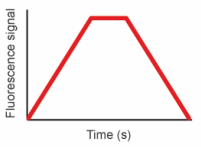
(A) Spinning disc confocal image from a movie of Ubq-CP110-GFP dynamics in the *Drosophila* early embryo over S-phase of cycle 12. Image is shown as a maximum intensity projection with background subtraction. The punctae correspond to individual centriole pairs. Panels below show a time-lapse of either Ubq-CP110-GFP or Ubq-GFP-Cep97 over cycle 12, produced by superimposing all centriole pairs (n = 150 for CP110, n = 115 for Cep97) at various time points within a single movie. **(B)** Graphs showing quantification of Ubq-CP110-GFP or Ubq-GFP-Cep97 fluorescence intensity at centrosomes in a single embryo over nuclear cycles 11-13. Centrosomal levels of these proteins oscillate, peaking in S-phase of each cycle. **(C-D)** Graphs showing quantification of Ubq-CP110-GFP and Ubq-RFP-Cep97 (C) or Plk4-GFP and Ubq-RFP-Cep97 (D) fluorescence intensity at centrosomes within the same embryo over nuclear cycles 11-13. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). Data in (B-D) are represented as mean ± SD. Imaging of (B-C) was carried out by the author and Mustafa Aydogan. Imaging of (D) is by M. Aydogan. Final figure produced by M. Aydogan for Hankins, Aydogan *et al.* (in preparation).

3.2.3.2. The peak of the CP110/Cep97 oscillation correlates with S-phase length

Having found that CP110-GFP and GFP-Cep97 levels oscillate at the centriole, we went on to image these proteins within individual cycles 11-13. M. Aydogan fitted these data against various non-linear regressions in order to determine a way of modelling the oscillation. He found that it preferentially fitted a Lorentzian model (**Figure 3.9 A**).

We used this regression to fit the CP110/Cep97 oscillation data from individual embryos and thus extract the time taken to reach the oscillation peak. We could also determine the length of S-phase in these embryos (defined as the time taken to progress from centrosome separation to nuclear envelope breakdown) and so test whether S-phase length correlates with the time of the peak. We found that the peak of the CP110/Cep97 oscillation correlated strongly with S-phase length (**Figure 3.9 B-G**). It has previously been shown that the plateau of the Sas-6 growth curve (that serves as a proxy for centriole cartwheel growth) does not correlate with S-phase length (Aydogan *et al.*, 2018), so I wanted to see how the dynamics of these proteins compared when imaged in the same embryo. I confirmed that the Sas-6-GFP plateau did not correlate with S-phase length (**Figure 3.9 J**) and found that it did not correlate with the peak of Ubq-RFP-Cep97 (**Figure 3.9 I**). This is to be expected, since the peak of the Ubq-RFP-Cep97 oscillation correlates strongly with S-phase length (**Figure 3.9 H**).

A

N = 14 embryos	Lorentzian	Gaussian	Linear increase, plateau, linear decrease	Linear increase, linear decrease
				
Deg. of freedom	13	13	11	12
R^2 value	0.99±0.01	0.98±0.01	0.92±0.10	0.98±0.01
% of fit preference	57%	14%	21%	7%

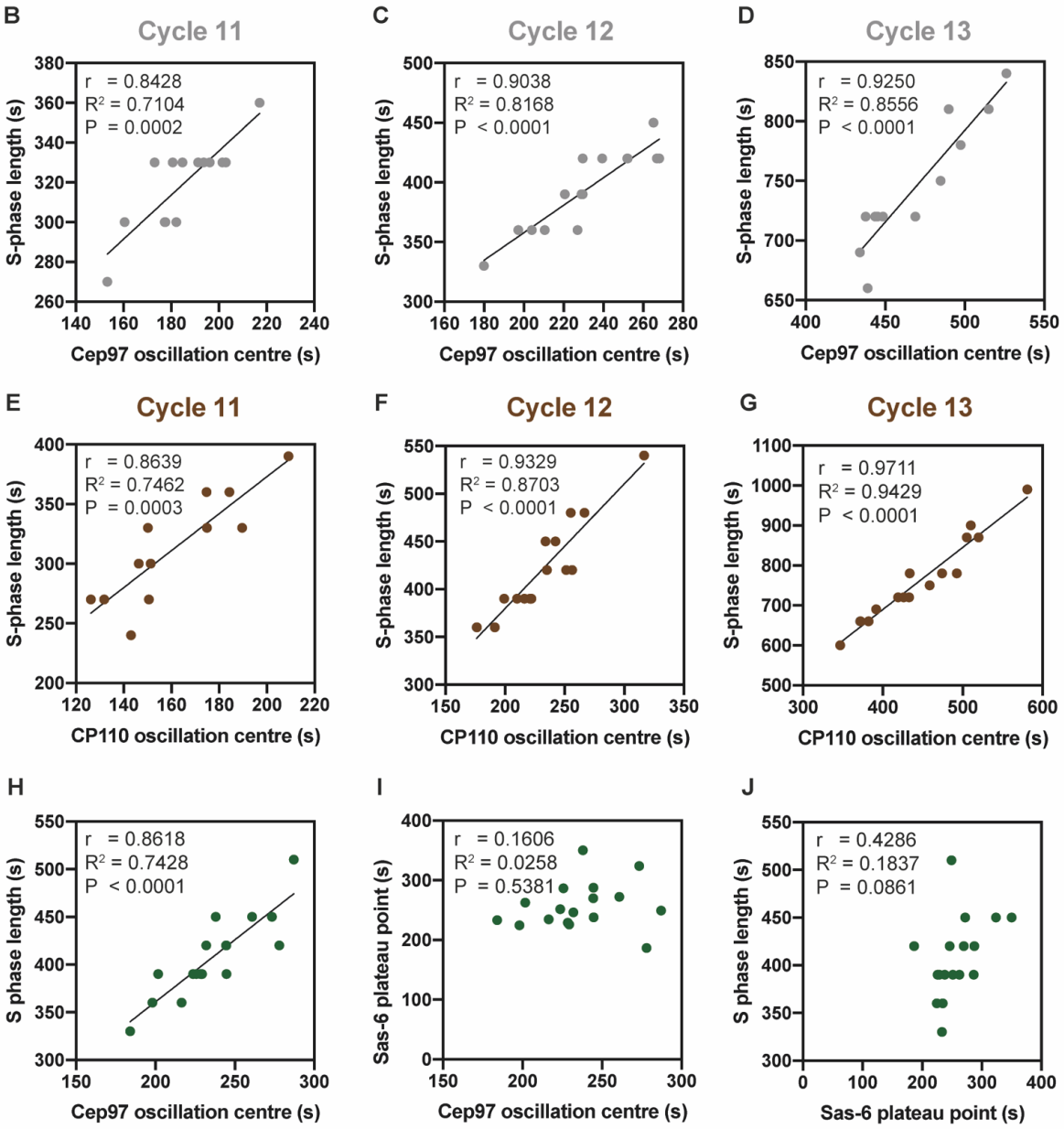


Figure 3.9 – The peak of the CP110/Cep97 oscillation correlates with S-phase length.

(A) To assess how to fit the CP110/Cep97 dynamic recruitment data, Ubq-GFP-Cep97 oscillation data were fitted against four separate regression models (illustrated in red): *(1) Lorentzian (2) Gaussian (3) Linear increase, plateau, linear decrease (4) Linear increase, linear decrease.* The majority of graphs preferentially fitted *(1)*, so this model was selected to fit all CP110/Cep97 data. R^2 values were also assessed, since these give an indication of the goodness-of-fit. Further details of these models are provided in **Chapter 2**. Analysis of Ubq-CP110-GFP data gave similar results (data not shown). Analysis and figure by Mustafa Aydogan for Hankins, Aydogan *et al.* (in preparation). **(B-H)** Graphs showing that there is a positive correlation between the time of the Ubq-GFP-Cep97 oscillation peak and S-phase length in cycle 11 (B), cycle 12 (C), and cycle 13 (D), between the Ubq-CP110-GFP oscillation peak and S-phase length in cycle 11 (E), cycle 12 (F), and cycle 13 (G), and between the Ubq-RFP-Cep97 oscillation peak and S-phase length in cycle 12 (H). **(I-J)** Graphs showing that the Sas-6-GFP plateau in cycle 12 does not correlate with the Ubq-RFP-Cep97 oscillation peak (I) or S-phase length (J). $N \geq 12$ embryos. *r*, R^2 and *p*-values are shown on the graphs for (B-J), and linear regressions have been fitted on those graphs where a statistically significant correlation was found.

3.2.3.3. CP110/Cep97 predominantly oscillate on the growing daughter centriole

We were curious as to whether this oscillation occurs on the mother or the daughter centriole, or both. This is not discernible from the spinning disc data, since it is not possible to resolve the mother and daughter centrioles in these images. I therefore imaged embryos expressing either Ubq-CP110-GFP or Ubq-GFP-Cep97 together with Asl-mCherry – a marker of mother centrioles (Novak *et al.*, 2014) – on the Zeiss 880 LSM fitted with an Airyscan detector, which can resolve the mother and daughter centrioles in living embryos. Once I had acquired images of these proteins over cycle 12, I could measure the GFP fluorescence intensity on mothers, and on daughters, in early-, mid- and late-S-phase. I found that the oscillation predominantly occurs on the daughter and, if anything, the signal on the mother centriole gradually decreases during S-phase (**Figure 3.10 A, B**).

I further refined this analysis by measuring the GFP intensity on mothers and daughters at 30 s intervals. Again, I found that the oscillation of CP110/Cep97 predominantly occurs on the daughter centriole. However, with this improved temporal resolution I also now observed a small oscillation on the mother (**Figure 3.10 C, D**). Interestingly, this oscillation appears to be slightly out of phase with the oscillation on the daughter, with the mother oscillation peaking earlier. This is most obvious in cycle 12, where the mother's oscillation seems to peak very early in S-phase before decreasing. This could explain why I observed an overall decreasing trend in mother GFP fluorescence in **Figure 3.10 B**. It would be important to test whether this oscillation also occurs on the mother in eCep97-GFP embryos. However, as discussed above, the low intensity of the centriolar signal in eCep97-GFP embryos makes such analysis challenging. Following the oscillation peak on both mothers and daughters, the fluorescence intensity appears to return to a level that is similar between cycles (**Figure 3.10 C, D**; we also observed this in **Figure 3.8 B**). This may represent a base level of CP110/Cep97 recruitment, though the intensity of the GFP signal at centrioles was often too low to measure reliably during mitosis.

I conclude that the CP110/Cep97 levels at the centriole oscillate during S-phase, and this oscillation occurs predominantly on the growing daughter centrioles. This raised the intriguing possibility that this oscillation might be important for regulating the growth of centriole MTs.

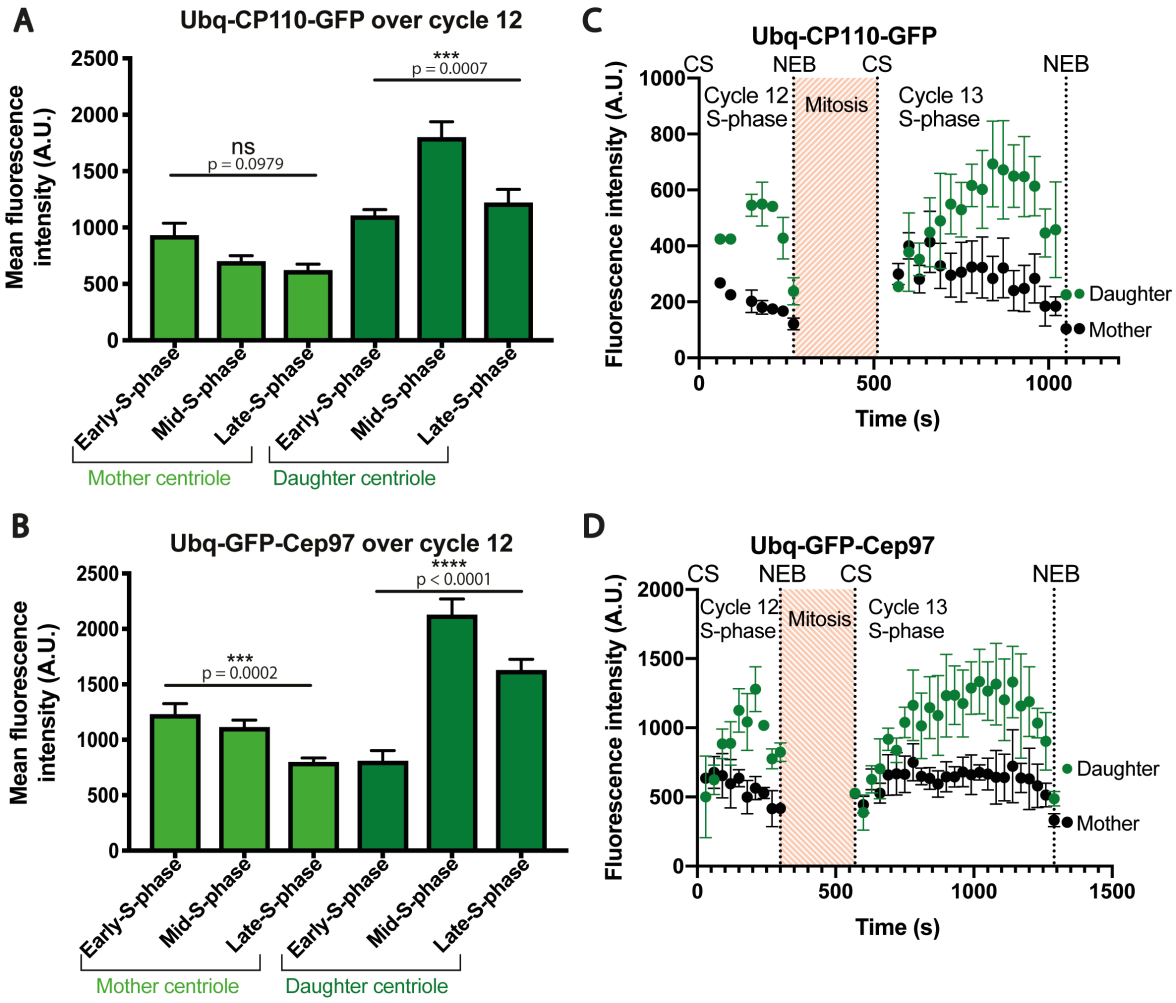


Figure 3.10 – The CP110/Cep97 oscillation occurs predominantly on the daughter centriole.

(A-B) Bar charts showing the mean fluorescence intensity of Ubq-CP110-GFP (A) or Ubq-GFP-Cep97 (B) on mother centrioles (light green) and daughter centrioles (dark green) at different periods in S-phase. Data are presented as mean ± SEM; N ≥ 7 embryos, within which individual centriole pairs were measured. ANOVA tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs. **(C-D)** Graphs showing the fluorescence intensity of Ubq-CP110-GFP (C) or Ubq-GFP-Cep97 (D) on mother or daughter centrioles in a single embryo over the course of S-phase in cycle 12 and cycle 13. Images were acquired every 30 s and fluorescence intensities were measured from micrographs at each timepoint, revealing that the oscillation predominantly occurs on the daughter centriole (green) but that a smaller oscillation also occurs on the mother (black). CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). The GFP signal was too weak in mitosis to reliably quantify. Data are presented as mean ± SD.

3.2.4. The properties of the CP110/Cep97 oscillation do not appear to regulate centriole MT growth

Recent work from our laboratory has shown that Plk4 oscillates at the centriole, and that the properties of this oscillation are important for regulating the growth rate and period of the central cartwheel structure (Aydogan *et al.*, 2020). We were therefore curious as to whether the CP110/Cep97 oscillation may play a similar role in the regulation of centriole MT growth. In particular, we wondered whether the peak of the oscillation might correlate to the time at which the centriole MTs stop growing, since this simple model might explain how these proteins are able to affect the growth rate, and potentially the growth period, of the centriole MTs.

In order to test this, I reasoned that it would be beneficial to examine the CP110/Cep97 oscillation in different doses of these distal-end proteins. Our results from exploring the dynamic growth of centriole MTs suggest that Ubq-CP110 embryos (overexpressing both CP110 and Cep97) might grow their centriole MTs for a slightly longer period of time than the WT embryos (where just Cep97 is overexpressed due to the Ubq-GFP-Cep97 marker). In other words, the plateau point of the MT distancing curve shifts later in S-phase when CP110/Cep97 are both overexpressed (**Figure 3.11 A**). If the peak of the oscillation determines the time at which MTs stop growing, or indeed if MT growth is inhibited once CP110/Cep97 recruitment surpasses a certain threshold, I would expect the peak to also shift later in S-phase when CP110/Cep97 are overexpressed (**Figure 3.11 B**).

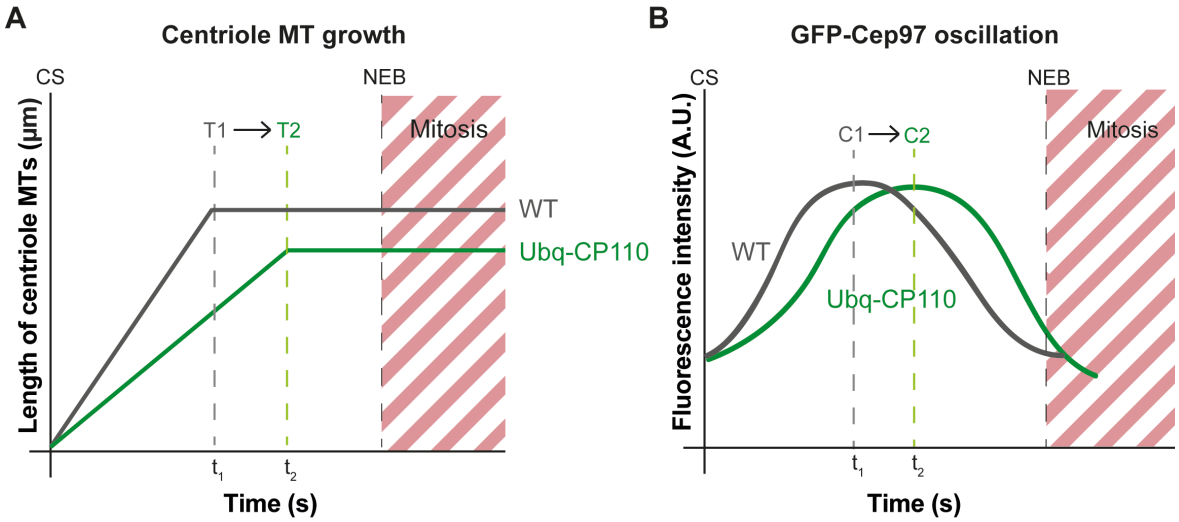


Figure 3.11 – Schematic summarising the predicted effect of CP110/Cep97 overexpression on the GFP-Cep97 oscillation, if the oscillation influences growth.

(A) Schematic summarising the results from **Figure 3.5** and **Figure 3.6**; centriole MTs (grey) in WT embryos grow in an approximately linear fashion and stop growing before the onset of mitosis. Their growth period is represented by *T1*. Centriole MTs in Ubq-CP110 embryos (green) grow slower and may also have a longer growth period (*T2*). Therefore, MTs may stop growing later in S-phase when CP110/Cep97 are overexpressed. **(B)** Schematic showing the expected effect of CP110/Cep97 overexpression on the GFP-Cep97 oscillation, if the peak of the oscillation triggers the centriole MTs to stop growing. If this model is correct, I would expect that the peak of the GFP-Cep97 oscillation in Ubq-CP110 embryos (*C2*) occurs later in S-phase than the peak of the GFP-Cep97 oscillation in WT embryos (*C1*). CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis).

3.2.4.1. Generation of a fly line expressing CP110-GFP at endogenous levels

To investigate the effects of CP110/Cep97 dosage on the CP110/Cep97 oscillation peak, I wanted to use a range of fly lines with different CP110/Cep97 expression levels. Thus far, the fly lines we have used to visualise CP110/Cep97 dynamics have expressed CP110-GFP/GFP-Cep97 under ubiquitin promoters. Using western blot, I demonstrated that expression of Cep97-GFP in embryos from the recently published eCep97-GFP fly line

(Dobbelaere *et al.*, 2020) was much lower than in Ubq-GFP-Cep97 embryos (**Figure 3.12 A**). This is consistent with results from Dobbelaere *et al.* 2020.

I also wanted to generate a fly line that expressed CP110-GFP under the *cp110* endogenous promoter. To do this, I cloned the *cp110* promoter into a plasmid containing *cp110* cDNA and a C-terminal GFP sequence. I sent this plasmid to the Cambridge Fly Facility, where it was injected into flies for integration via P-element insertion. I confirmed by western blot that the expression levels of CP110-GFP in these stocks were ~4-5x lower than in the Ubq-CP110-GFP strain (**Figure 3.12 B**). I hereafter refer to this fly line as ***'eCP110-GFP'***.

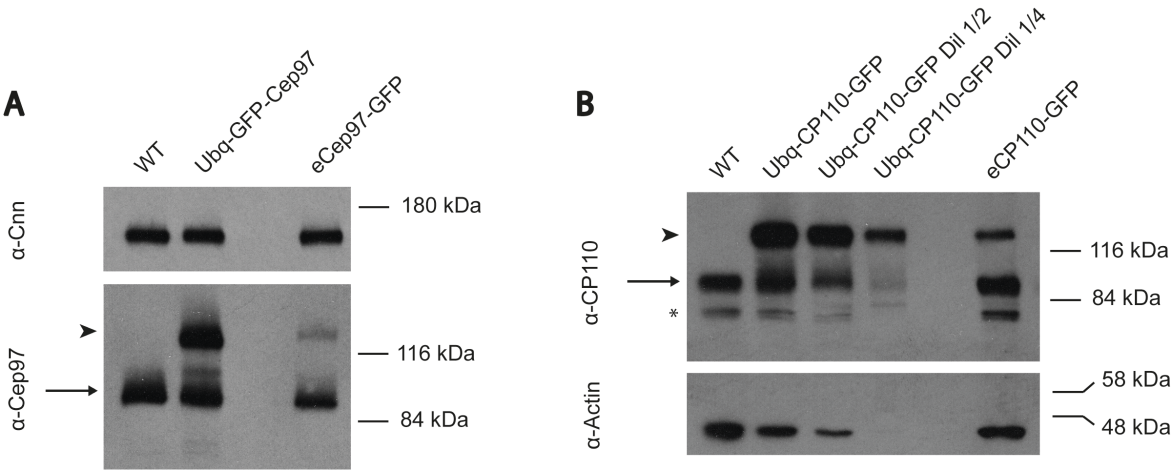


Figure 3.12 – Western blots to assess the expression of eCep97-GFP and eCP110-GFP.

(A) Western blots comparing the level of Cep97 expression in WT embryos expressing one copy of either eCep97-GFP or Ubq-GFP-Cep97. The endogenous Cep97 band is indicated with an arrow. The band corresponding to the GFP-tagged protein is indicated with an arrowhead. Cnn is shown as a loading control. N = 1. **(B)** Blots comparing the level of CP110 expression in WT embryos expressing one copy of either eCP110-GFP or Ubq-CP110-GFP. Serial dilutions of the Ubq-CP110-GFP sample were also analysed. The endogenous CP110 band is indicated with an arrow. Non-specific bands are indicated with *. The band corresponding to the GFP-tagged protein is indicated with an arrowhead. The results demonstrate that expression of CP110-GFP is ~4-5x lower in eCP110-GFP embryos than in Ubq-CP110-GFP embryos (relative levels were estimated by eye). Actin is shown as a loading control. N = 1.

3.2.4.2. Higher levels of CP110/Cep97 expression lead to faster, and higher levels of, centriole recruitment

Now that I had fly lines expressing GFP-tagged versions of both CP110 or Cep97 at two different CP110/Cep97 expression levels, I was able to compare the oscillation parameters between each genotype. It is important to note that, in the following experiments, the transgenes coding for the GFP-tagged versions of CP110 or Cep97 are expressed in embryos that also possess the native CP110/Cep97 loci. Endogenous, untagged CP110/Cep97 molecules are therefore also present in these embryos, so it is conceivable that some competition may occur between the tagged and untagged proteins for recruitment to the centriole. This strategy was adopted due to time constraints, and several studies in the lab have shown that the recruitment kinetics of fluorescently tagged fusion proteins is not dramatically altered by the presence of any untagged endogenous protein (with the obvious exception that less of the tagged protein is invariably recruited due to the presence of the untagged protein). Nevertheless, it will be interesting in future to repeat these experiments in *cp110*/*cep97* null embryos.

I found that the intensity acquired (the difference between initial fluorescence intensity and the intensity at the oscillation peak) was significantly higher in Ubq-GFP-Cep97 embryos than in eCep97-GFP embryos (**Figure 3.13 A, C**). I also found that the peak of the oscillation occurred slightly, but significantly, later in S-phase in eCep97-GFP embryos than in Ubq-GFP-Cep97 embryos (**Figure 3.13 A, C**). Because the peak of the CP110/Cep97 oscillation correlates with S-phase length (**Figure 3.9**), I also calculated a 'corrected' time to peak by normalising the centre of the oscillation to S-phase length, to account for slight variations in S-phase length between the datasets. This correction meant

that the differences in the peak time between eCep97-GFP and Ubq-GFP-Cep97 embryos were no longer statistically significant (**Figure 3.13 E**).

I carried out a similar analysis to compare the peak of the eCP110-GFP oscillation with the Ubq-CP110-GFP peak. Again, the intensity acquired was significantly higher for Ubq-CP110-GFP than for eCP110-GFP (**Figure 3.13 B, D**). However, the difference between the centre of the peaks was not statistically significant in either the original or S-phase corrected data (**Figure 3.13 B, D, F**). There therefore did not appear to be any strikingly significant differences in the oscillation peak time between these fly lines, though there is a significant increase in the amount of CP110/Cep97 recruited to the centriole when these proteins are expressed at higher levels.

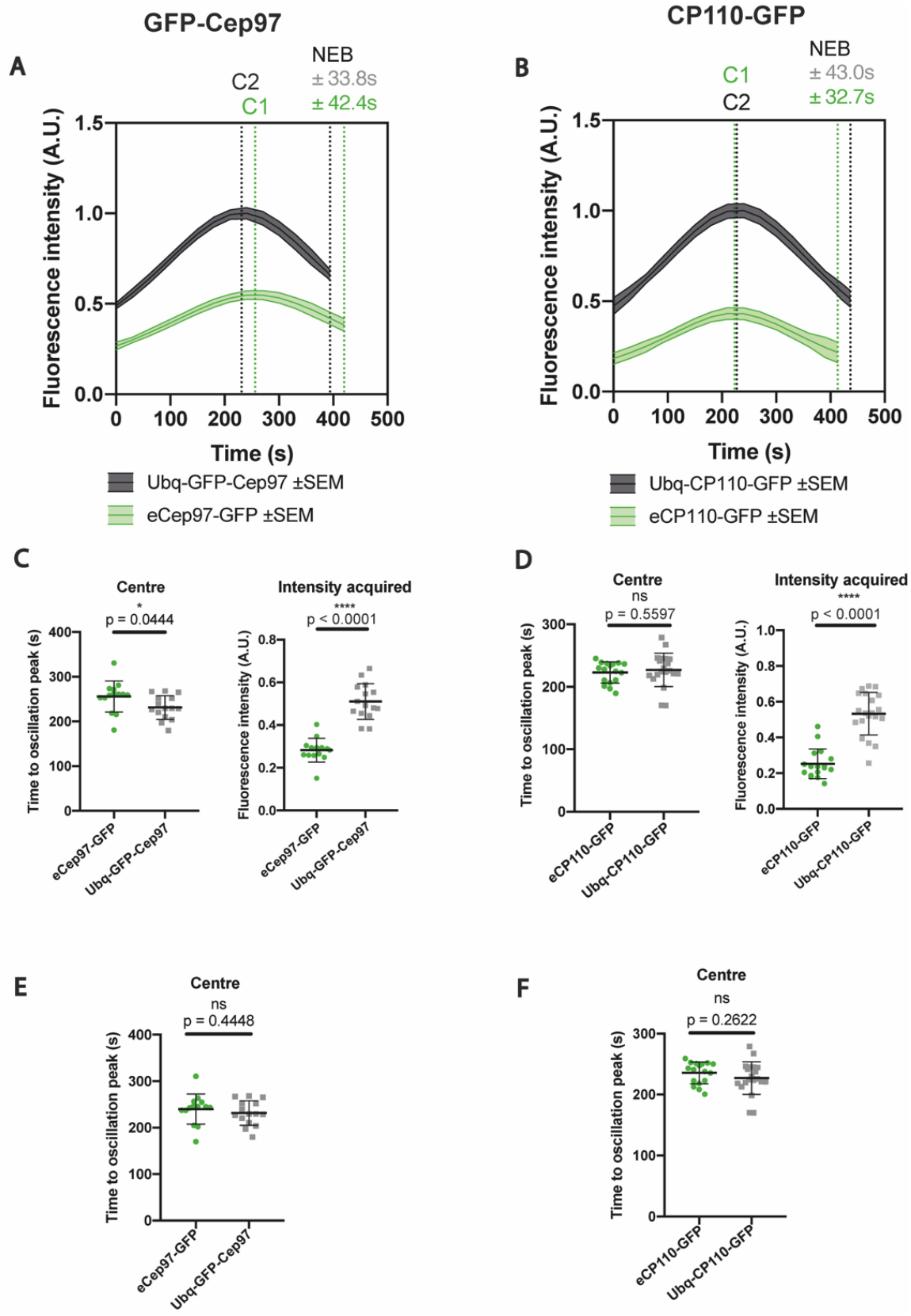

GFP-Cep97
A
NEB
± 33.8s
C2
C1
± 42.4s
Fluorescence intensity (A.U.)
1.5
1.0
0.5
0.0
0
100
200
300
400
500
Time (s)
Ubq-GFP-Cep97 ±SEM
eCep97-GFP ±SEM
CP110-GFP
B
NEB
± 43.0s
C1
C2
± 32.7s
Fluorescence intensity (A.U.)
Time (s)
Ubq-CP110-GFP ±SEM
eCP110-GFP ±SEM
C
Centre
*
p = 0.0444
Time to oscillation peak (s)
400
300
200
100
0
eCep97-GFP
Ubq-GFP-Cep97
Intensity acquired
****
p < 0.0001
Fluorescence intensity (A.U.)
0.8
0.6
0.4
0.2
0.0
eCep97-GFP
Ubq-GFP-Cep97
D
Centre
ns
p = 0.5597
Time to oscillation peak (s)
300
200
100
0
eCP110-GFP
Ubq-CP110-GFP
Intensity acquired
****
p < 0.0001
Fluorescence intensity (A.U.)
1.0
0.8
0.6
0.4
0.2
0.0
eCP110-GFP
Ubq-CP110-GFP
E
Centre
ns
p = 0.4448
Time to oscillation peak (s)
400
300
200
100
0
eCep97-GFP
Ubq-GFP-Cep97
F
Centre
ns
p = 0.2622
Time to oscillation peak (s)
300
200
100
0
eCP110-GFP
Ubq-CP110-GFP

Figure 3.13 – Comparing the recruitment dynamics of eCep97-GFP with Ubq-GFP-Cep97, and eCP110-GFP with Ubq-CP110-GFP.

(A-B) Average curves showing the Cep97 oscillation over S-phase of cycle 12 in eCep97-GFP and Ubq-GFP-Cep97 embryos (A) or the CP110 oscillation in eCP110-GFP and Ubq-CP110-GFP embryos (B). C1 indicates position of the eCep97-GFP/eCP110-GFP peak, C2 indicates position of the Ubq-GFP-Cep97/Ubq-CP110-GFP peak. NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(C-D)** Scatter plots showing quantification of time to the peak (or *Centre*) of the oscillation and of intensity acquired in either Ubq-GFP-Cep97 and eCep97-GFP (C) or Ubq-CP110-GFP and eCP110-GFP (D). **(E-F)** Centres from (C) and (D) corrected for S-phase length, to account for differences in S-phase length between the datasets. For (C-F), data are presented as mean ± SD; N ≥ 14 embryos, n ≥ 15 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

However, when I compared the time to peak of Ubq-GFP-Cep97 in either WT or Ubq-CP110 embryos (i.e. the same genotype used to investigate the effect of CP110 overexpression on MT growth in **Figure 3.5**) I found that the oscillation peaked earlier in embryos where CP110 is overexpressed; this difference was statistically significant (**Figure 3.14 A, C**), and even more so when I corrected for variations in S-phase length between the datasets (**Figure 3.14 E**). The intensity acquired was also higher in embryos expressing both Ubq-GFP-Cep97 and Ubq-CP110 than in those embryos where only Cep97 is overexpressed (**Figure 3.14 A, C**). Similarly, I found a strong effect of Cep97 overexpression on the oscillation of Ubq-CP110-GFP, again observing that the oscillation peaks earlier, and with higher amplitude, in S-phase in embryos expressing Ubq-CP110-GFP alongside Ubq-Cep97 (**Figure 3.14 B, D, F**).

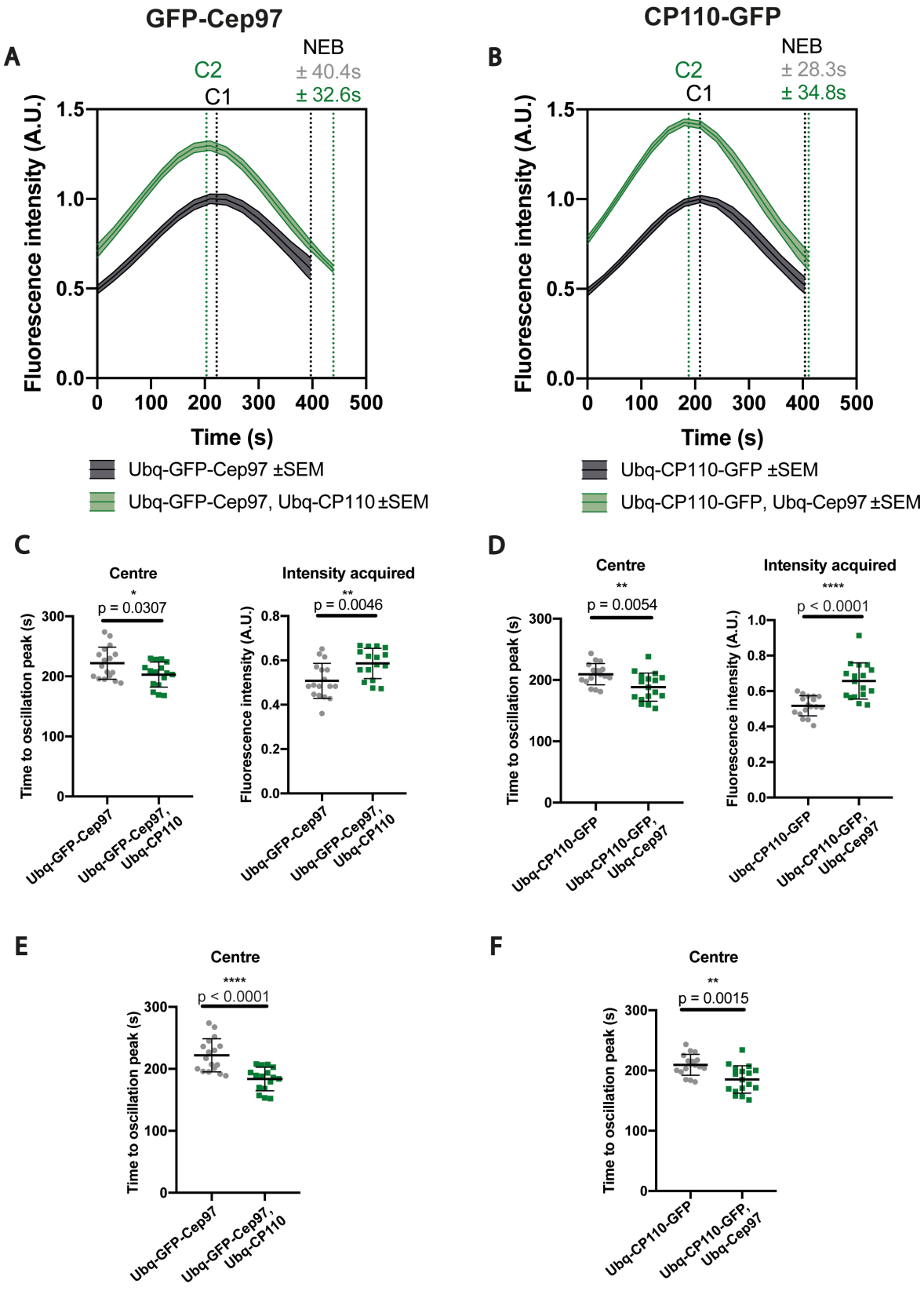

GFP-Cep97
A
NEB
± 40.4s
± 32.6s
C2
C1
Fluorescence intensity (A.U.)
Time (s)
Ubq-GFP-Cep97 ±SEM
Ubq-GFP-Cep97, Ubq-CP110 ±SEM
CP110-GFP
B
NEB
± 28.3s
± 34.8s
C2
C1
Fluorescence intensity (A.U.)
Time (s)
Ubq-CP110-GFP ±SEM
Ubq-CP110-GFP, Ubq-Cep97 ±SEM
C
Centre
p = 0.0307
Time to oscillation peak (s)
Intensity acquired
p = 0.0046
Fluorescence intensity (A.U.)
Ubq-GFP-Cep97
Ubq-GFP-Cep97, Ubq-CP110
D
Centre
p = 0.0054
Time to oscillation peak (s)
Intensity acquired
p < 0.0001
Fluorescence intensity (A.U.)
Ubq-CP110-GFP
Ubq-CP110-GFP, Ubq-Cep97
E
Centre
p < 0.0001
Time to oscillation peak (s)
Ubq-GFP-Cep97
Ubq-GFP-Cep97, Ubq-CP110
F
Centre
p = 0.0015
Time to oscillation peak (s)
Ubq-CP110-GFP
Ubq-CP110-GFP, Ubq-Cep97

Figure 3.14 – Overexpression of both CP110 and Cep97 leads to faster, and higher levels of, centriole recruitment.

(A-B) Average curves showing the Cep97 oscillation over S-phase of cycle 12 in Ubq-GFP-Cep97 embryos and Ubq-GFP-Cep97 embryos where Ubq-CP110 is also expressed (A) or the CP110 oscillation in Ubq-CP110-GFP embryos and Ubq-CP110-GFP embryos where Ubq-Cep97 is also expressed (B). C1 indicates position of the Ubq-GFP-Cep97/Ubq-CP110-GFP peak, C2 indicates position of the Ubq-GFP-Cep97/Ubq-CP110-GFP peak when Ubq-CP110/Ubq-Cep97 are also expressed. NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(C-D)** Scatter plots showing quantification of time to the peak (or *Centre*) of the GFP-Cep97 oscillation and of intensity acquired in either Ubq-GFP-Cep97 and Ubq-GFP-Cep97 + Ubq-CP110 (C) or Ubq-CP110-GFP and Ubq-CP110-GFP + Ubq-Cep97 (D). **(E-F)** Centres from (C) and (D) corrected for S-phase length, to account for differences in S-phase length between the datasets. For (C-F), data are presented as mean ± SD; N ≥ 16 embryos, n ≥ 33 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs. Imaging by M. Aydogan, analysis by the author.

Overall, these results suggest that increasing the level of *both* CP110 and Cep97 expression in embryos leads to the peak of the oscillation occurring earlier in S-phase. This is most clearly observed when comparing the Ubq-GFP-Cep97 oscillation in WT and Ubq-CP110 embryos (**Figure 3.14 E**). This dataset is also the most readily comparable to our data on MT growth dynamics (**Figure 3.5**) since the same genotype was used in both experiments. I therefore conclude that the peak of the CP110/Cep97 oscillation occurs earlier in S-phase when both CP110 and Cep97 are overexpressed. This is in the opposite direction to what I would expect if the peak of the oscillation influences the time when centriole MTs stop growing (**Figure 3.11**), since I find that the MT growth period may be slightly longer in CP110/Cep97 overexpression embryos.

I consistently found that more CP110/Cep97 was recruited to centrioles when one/both of these proteins were expressed at higher levels. This was clear when comparing the

intensity acquired in eCep97-GFP and Ubq-GFP-Cep97 embryos (**Figure 3.13 C**) and in eCP110-GFP and Ubq-CP110-GFP embryos (**Figure 3.13 D**). I observed the same effect when comparing the Ubq-GFP-Cep97 intensity acquired in WT and CP110 overexpression embryos (**Figure 3.14 C**), and when comparing the Ubq-CP110-GFP intensity in WT and Cep97 overexpression embryos (**Figure 3.14 D**). These data strongly suggest that daughter centriole MTs do not stop growing when some threshold level of CP110/Cep97 is reached; in embryos overexpressing CP110 and/or Cep97, the centrioles would always acquire such a threshold-level earlier than in the WT embryos because the proteins are recruited faster, and thus the centriole MTs should stop growing earlier in S-phase. My earlier data suggested that, if anything, the centriole MTs grew slower and for a slightly longer period when these proteins were overexpressed (**Figure 3.5**). I conclude that the CP110/Cep97 oscillation does not play an important role in determining the growth period of the centriole MTs.

Furthermore, the amplitude of the GFP-Cep97 oscillation does not correlate with the growth rate of the MTs (**Figure 3.15**), so it is also challenging to determine a clear link between the properties of the oscillation and the growth rate of the centriole MTs. Despite the absence of a statistically significant correlation, the graph in **Figure 3.15** does indicate that some of those centriole MTs that grow the slowest have accumulated higher levels of CP110/Cep97 at their growing ends. However, it is possible that higher levels of CP110/Cep97 do not directly influence the growth of the MTs; instead, higher levels of CP110/Cep97 at centriole MTs may simply reflect slightly higher global levels of these proteins (see below for discussion on the effect of altering the global levels of CP110/Cep97 on cartwheel growth). Overall, therefore, we were unable to identify any direct connection between the parameters of the CP110/Cep97 oscillation and the growth

period of the centriole MTs, although I cannot rule out the possibility that these proteins might affect the growth rate in some way.

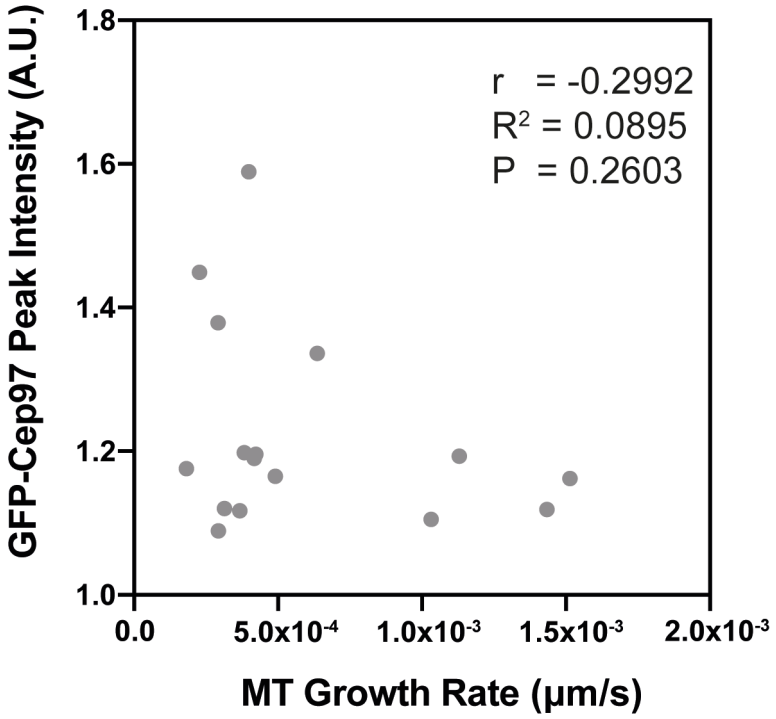


Figure 3.15 – The growth rate of the MTs does not correlate with the amplitude of the GFP-Cep97 signal.

Graph showing that there is no correlation between the growth rate of the centriole MTs (quantified in **Figure 3.5**) and the peak intensity of the GFP-Cep97 oscillation (extracted from the original images analysed in **Figure 3.5** – see **Appendix A1** for more details). N = 16 embryos. *r*, R^2 and *p*-values are shown on the graph.

3.2.5. CP110/Cep97 play a surprising role in cartwheel growth regulation

Overexpressing CP110/Cep97 slows centriole MT growth (**Figure 3.5**), but we could find no obvious connection between the CP110/Cep97 oscillation parameters and the centriole

MT growth dynamics. We wondered, therefore, whether the dosage of CP110/Cep97 might influence the growth of the centriole MTs by influencing the growth of the cartwheel. The growth dynamics of the centriole MTs strikingly resembled those of the cartwheel (**Figure 3.3**, **Figure 3.5**; Aydogan *et al.*, 2018). Plk4 is the master regulator of cartwheel growth, and reducing the levels of Plk4 in the embryo results in both the cartwheel and the centriole MTs growing at a slower rate and for a longer period (Aydogan *et al.*, 2018), indicating a relationship between the cartwheel and the MT growth processes. Intriguingly, although CP110/Cep97 and Plk4 localise to opposite ends of the growing daughter centriole (Kleylein-Sohn *et al.*, 2007), Plk4 interacts with CP110 (Galletta *et al.*, 2016a) and phosphorylates it in human cells (Lee *et al.*, 2017), while depletion of CP110 inhibits the ability of Plk4 to promote centriole over-duplication (Kleylein-Sohn *et al.*, 2007).

To test this possibility, M. Aydogan imaged Sas-6-GFP incorporation (with Sas-6-GFP fluorescence intensity serving as a proxy measure for cartwheel size) in embryos expressing different levels of either CP110 or Cep97. This data was then analysed independently by myself and by M. Aydogan. Our analyses both gave the same, surprising result: increasing the dose of CP110 or Cep97 leads to the cartwheel growth rate being lower and the cartwheel growth period being longer (**Figure 3.16**). By contrast, cartwheels in *cp110* or *cep97* null embryos grew faster for a shorter period of time (**Figure 3.16**). The overall result of this perturbation is that cartwheels end up being slightly shorter under CP110/Cep97 overexpression conditions compared to WT, and slightly longer when CP110/Cep97 are absent. The latter observation is interesting, since it is consistent with measurements taken from EM images of centrioles in *cp110* null *Drosophila* wing discs (Franz *et al.*, 2013).

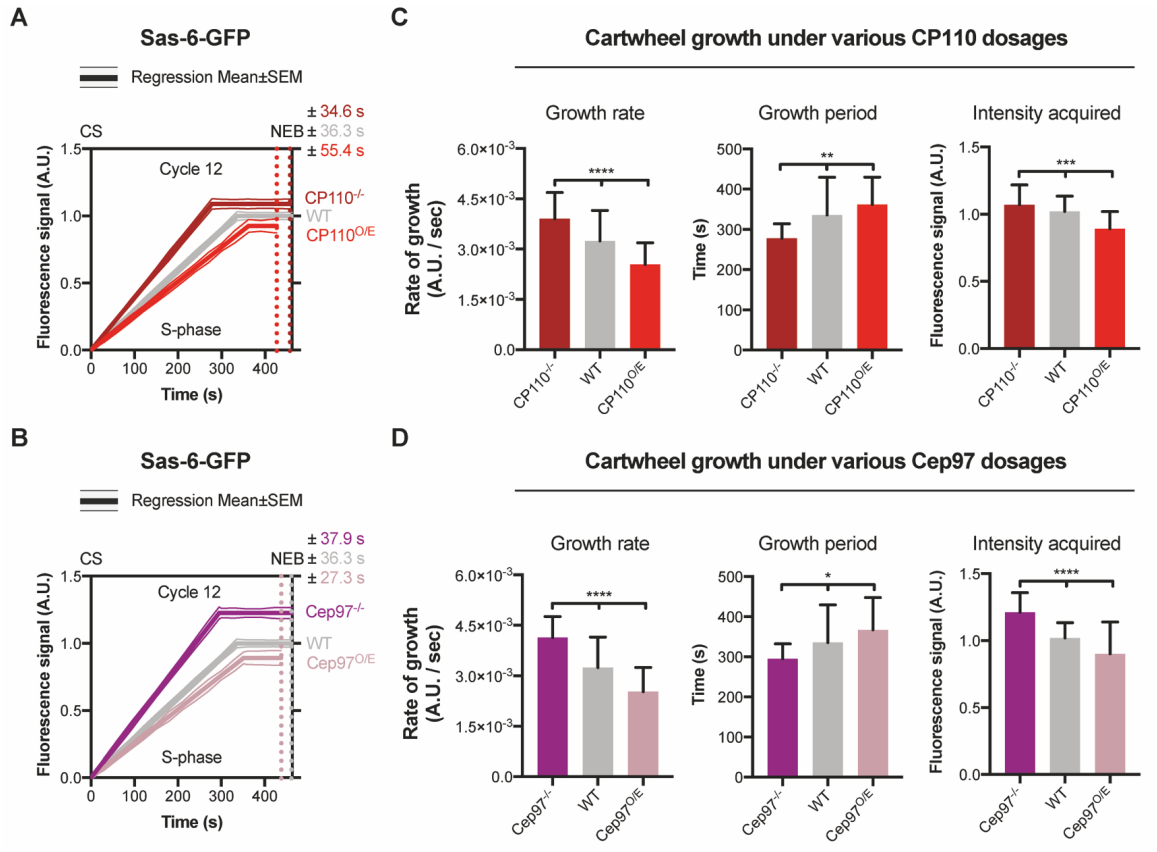


Figure 3.16 – Dynamics of cartwheel growth in embryos expressing varying doses of CP110/Cep97.

(A-B) Average curves showing the incorporation of Sas-6-GFP over S-phase of cycle 12 in embryos expressing different levels (as indicated on the graphs) of CP110 (A) or Cep97 (B). Sas-6-GFP fluorescence intensity is used here as a proxy for cartwheel size. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(C-D)** Bar charts showing quantification of cartwheel growth rate, cartwheel growth period and final cartwheel size (given by the proxy measurement of 'intensity acquired') in embryos expressing different levels of CP110 (C) or Cep97 (D). Data are presented as mean ± SD; N ≥ 14 embryos, n ≥ 40 centrioles per embryo. ANOVA tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are indicated on the graphs: * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$, **** represents $p<0.0001$. Imaging by Mustafa Aydogan, analysis by the author and M. Aydogan. Final figure produced by M. Aydogan for Hankins, Aydogan *et al.* (in preparation).

3.2.6. CP110/Cep97 may interact with Plk4 at the centriole

Since the dynamics of cartwheel growth are controlled by the oscillatory recruitment of Plk4 to the centriole (Aydogan *et al.*, 2020), we next investigated whether the dose of CP110/Cep97 might influence the dynamics of Plk4 recruitment. Remarkably, experiments by M. Aydogan revealed that the amplitude of the Plk4 oscillation increased in the absence of CP110 or Cep97, and decreased when either protein was overexpressed (**Figure 3.17**). This suggests that the CP110/Cep97 complex (which is concentrated at the distal-end of the growing daughter) is somehow able to influence the recruitment of Plk4 at the proximal-end. This may explain why altering the dosage of CP110/Cep97 affects the growth parameters of the cartwheel.

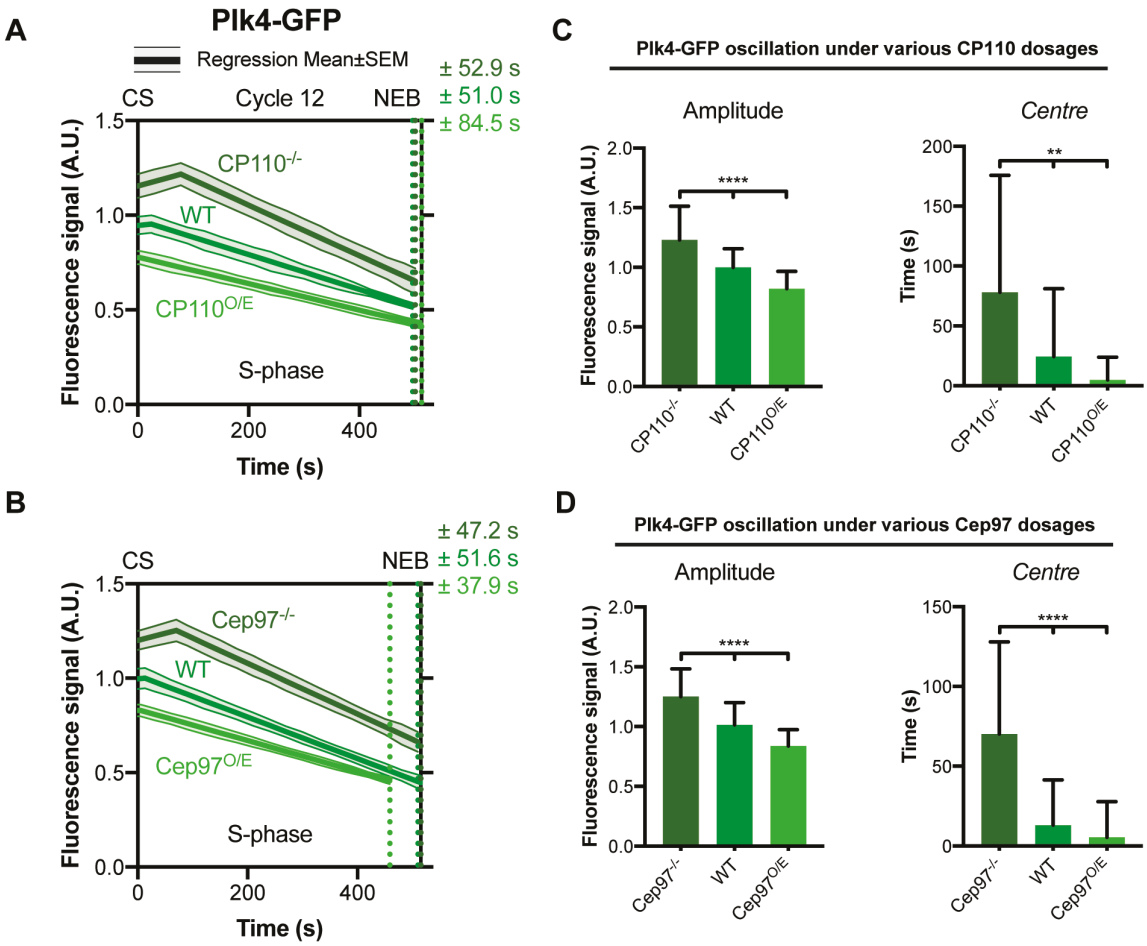


Figure 3.17– Dynamics of Plk4 centriolar recruitment in embryos expressing varying doses of CP110/Cep97.

(A-B) Average curves comparing the recruitment of Plk4-GFP to centrioles over S-phase of cycle 12 in embryos expressing different levels (as indicated on the graphs) of CP110 (A) or Cep97 (B). CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(C-D)** Bar charts showing quantification of the amplitude and time to peak (or *Centre*) of the Plk4 oscillation in embryos expressing different levels of CP110 (C) or Cep97 (D). Data are presented as mean ± SD; N ≥ 16 embryos, n ≥ 45 centrioles per embryo. ANOVA tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are indicated on the graphs: ** represents $p<0.01$, **** represents $p<0.0001$. Figure by M. Aydogan for Hankins, Aydogan *et al.* (in preparation).

Since CP110/Cep97 and Plk4 localise to opposite ends of the growing daughter centriole (Kleylein-Sohn *et al.*, 2007), we wondered whether these proteins might instead be interacting in the cytoplasm. I therefore performed western blots to assess the global level of CP110/Cep97 expression under different Plk4 conditions. I compared WT embryos with those expressing a reduced level of Plk4. I also looked at embryos expressing a copy of Plk4 with reduced kinase activity; for this sample, my control was WT embryos expressing an untagged copy of Plk4 driven by its endogenous promoter (Aydogan *et al.*, 2018). I found no statistically significant difference in CP110/Cep97 levels under the various Plk4 conditions tested (**Figure 3.18**). These results are consistent with measurements of CP110/Cep97 cytoplasmic concentrations made via Fluorescence Correlation Spectroscopy (FCS; Thomas Steinacker, personal communication). Plk4 is present at very low levels in embryos, and it is very difficult to detect in western blotting experiments. Thomas Steinacker (Raff Lab, University of Oxford) therefore used Peak Counting Spectroscopy (Aydogan *et al.*, 2020), to measure the cytoplasmic concentration of Plk4 under different doses of CP110/Cep97, and found that this remained unchanged (T. Steinacker, personal communication). Thus, cytoplasmic Plk4 levels do not appear to affect the cytoplasmic levels of CP110/Cep97, and vice versa.

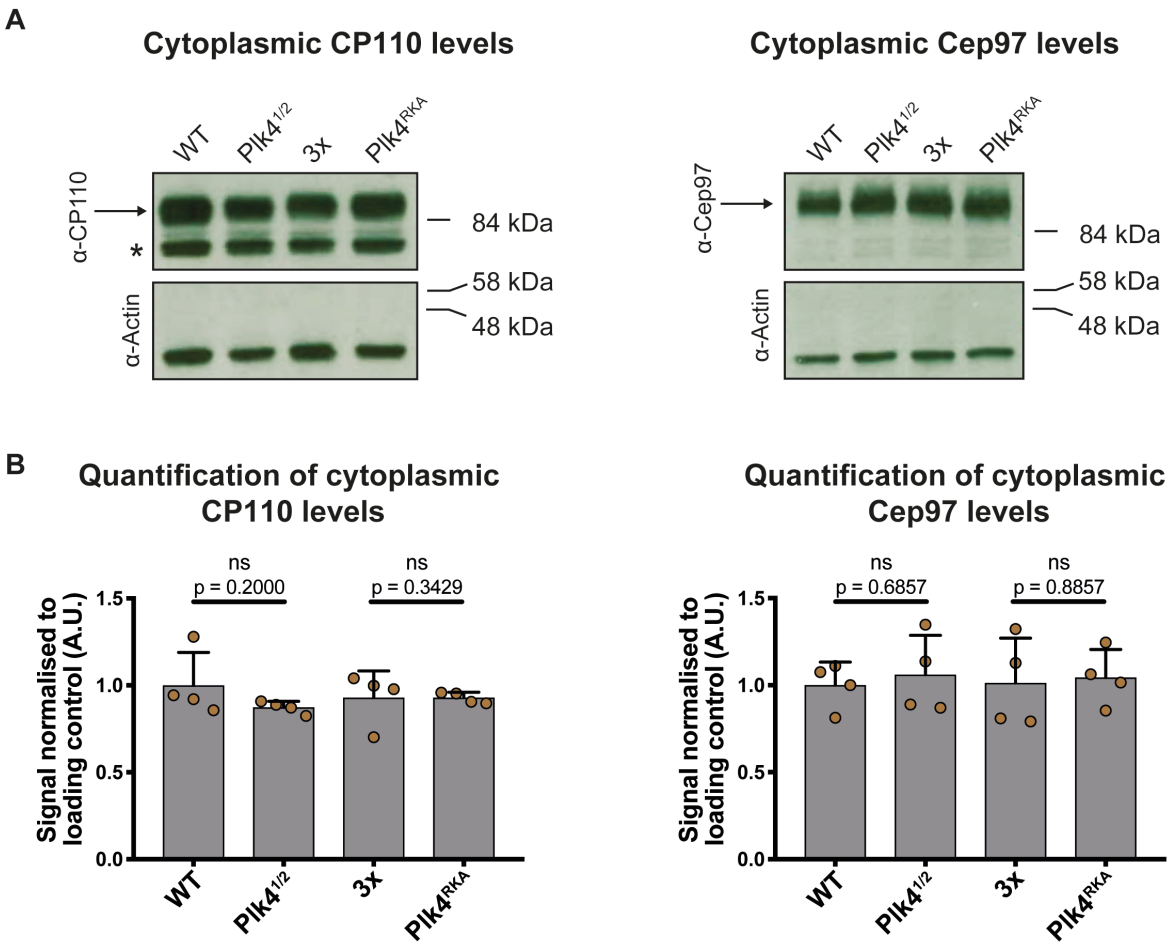


Figure 3.18 – Western blots comparing CP110/Cep97 expression in embryos under different Plk4 conditions.

(A) Western blots comparing the level of endogenous CP110 or Cep97 expression in embryos under different Plk4 conditions: (1) WT embryos (*WT*); (2) embryos heterozygous for the Plk4 mutation (*Plk4*$^{1/2}$); (3) WT embryos expressing an untagged copy of Plk4 driven by its endogenous promoter (*3x*); (4) WT embryos expressing an untagged copy of Plk4 with reduced kinase activity (*Plk4*^{RKA}). CP110 and Cep97 bands are indicated with arrows. Non-specific bands are indicated with *. Actin is shown as a loading control. **(B)** Bar charts showing quantification of the levels of CP110 or Cep97 (normalised to their respective actin loading control levels). Data are presented as mean ± SD; N = 4 (from four technical repeats of the blots in (A)). T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

It was also possible that CP110/Cep97 might interact with Plk4 in the cytoplasm in a way that does not affect the overall levels of these proteins. For example, CP110/Cep97 might be phosphorylated by Plk4. I therefore decided to also carry out an immunoprecipitation (IP) experiment to test for the existence of such an interaction (**Figure 3.19**). I collected embryos from flies lacking endogenous Plk4 and expressing Plk4-GFP under its endogenous promoter. I also collected WT embryos to act as a control. I produced lysates from these samples; this represented the '**input**' sample. I used an α-GFP antibody bound onto beads to carry out the pulldown, and collected samples of the '**unbound**' fraction that passed through the column. I then eluted the sample that had bound to the beads by boiling in sample buffer and collected this (the '**bound**' sample). Western blotting with these samples demonstrated that I had successfully pulled down Plk4-GFP (**Figure 3.19**). Probing for CP110 and Cep97 then allowed me to compare whether levels of these proteins were enriched in the Plk4-GFP 'bound' sample compared with the WT control. I did not see any enrichment here, suggesting that I had not pulled down more CP110/Cep97 along with the Plk4-GFP. There appears to have been an issue of protein stickiness on the beads, however, since I also detected protein in the 'bound' lane of the WT sample. This makes the results more difficult to interpret. It is also possible that any potential interaction between CP110/Cep97 and Plk4 might be transient, which would make it less likely to be detectable via IP. Therefore, these results do not exclude the possibility that there is an interaction between CP110/Cep97 although, given our overall results, I think it is likely that this occurs at the centriole rather than in the cytoplasm.

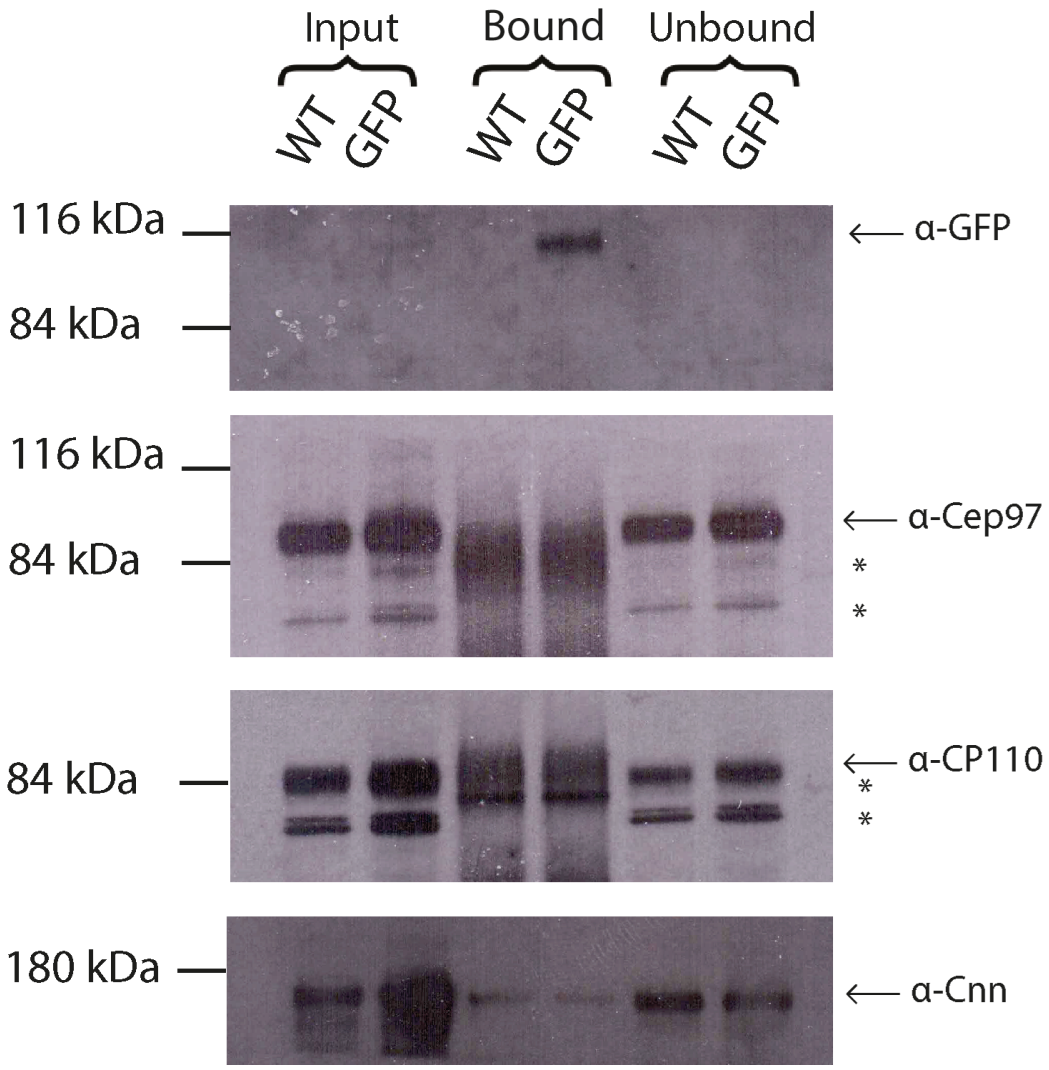


Figure 3.19 – Immunoprecipitation assay using embryo extract to test for an interaction between Plk4 and CP110/Cep97.

Western blot showing the result of an immunoprecipitation experiment. The lanes corresponding to embryos expressing Plk4-GFP are indicated with 'GFP'. An α-GFP antibody was used for the pull down. WT embryos were used as a control, and the lanes corresponding to these control samples are indicated with 'WT'. The input, bound, and unbound fraction lanes are indicated with labels above the blot. The bands for GFP, Cep97, CP110 and Cnn are indicated with arrows. Non-specific bands are indicated with *. As expected, the GFP band was only visible in the Plk4-GFP bound lane. No enrichment of either the CP110 or Cep97 signal was observed in the bound lane when comparing the Plk4-GFP lane with the WT control lane. Cnn is shown as a loading control. N = 1.

3.3. Discussion

In this Chapter, I have demonstrated that CP110/Cep97 are recruited to centrioles in living *Drosophila* embryos, and that they are required for each other's recruitment. This is consistent with reports from other organisms (Kleylein-Sohn *et al.*, 2007, Sonnen *et al.*, 2012, Spektor *et al.*, 2007). This complex was known to play a role in centriole length regulation (Kohlmaier *et al.*, 2009, Schmidt *et al.*, 2009, Tang *et al.*, 2009), with a particularly important role in controlling centriole MT length in flies (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013). However, imaging centriole MT growth in living systems is challenging, so it was not clear how these proteins influence centriole size, or whether they play a role during the process of centriole MT growth. Through our development of an automated assay for the dynamic tracking of centriole MT growth, we have shown that overexpression of CP110/Cep97 in *Drosophila* embryos results in slower centriole MT growth, possibly for a slightly longer period. These centriole MTs appear to acquire a shorter final length than their WT counterparts.

We also found that CP110/Cep97 levels oscillate at the centriole, with levels peaking in approximately mid-S-phase. During my DPhil, a study was published suggesting that Cep97-GFP does not exhibit oscillatory recruitment when expressed at endogenous levels in *Drosophila* embryos, and that it is only recruited to new daughter centrioles once they have finished growing (Dobbelaere *et al.*, 2020). However, my imaging experiments with this same fly line have revealed that the oscillation is maintained in this genotype, with eCep97-GFP levels at the centriole increasing from early S-phase (**Figure 3.13 A**). A possible reason for this discrepancy is that the published study used fluorescence recovery after photobleaching (FRAP) to infer protein dynamics, and it is challenging to generate

high *n* numbers using this analysis. By contrast, my experiments investigated the fluorescence intensities of hundreds of centriole pairs, from multiple embryos. I am therefore confident of my conclusion that CP110 and Cep97 are recruited to the growing daughter centriole in an oscillatory manner.

In contrast to the example of the Plk4 oscillation's role in determining cartwheel growth dynamics, the properties of the CP110/Cep97 oscillation do not appear to directly influence the growth rate or growth period of the centriole MTs. Instead, we found that perturbing CP110/Cep97 levels in the embryo can affect the growth rate and growth period of the cartwheel structure. For example, in *cp110* null embryos, we found that the cartwheel grew faster and for a shorter period of time than in WT embryos. Its final length was also greater than the cartwheels of WT embryos. This result is particularly interesting given that EM data from Franz *et al.* 2013 suggests that the cartwheel structure is slightly, but significantly, longer in *cp110* null embryos than in WT embryos.

This effect on cartwheel growth dynamics is possibly mediated via an interaction between CP110/Cep97 and Plk4. It seems most likely, given our results, that such an interaction occurs at the centriole. For example, I was unable to detect an interaction in cell extract using immunoprecipitation; however, any interaction might be too transient to detect via this approach. CP110 is a Plk4 substrate in human cells (Lee *et al.*, 2017), and it would therefore be interesting to test for an interaction between Plk4 and CP110/Cep97 using an *in vitro* kinase assay to see if Plk4 phosphorylates CP110/Cep97. This assay has the added advantage over immunoprecipitation of being able to test for direct interactions, since any interaction detected within a cell lysate might be mediated via intermediate proteins.

However, attempting a kinase assay would be challenging, since it would involve purification of full-length functional CP110/Cep97 and Plk4.

Interestingly, we found that cartwheels grew slower and for a longer period when CP110/Cep97 were overexpressed. Similarly, centriole MTs grew slower (and possibly for a longer period) when CP110/Cep97 were overexpressed. Given this observation, and considering that the CP110/Cep97 oscillation does not appear to directly influence centriole MT growth, it seems possible that the real driver of centriole MT growth could be the growth of the underlying cartwheel structure. The centriole MTs are physically linked to the cartwheel by the pinhead (Guichard *et al.*, 2013) and it is possible that these MTs need to be stabilised by their association with the cartwheel as they grow (Hirono, 2014). If this is the case, and the cartwheel drives the incorporation of the centriole MTs, then any changes in cartwheel growth rate or period might be expected to translate to the centriole MTs. Thus, the effect we see on centriole MT growth when CP110/Cep97 are overexpressed might be mediated via CP110/Cep97 interacting with Plk4 to influence cartwheel growth (and thus MT growth) rather than by a direct effect of CP110/Cep97 upon the MTs. Alternatively, it could be that CP110/Cep97 are also binding to the cartwheel structure, so measurements of GFP-Cep97 distancing reflect the growth of the cartwheel structure as well as that of the centriole MTs. To test if the cartwheel really drives MT growth, it would be interesting to investigate whether certain perturbations that affect cartwheel growth (such as the dose of Ana2; M. Aydogan, personal communication) also consistently affect centriole MTs in the same way, and vice versa. Finally, it is also possible that there is some form of 'mechanical' interaction between the proximal- and distal-ends of the centriole. This idea is further explored in **Chapter 7**.

It was also possible that the increase in cartwheel growth rate seen in *cp110*/*cep97* null embryos could be due to the loss of the distal-end 'cap', thus potentially allowing Sas-6-GFP to incorporate from the distal- as well as the proximal-end. However, FRAP experiments have shown that Sas-6-GFP incorporation is still proximal and exclusive to the daughter in *cp110/cep97* null embryos (M. Aydogan and A. Wainman, personal communication). Furthermore, this would not explain why the growth period of these cartwheels is reduced, and does not address the effect on cartwheel dynamics that we observed in embryos overexpressing CP110/Cep97.

My results suggest that the CP110/Cep97 oscillation does not time the growth period of the centriole MTs. I cannot rule out the possibility that CP110/Cep97 might directly influence the growth rate of centriole MTs, though I have not found any strong evidence for this. However, we do know that loss of these proteins does cause dramatic extensions of the centriole MTs in other tissues (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013). It is tempting to postulate that CP110/Cep97 might perform two separate roles:

(1) Capping the MTs of grown centrioles and preventing them from over-elongating.
(2) Influencing the growth of the cartwheel during centriole biogenesis in S-phase.

I suggest that CP110/Cep97's role in capping MTs might be particularly important in cells that undergo a second round of centriole elongation during G_2-phase of the cell cycle (Tates, 1971). I have imaged Plk4-mNeonGreen in cycle 14, when cellularisation begins in the *Drosophila* embryo and cycles start to acquire G_2-phases (Edgar and O'Farrell, 1990). In this cycle, I observed that Plk4 oscillates, peaking around 10 minutes into S-phase before decreasing again (**Figure 3.20 A**). I have also looked at Ubq-CP110-GFP and Ubq-

GFP-Cep97 recruitment dynamics in cycle 14. By contrast, these proteins appear to be recruited to the centrioles and then remain there at high levels (**Figure 3.20 B, C**). It is possible that, if I could image for longer (the centrioles begin to move out of the plane of view after a certain point in the cycle, as cellularisation takes place) then I would see these levels decrease again, since we know that they usually peak later in S-phase than Plk4 does. However, it is also possible that these proteins are being recruited to centrioles and remaining there at high levels. Wing disc cells also have a G_2-phase, though they normally do not undergo a dramatic elongation during this process; in flies, this secondary phase of elongation is most prominently observed in spermatocyte centrioles (Tates, 1971) and CP110 is removed from spermatocyte centrioles at elongation onset (Franz *et al.*, 2013, Shoda *et al.*, 2021). One possibility is that wing disc centriole MTs are usually capped by CP110/Cep97, thus preventing their extension, and that the loss of these proteins in *cp110* or *cep97* null wing discs then leads to MT over-elongation.

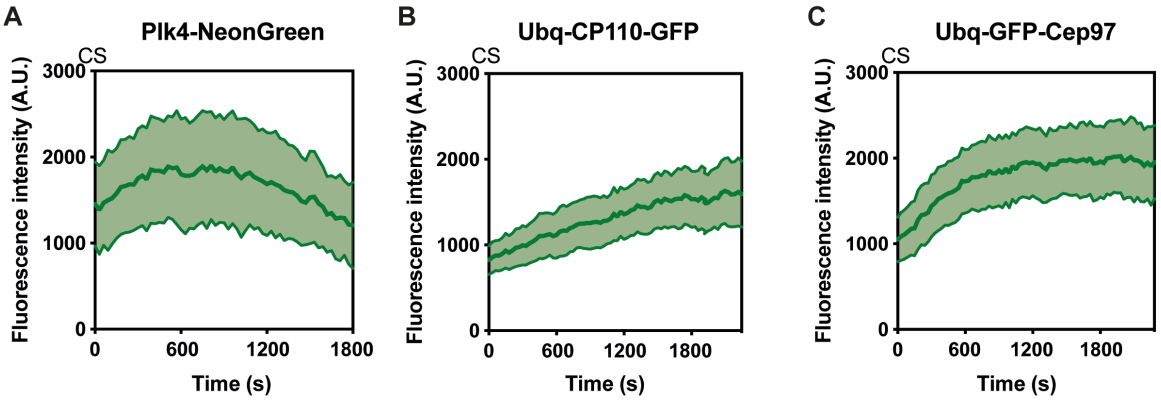


Figure 3.20 – Centriolar recruitment of Plk4, CP110 and Cep97 in nuclear cycle 14.

(A-C) Graphs showing quantification of ePlk4-mNeonGreen (A), Ubq-CP110-GFP (B), or Ubq-GFP-Cep97 (C) fluorescence intensity at centrosomes in a single embryo during S-phase of cycle 14. CS = centrosome separation (marking the start of S-phase). Data are presented as mean ± SD.

There are two isoforms of CP110 in *Drosophila,* a short form and a long form, and only the short form is expressed in testes (Franz *et al.*, 2013). By contrast, wing discs and embryos only express the long form (Franz *et al.*, 2013). It has been suggested that the short isoform of CP110 is unable to cap MTs, since the long isoform fully rescues the *cp110* null phenotype of over-elongated centriole MTs, whereas expression of the short isoform in *cp110* null wing discs does not fully prevent the formation of these long structures (Franz *et al.*, 2013). This distinction could explain why spermatocyte centrioles undergo a second elongation during G_2 while wing disc centrioles do not.

Crucially, it is not clear whether the centrioles observed by EM in Franz *et al.* 2013, or in Dobbelaere *et al.* 2020, are mothers, daughters or both. Hence it is unclear whether these elongated projections form during centriole biogenesis or later, on mature mothers. Unfortunately, it is not possible to carry out the equivalent experiment using our GFP-Cep97 distancing algorithm, since *cp110* null embryos do not recruit GFP-Cep97. Otherwise, it would be possible to check if daughters in these embryos develop overly long MTs as they grow. If so, this assay would also indicate how long these extensions are.

Finally, this Chapter also raises some interesting questions about how CP110 and Cep97 are recruited to the centriole. This information would offer insight into how these proteins might be interacting with the centriole MTs, and how they could interact with Plk4. This question of how CP110/Cep97 might localise at the centriole is a major subject of the following thesis chapters.

Chapter 4: Exploring the CP110/Cep97 Interaction Network – Cep104

4.1. Introduction

4.1.1. The mechanism by which CP110/Cep97 are recruited to centrioles is unknown

In the previous chapter, I presented evidence that CP110 and Cep97 form a functional complex at the distal-end of centrioles, and that their levels oscillate over the course of the cell cycle. This oscillation does not appear to be important for timing centriole MT growth; however, it seems that CP110/Cep97 play an unexpected role in influencing the growth parameters of the central cartwheel structure, potentially via an interaction with the kinase Plk4.

Our work suggests that such an interaction may take place at the centriole, rather than in the cytoplasm. However, the means by which such an interaction with Plk4 might be mediated remain unclear. A previous study showed that human Plk4 can phosphorylate CP110, yet this phosphorylation site is not conserved in flies (Lee *et al.*, 2017). Furthermore, it is not known how CP110/Cep97 are even recruited to the centriole in the first place. Improving our understanding of this process would be an important starting point to further investigate both how CP110/Cep97 might interact with Plk4, as well as how CP110/Cep97 interact with the centriole MTs.

4.1.2. CP110/Cep97 are implicated in MT length control yet have no known MT-binding domains

Understanding the mechanism of CP110/Cep97 recruitment to the centriole should provide an insight into how these proteins bind to centriole MTs. CP110/Cep97 are thought to act as a 'capping complex' to prevent the over-elongation of centriole MTs (Kohlmaier *et al.*, 2009, Schmidt *et al.*, 2009, Tang *et al.*, 2009). They are known to be required for each other's recruitment in human cells (Spektor *et al.*, 2007) and in the previous Chapter I demonstrated that this is also the case in early *Drosophila* embryos. However, the tertiary structure of these proteins is unknown, and there is no evidence that CP110/Cep97 can bind directly to MTs (Rezabkova *et al.*, 2016). It is therefore suspected that these proteins might be recruited to the centriole MTs via an interaction partner (Jiang *et al.*, 2012, Rezabkova *et al.*, 2016, Satish Tammana *et al.*, 2013).

4.1.3. CP110/Cep97 interact with other distal-end-binding proteins, including Cep104

CP110 and Cep97 have been identified as the interaction partners of several other proteins that localise to the distal-ends of centrioles, including Cep104 (Al-Jassar *et al.*, 2017, Jiang *et al.*, 2012, Rezabkova *et al.*, 2016). Cep104 co-localises with the CP110/Cep97 complex on both the mother and daughter centrioles (Frikstad *et al.*, 2019, Gilani, 2018, Jakobsen *et al.*, 2011, Jiang *et al.*, 2012). It is strongly implicated in ciliogenesis, tracking the growing ciliary tip during axoneme extension (Satish Tammana *et al.*, 2013), and depletion of Cep104 results in reduced cilia/flagella formation (Frikstad *et al.*, 2019, Jiang *et al.*, 2012,

Louka *et al.*, 2018, Satish Tammana *et al.*, 2013). There is growing evidence that Cep104 promotes axoneme extension via its tubulin-binding TOG domain, which may facilitate tubulin subunit incorporation in a similar manner to XMAP215 and CLASP proteins (Akhmanova and Steinmetz, 2015, Al-Bassam and Chang, 2011, Al-Jassar *et al.*, 2017, Louka *et al.*, 2018, Rezabkova *et al.*, 2016, Yamazoe *et al.*, 2020).

The removal of CP110/Cep97 from the centriole appears to be an important step that licences the mother centriole for ciliogenesis (Spektor *et al.*, 2007). Cep104 does not seem to be required for this removal stage, since human RPE cells depleted of Cep104 fail to generate cilia but are still able to remove the capping complex (Frikstad *et al.*, 2019, Yamazoe *et al.*, 2020). This further supports the idea that Cep104 largely functions in ciliogenesis by promoting axoneme extension. However, overexpression of Cep104 can promote recruitment of Cep97 to MTs (Jiang *et al.*, 2012), so it has been suggested that Cep104 might act to recruit or stabilise CP110/Cep97 at the mother centriole until something triggers their release at the onset of ciliogenesis (Jiang *et al.*, 2012, Rezabkova *et al.*, 2016, Satish Tammana *et al.*, 2013). This could explain how CP110/Cep97 are successfully recruited to the growing ends of centriole MTs, since Cep104 possesses an SxIP motif which interacts with EB1, a protein that tracks growing MT ends (Jiang *et al.*, 2012, Maurer *et al.*, 2011, Rezabkova *et al.*, 2016, Tirnauer and Bierer, 2000). I was therefore curious as to whether Cep104 might play a role in CP110/Cep97's recruitment to the centriole, perhaps acting as an intermediate to facilitate the capping complex's interaction with centriole MTs.

Finally, I was also curious as to whether other distal-end-binding proteins might also play a role in centriole growth regulation during centriole biogenesis, since we have uncovered a surprising role for CP110/Cep97 in cartwheel growth regulation (**Chapter 3**).

4.1.4. Key findings in this chapter

In this Chapter, I begin by demonstrating that Cep104 is recruited to *Drosophila* centrioles. Next, I describe the generation of a *cep104* mutant using CRISPR/Cas9 technology. Using this mutant, I demonstrate that CP110/Cep97 are still recruited to the centriole even in the absence of full-length Cep104. Additionally, I show that CP110/Cep97 still maintain their oscillatory behaviour at the centriole in the absence of Cep104, and that the parameters of this oscillation are undisturbed. I demonstrate that Cep104 itself is also recruited to centrioles in an oscillatory manner. This recruitment is dramatically reduced when CP110/Cep97 are absent. This is interesting since it suggests that Cep104 might be acting downstream of the distal cap proteins, or might be performing an independent function. Finally, I show that Cep104 levels in the embryo appear to have little effect on the growth of the daughter centriole during centriole biogenesis, in contrast to CP110/Cep97.

4.2. Results

4.2.1. Cep104 localises to *Drosophila* centrioles

First, I confirmed that Cep104 is recruited to centrioles in the early *Drosophila* embryo by generating flies expressing Ubq-Cep104-GFP and Ubq-RFP-Cep97 (**Figure 4.1 A**). I observed that the Cep104-GFP signal co-localised with RFP-Cep97 for the duration of the cell cycle, which confirmed that Cep104 does indeed localise to the centrioles, as previously reported in other organisms (Frikstad *et al.*, 2019, Jakobsen *et al.*, 2011, Jiang *et al.*, 2012, Louka *et al.*, 2018, Satish Tammana *et al.*, 2013). Furthermore, the signal appeared to grow brighter as S-phase progressed.

To investigate the centriolar localisation of Cep104 at higher resolution, I produced wing disc squashes from third-instar larvae of a *Drosophila* line expressing Ubq-Cep104-GFP. These squashes were stained using an α-GFP nanobody (to enhance the Cep104-GFP fluorescence signal) and an α-Asl antibody. The latter would recognise Asl, a known marker of mother centrioles (Novak *et al.*, 2014). These wing discs were then imaged using 3D-SIM super-resolution microscopy. I observed that Cep104-GFP forms a ring within the Asl-mCherry signal, demonstrating that Cep104 localises to the mother centriole (**Figure 4.1 B**). In addition, some Cep104 localises as a dot in the vicinity of the mother centriole, which suggests that Cep104 is also localised to the daughter centriole. The ring-like localisation of Cep104 on the mother centriole was strikingly reminiscent of the ring formed by CP110/Cep97 (**Figure 3.1**).

I then passed these images through a script that can measure the radius of this GFP ring. This script has been published (Gartenmann *et al.*, 2017) and subsequently modified by Isaac Wong (Raff Lab, University of Oxford). It first uses the ring formed by the Asl signal to identify mother centrioles that are oriented directly towards the microscope objective, before taking measurements of the GFP ring's diameter. Unfortunately, the script was unable to identify any rings that were appropriate for measurement, most likely due to the background signal in the images being too high. It would therefore be beneficial to try to troubleshoot this experiment in the future. For example, the signal:noise ratio might be improved by reducing the antibody concentration and incubating for a longer period of time at 4°C, or blocking using sera from the antibody source species (Ramos-Vara, 2005). Increasing the number of washes or modifying the pH of the buffers used might also help reduce the background signal. However, it is interesting to observe that Cep104 seems to be localising at the distal-end of centrioles as a ring in a similar manner to CP110/Cep97.

Finally, I confirmed that Cep104 also localises to both mother and daughter centrioles in *Drosophila* early embryos by imaging embryos expressing Ubq-Cep104-GFP using a Zeiss 880 LSM microscope fitted with Airyscan detector. Airyscan Processing allowed me to resolve the Cep104 signal as two puncta, which correspond to the mother and daughter (**Figure 4.1 C**). This preliminary analysis revealed that the centriole pairs often exhibited an asymmetry in Cep104-GFP fluorescence intensity. It would be interesting to use an Asl marker to investigate whether the brighter signal corresponds to the mother or the daughter, as I did with CP110/Cep97 in **Chapter 3**.

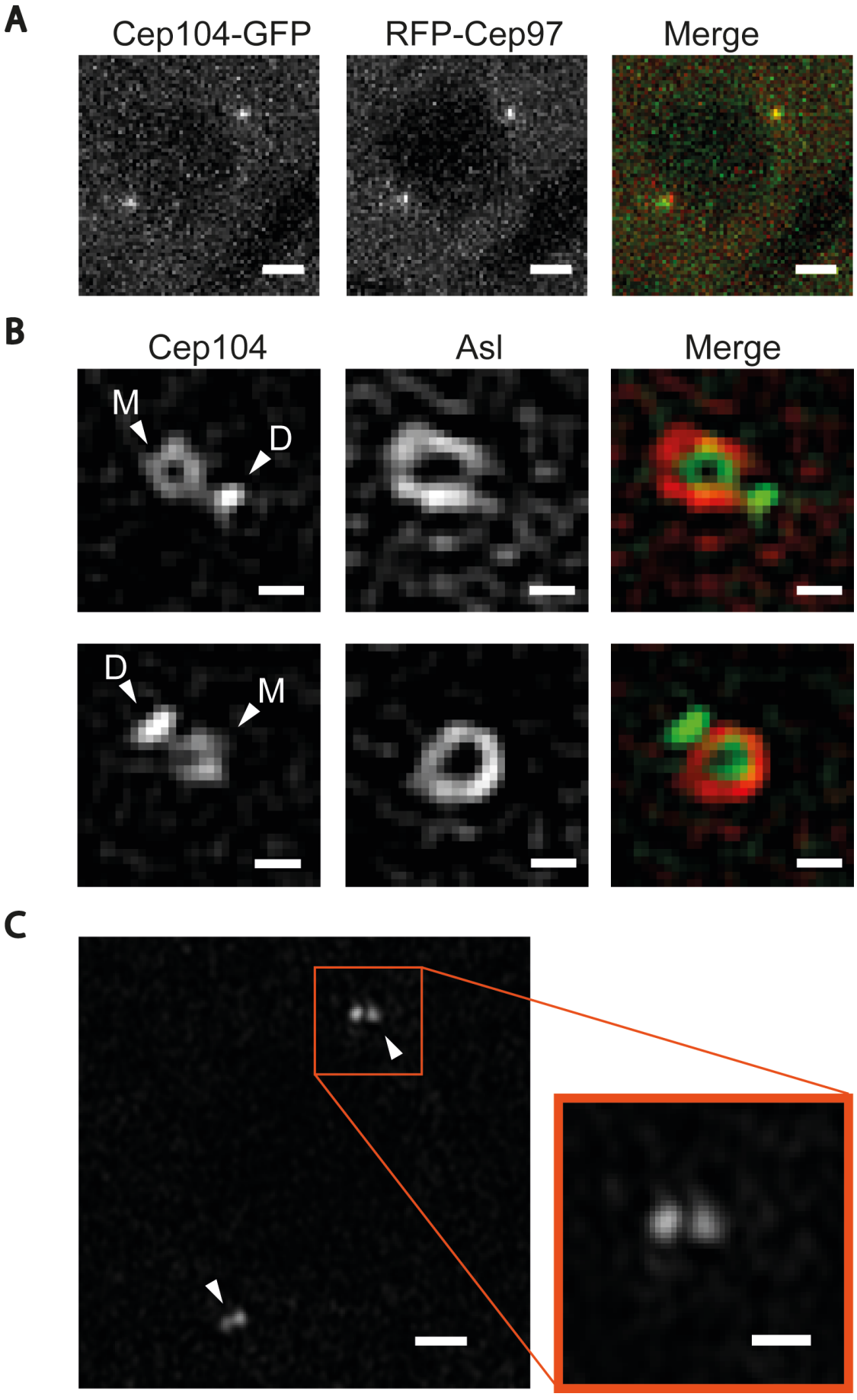

A
Cep104-GFP
RFP-Cep97
Merge
B
Cep104
Asl
Merge
M
D
D
M
C

Figure 4.1 – Cep104 is recruited to both mother and daughter centrioles in *Drosophila*.

(A) Spinning disc confocal images showing Ubq-Cep104-GFP localisation in cycle 12 *Drosophila* embryos. The Cep104-GFP signal co-localises with centriole protein Ubq-RFP-Cep97, confirming that Cep104 localises to centrioles. Scale bar represents 2.5 μm. **(B)** 3D-SIM images showing centrioles in fixed wing disc preparations, stained for Asl (a marker of the mother centriole; Novak *et al.* 2014) and Cep104-GFP. Cep104 localises to both mother (M) and daughter (D) centrioles. The Cep104-GFP signal on the mother centriole forms a ring inside the larger Asl ring. Scale bar represents 250 nm. These micrographs are presented as best-focussed slices rather than maximum intensity projections. **(C)** Zeiss 880 LSM Airy-processed images showing Cep104-GFP localisation in cycle 12 *Drosophila* embryos. Centriole pairs are indicated with arrow heads. Inset (orange border) shows one of the pairs in greater detail. The Cep104 signal can be resolved as two puncta, supporting the observation in (B) that Cep104 appears to localise to both the mother and daughter centriole. There may be an asymmetry in the fluorescence intensity of mother and daughter. Scale bar in main image represents 1 μm, scale bar in inset represents 0.5 μm.

4.2.2. CP110/Cep97 are still recruited to centrioles in the absence of Cep104

4.2.2.1. Generation of a *cep104* CRISPR knockout *Drosophila* line

Unlike the human homologue, Cep104 in *Drosophila* does not contain an SxIP motif (Al-Jassar *et al.*, 2017), a sequence that might interact with EB1 to facilitate human Cep104's recruitment to growing MT ends (Jiang *et al.*, 2012, Rezabkova *et al.*, 2016). Furthermore, immunofluorescence experiments have not detected EB1 at the centrioles in *Drosophila* (Saroj Saurya, personal communication). However, the fact that Cep104 still localised to centrioles in both early fly embryos and wing discs in a similar manner to CP110/Cep97 suggested that it might somehow still be localising to the centriole MTs, perhaps through an alternative domain or via another protein intermediate. I was therefore interested in investigating whether the Cep104 that I see at the centriole might be playing a role in CP110/Cep97 recruitment.

To do this, I generated a fly line where Cep104 is genetically ablated using CRISPR/Cas9 technology (Bassett *et al.*, 2014). I first mapped the gene region of *cep104* using data from Flybase (Larkin *et al.*, 2020). During this process, I found that there was some overlap between the region encoding the N-terminal half of Cep104 and the gene *CG10165* (Larkin *et al.*, 2020). Disrupting this other gene could potentially cause unintended effects, so I decided to design gRNAs that would target the last 5 exons at the C-terminus of the *cep104* gene, since these exons do not overlap with other known genes (**Figure 4.2 A**). I also sequenced the *cep104* gene region in the nos-Cas9 parental fly line that I would be using; the nos promoter drives Cas9 expression in the germline (Port *et al.*, 2014). This sequencing allowed me to check for any single nucleotide polymorphisms (SNPs) in the nos-Cas9 strain that could affect guide RNA (gRNA) binding. I found some SNPs, but they did not affect the relevant target sequences.

I used an online gRNA design tool to check that the gRNAs I had chosen were predicted to have 0 or less than 2 off-target effects in other parts of the *Drosophila* genome (Gratz *et al.*, 2014). I found that some gRNAs had predicted off-target effects in an unannotated Flybase sequence. This could potentially represent a duplicated region of the fly genome; during nucleotide sequencing of this region in the nos-Cas9 flies, I had also observed some double-peaks which further supported this conclusion. I therefore decided to also avoid this potential 'duplicated region' when designing my gRNAs (**Figure 4.2 A**).

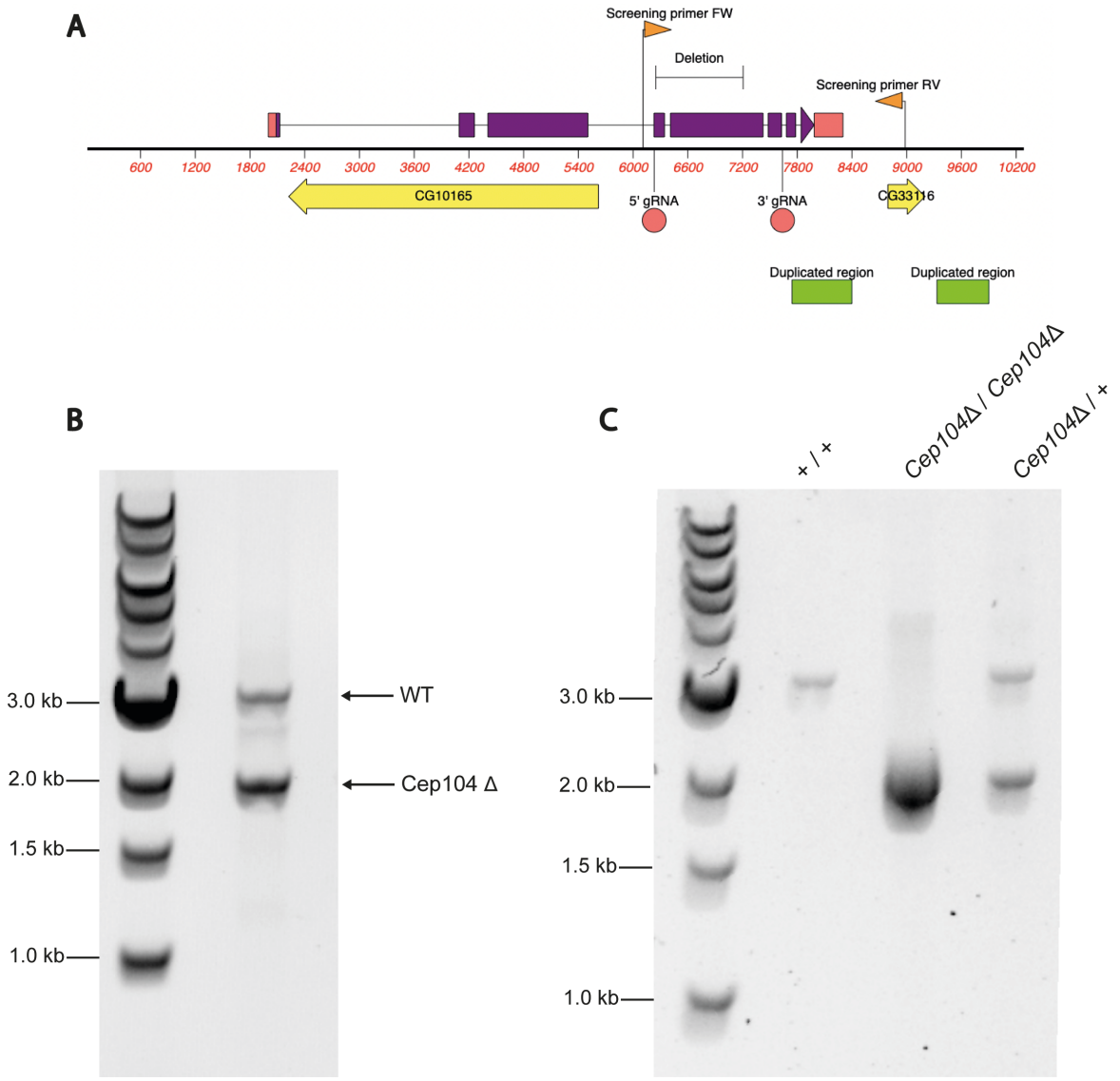


Figure 4.2 – Generation of a *cep104* knockout fly line using CRISPR/Cas9.

(A) Map showing the exons (purple boxes; arrow indicates direction of transcription) and 3' and 5' UTRs (red boxes) of the *Drosophila cep104* gene. Guide RNAs were designed to bind at the regions marked by red circles. These regions were chosen to avoid overlapping genes *CG10165* and *CG33116* (yellow boxes; arrows indicate direction of transcription) and potential regions of gene duplication (green boxes). Screening primers were also designed (orange arrows) to anneal either side of the region targeted for deletion; this would allow me to screen flies for deletion events. The resulting deletion was mapped and is denoted by the black bracket above the *cep104* gene region. **(B)** PCR product from a fly heterozygous for the *cep104* deletion shows the size of the region between the screening primers in either WT or in *cep104* deletion cases. When the *cep104* gene region is WT, the band produced is ~3 kb. A deletion event causes the size of this band to decrease to ~2 kb. This is consistent with the size of the deletion (958 bp) which I determined by sequencing the *cep104* deletion mutants. **(C)** PCR product from flies that are WT (+/+), homozygous for the *cep104* deletion or heterozygous for the *cep104* deletion, confirming the identity of the PCR bands shown in (B).

In order to screen for progeny where a deletion event had occurred, I also designed PCR primers that would anneal at either side of the gRNA binding sites (indicated by the orange arrows in **Figure 4.2 A**). This would allow me to select flies that produced a second, smaller PCR product (**Figure 4.2 B, C**), indicating that they were heterozygous for the *cep104* deletion.

I cloned my gRNAs into pCFD4 plasmids, chosen because this vector can drive expression of two gRNAs (Port *et al.*, 2014), and sent them to the Cambridge Fly Facility for the generation of fly lines. Two strategies for this were attempted:

4.2.2.1.1. Strategy 1: transient expression of gRNAs in nos-Cas9 flies

At the Fly Facility, constructs were injected into nos-Cas9 larvae (Port *et al.*, 2014) for transient gRNA expression. I placed these larvae in quarantine until flies began to eclose. I collected males and virgin females, which I placed into separate vials and crossed with SM5/Gla for balancing (**Figure 4.3**).

Once these crosses produced larvae, I lysed the founders and screened their microlysates for deletions via PCR, using the screening primers I had designed. I did not observe any PCR products corresponding to deletion events.

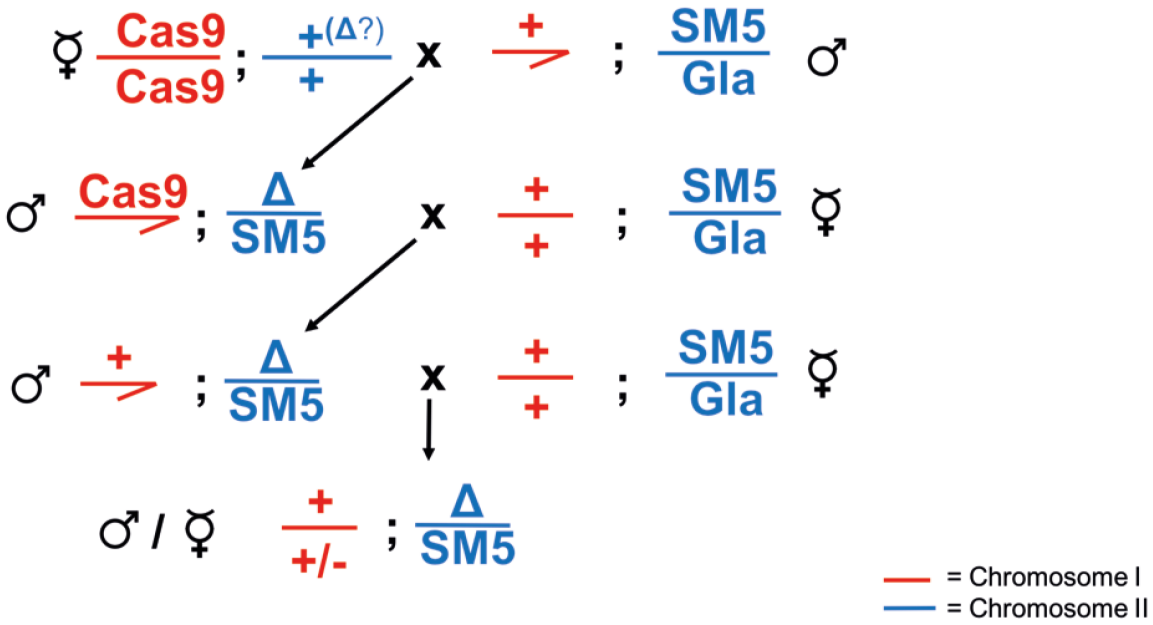


Figure 4.3 – Crossing scheme for transient gRNA expression strategy.

Crossing scheme for generating a balanced *cep104* deletion stock starting with a female founder with transiently expressed gRNAs. The crossing scheme for starting with a male founder was the same, but had one fewer crossing step when crossing with the SM5/Gla flies to eliminate the Cas9. Δ indicates a suspected deletion.

4.2.2.1.2. Strategy 2: transgenic expression of gRNAs

Since Strategy 1 failed to produce a deletion, I attempted an alternative approach. At the Fly Facility, constructs were injected into BL25709 flies for transgenic integration of gRNAs. I crossed these transgenic flies with nos-Cas9 flies (Port *et al.*, 2014) before collecting virgin female progeny from this cross and placing them in separate vials for crossing with SM5/Gla males (**Figure 4.4**). I lysed and screened these founders as before. This time, I observed deletion events with an efficiency of ~1 in 50. Of three recovered deletions, one female founder transmitted the deletion to its male offspring, which I

selected to set up the balanced deletion stock. The difference in size between the WT and the *cep104* deletion PCR products is shown in **Figure 4.2 B, C**.

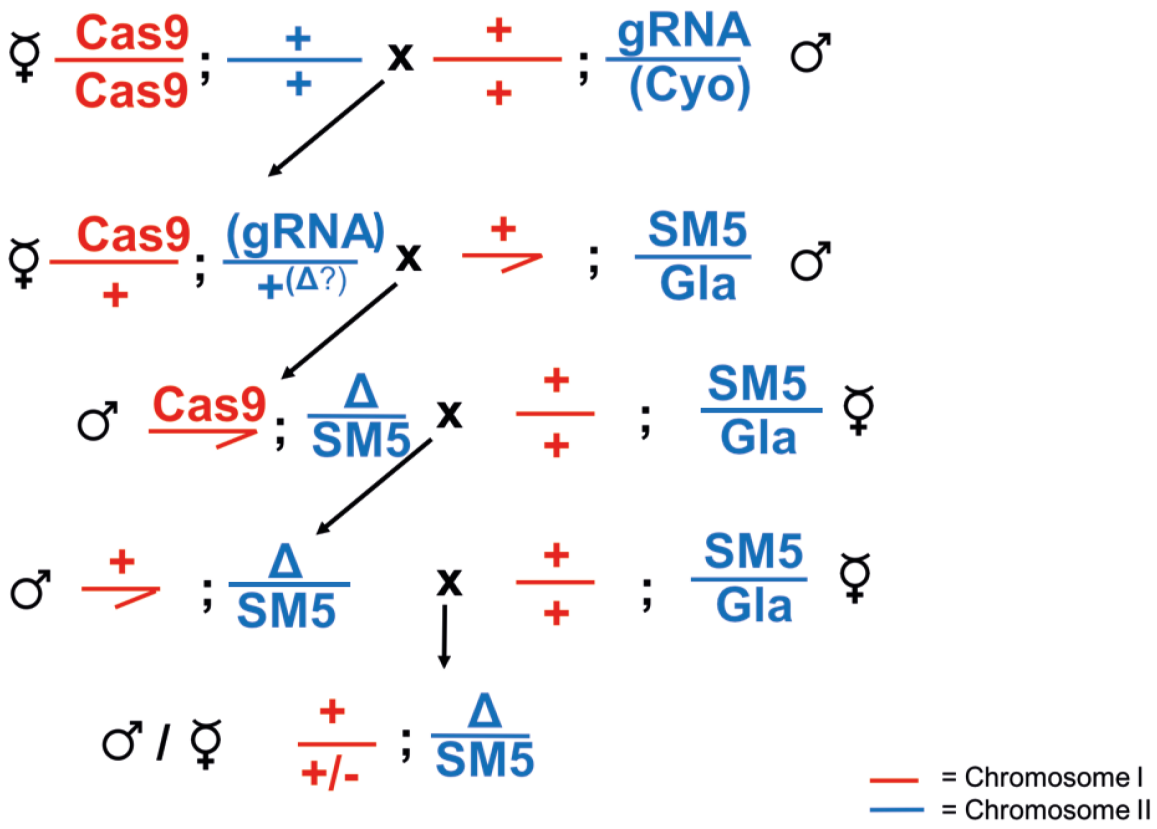


Figure 4.4 – Crossing scheme for transgenic gRNA integration strategy.

Crossing scheme for generating a balanced *cep104* deletion stock starting with Cas9 flies and flies stably expressing the gRNAs. These two fly strains were crossed and their female progeny were used as founders to set up crosses with SM5/Gla. Δ indicates a suspected deletion.

DNA sequencing demonstrated that a region of 958 bp had been deleted in the C-terminus of the *cep104* gene (this deletion is indicated on the gene region schematic in **Figure 4.2 A**). The deletion does not extend across the full length of the region between the two gRNAs that I designed. I suspect that this is because the 5'gRNA was more efficient than the 3'gRNA and may have caused a break in the DNA that was then chewed back to create

the deletion. This deletion removes exon 4 and the majority of exon 5 (the region coding for amino acids 442-739).

I also wanted to confirm the absence of full-length Cep104 from my mutants via western blot. I purified from rabbit sera an α-Cep104 antibody generated by Metta Pratt (Raff Lab, University of Oxford). This antibody is designed to recognise the N-terminus of Cep104, binding in the range of amino acids 1-358. Hence, if any short versions of the Cep104 protein remained following my genetic ablation of the C-terminal region, these should also be recognised by this antibody.

I then tested the purified antibody on samples of fixed *Drosophila* embryos. Unfortunately, no bands were detected (data not shown). This remained the case even when I used a higher concentration of embryos in my samples, or when I attempted to use the original sera to probe the samples.

Since Cep104 expression is higher in brains than in embryos (Larkin *et al.*, 2020) I dissected out brains from third-instar larvae and fixed these for western blotting. I then tested both the sera and the purified antibody on these. As before, the purified antibody did not yield any bands. However, using the sera I was now able to detect bands of approximately the expected size (104 kDa). To determine which of these bands corresponded to Cep104, I ran a western blot where I dissected out brains from WT, Ubq-Cep104-GFP and homozygous *cep104* mutant larvae (**Figure 4.5 A**). I observed that one of the bands of the approximate expected size disappeared in the *cep104* mutant larvae.

The Cep104-GFP band was also detected to run at a higher molecular weight than this band, as expected.

To further confirm the identity of this band, I also ran western blots where I dissected out the brains of WT, Ubq-Cep104-GFP, and Ubq-Cep104 (untagged) larvae. Since untagged Cep104 is thought to be overexpressed in the latter fly line, I reasoned that the Cep104 band should increase in size in these Ubq-Cep104 brain samples. Indeed, this was the case, both confirming the identity of the band and demonstrating that Cep104 is ~8x overexpressed in the Ubq-Cep104 brains compared with WT brains (**Figure 4.5 B**; levels estimated by eye).

Having identified the Cep104 band, and having observed that this band was absent in the *cep104* mutant larvae, I was therefore able to confirm that the full-length version of the Cep104 protein is not present at detectable levels in flies that are homozygous for the CRISPR deletion.

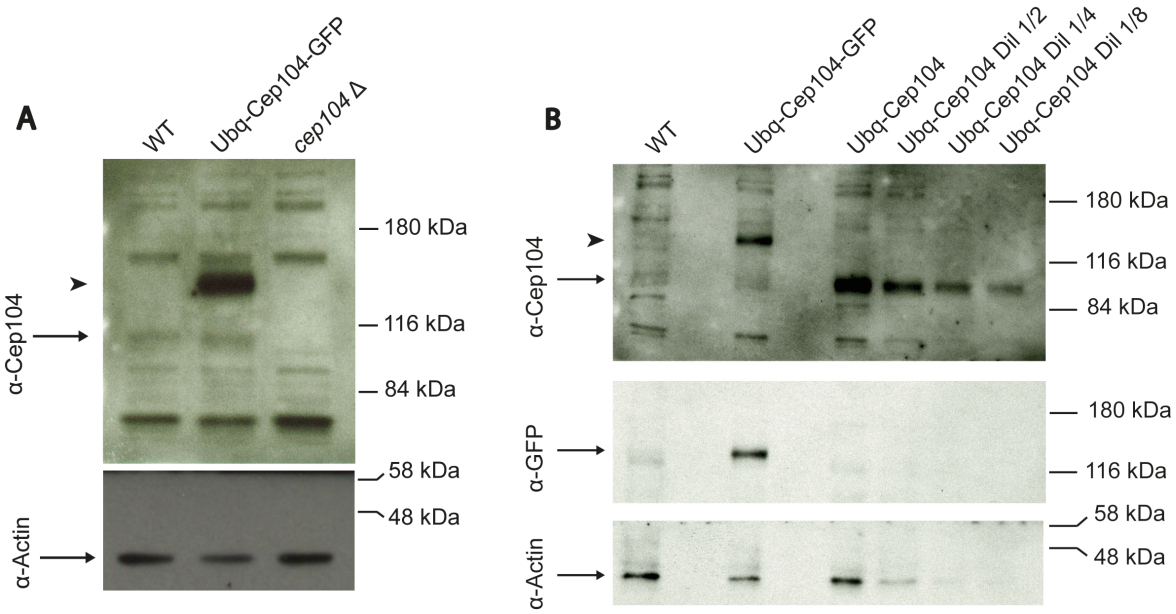


Figure 4.5 – Western blots confirm the absence of full-length Cep104 in the deletion stock.

Western blots using α-Cep104 sera to probe samples from *Drosophila* larval brains. **(A)** Brains from WT larvae, larvae expressing both endogenous Cep104 and Ubq-Cep104-GFP, and larvae homozygous for the *cep104* CRISPR deletion (*cep104* Δ) were assayed. A band of the expected size for Cep104 was detected in the WT and Cep104-GFP brains, but not in the brains of larvae homozygous for the deletion. This band is indicated with an arrow. In the Ubq-Cep104-GFP brain sample, a band corresponding to GFP-tagged Cep104 (arrowhead) was detected running at a higher molecular weight than the Cep104 band, as expected. Actin is shown as a loading control. N = 2 technical repeats. **(B)** Brains from WT larvae, larvae expressing both endogenous Cep104 and Ubq-Cep104-GFP, and larvae expressing one copy of untagged Ubq-Cep104 were assayed. Serial dilutions of the Ubq-Cep104 sample were also analysed. The band corresponding to Cep104 (arrow) is stronger in the Ubq-Cep104 sample than in the WT or Ubq-Cep104-GFP sample. The band corresponding to GFP-tagged Cep104 is indicated with an arrowhead. Actin is shown as a loading control. N = 2 technical repeats.

The deletion in the *cep104* mutant fly line roughly corresponds to the region of the human protein containing the TOG domain (**Figure 4.6 A, B**). However, the region of the gene coding for the N-terminus of the protein was left intact. This could not be avoided due to overlapping genes in the *cep104* gene region (**Figure 4.2 A**). In humans, amino acids 1-163 of Cep104 are implicated in Cep97 binding (Jiang *et al.*, 2012, Rezabkova *et al.*, 2016). To test whether a shortened form of Cep104 is expressed in the mutant fly line, I

used the α-Cep104 sera to repeat the blot in **Figure 4.5 A**, this time probing the whole membrane to check if any bands are present in the *cep104* mutant sample that run at a lower molecular weight than the full-length Cep104 band. I did not observe any, suggesting that a shortened form of Cep104 is not expressed, or at least not at detectable levels (**Figure 4.6 C**).

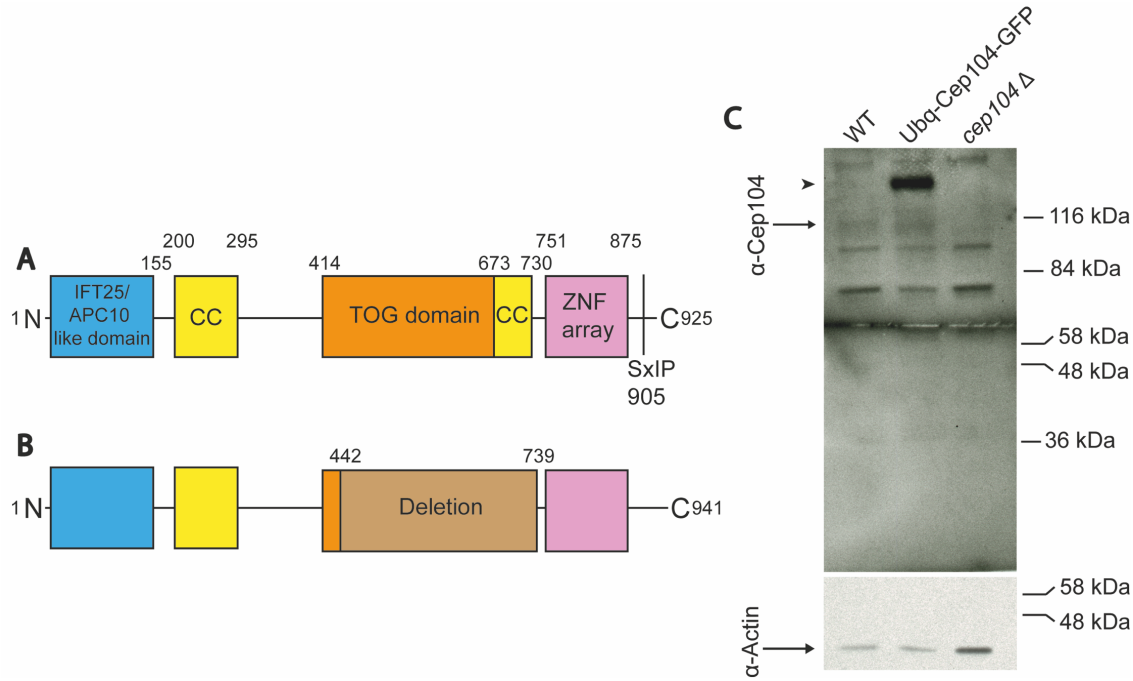


Figure 4.6 – Expression of a shortened version of Cep104 was not detectable via Western blotting.

(A) Schematic of the human Cep104 protein, showing its key domains (adapted from Al-Jassar *et al.* 2017). **(B)** Schematic showing which regions of *Drosophila* Cep104 I have deleted in my CRISPR mutant. The regions roughly corresponding to the known human Cep104 domains are illustrated in the same colours as (A). **(C)** Western blot using α-Cep104 sera to probe samples from *Drosophila* larval brains. Brains from WT larvae, larvae expressing both endogenous Cep104 and Ubq-Cep104-GFP, larvae homozygous for the *cep104* CRISPR deletion (*cep104* Δ) were assayed. The band corresponding to Cep104 is indicated with an arrow. This band is absent in the *cep104* mutant brains. The band corresponding to GFP-tagged Cep104 is indicated with an arrowhead. No bands running at a lower molecular weight are apparent in the *cep104* mutant brain sample. Actin is shown as a loading control. N = 1.

4.2.2.2. *cep104* knockout flies are viable, fertile and coordinated

I recovered some homozygous *cep104* mutants from my balanced stock, indicating that these flies are viable despite lacking the full-length Cep104 protein. These flies appeared to be coordinated, suggesting that they had functional cilia. I went on to carry out negative geotaxis assays using these flies, testing how far 10 homozygous mutants climbed compared with 10 flies from the parental line (which expressed the gRNAs but not Cas9, so should not have experienced any deletion events) I had generated. I had attempted to generate an alternative control by following the crossing scheme in **Figure 4.4** and selecting progeny that did not exhibit a deletion when screened via PCR. However, when I sequenced these flies, they all had small deletions near the 5'gRNA binding site. This further supports my conclusion that the 5'gRNA was particularly potent, and that the deletion I managed to generate was likely a result of a DNA cut at the 5'gRNA site. I carried out four technical repeats of the negative geotaxis assay, and found that there was no significant difference in the mean distance climbed by the flies in each condition (**Figure 4.7 A**). I also ranked the flies from each technical repeat by distance climbed, and calculated the mean distance for each rank (**Figure 4.7 B**). Again, this analysis showed no significant difference between the mutant and its parental line. Finally, I counted the number of flies that climbed at least 8 cm in 10 s and compared this measurement. Again, flies homozygous for the *cep104* deletion did not appear to perform differently from the parental control line, though numbers were of course small here since this statistic is calculated from four repeats (**Figure 4.7 C**).

To test whether the *cep104* knockout flies were fertile, I selected males and virgin females that were homozygous for the *cep104* deletion and placed them together in vials of fly

food. A few days later, I observed that these matings had produced larvae, demonstrating that both male and female flies homozygous for the *cep104* deletion are fertile.

These results are surprising, since Cep104 is thought to play an important role in ciliogenesis. However, this coordinated, fertile phenotype is reminiscent of the *cp110* null flies, which are also grossly normal (Franz *et al.*, 2013). *cep97* null flies also appear normal, yet negative geotaxis assays reveal behavioural defects and their chordotonal organs are disrupted (Dobbelaere *et al.*, 2020). It will be interesting to investigate how the morphology of cilia in the sensory neurons of *cep104* null flies compares with WT.

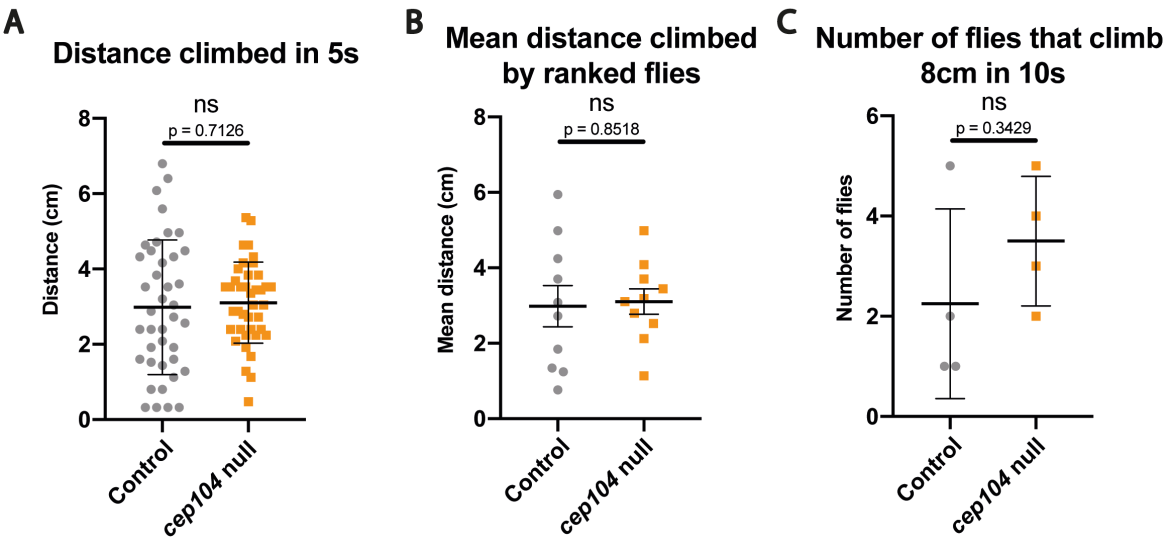


Figure 4.7 – Negative geotaxis assays to assess the coordination of *cep104* mutant flies.

(A) Graph showing quantifications of the distance climbed in 5 s by flies that are homozygous for the *cep104* deletion, compared with a parental control fly line. Data are presented as mean ± SD; N = 40. **(B)** Graph showing the mean distance (quantified from the data in (A)) climbed by flies ranked in ascending order of distance climbed. Data are presented as mean ± SEM; N = 4 technical repeats, n = 10 flies per repeat. **(C)** Graph showing the number of flies homozygous for the *cep104* deletion that climbed more than 8 cm in 10 s, compared with a parental control fly line. Data are presented as mean ± SD; N = 4 technical repeats. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

4.2.2.3. GFP-Cep97 and CP110-GFP are still recruited to centrioles in the absence of Cep104

To address whether Cep104 might play a role in recruiting CP110/Cep97 to the ends of centrioles, I generated fly lines expressing either Ubq-CP110-GFP or Ubq-GFP-Cep97 in either a WT or *cep104* mutant background. Embryos from these lines were investigated with spinning disc confocal microscopy so that the localisation of CP110-GFP/GFP-Cep97 could be observed. In both cases, both CP110-GFP and GFP-Cep97 were able to localise to centrioles (**Figure 4.8**). The GFP signal appeared to be brighter in S-phase than in mitosis, which is consistent with the CP110/Cep97 recruitment dynamics described in **Chapter 3.** This result suggests that Cep104 is not required for the recruitment of these proteins to centrioles.

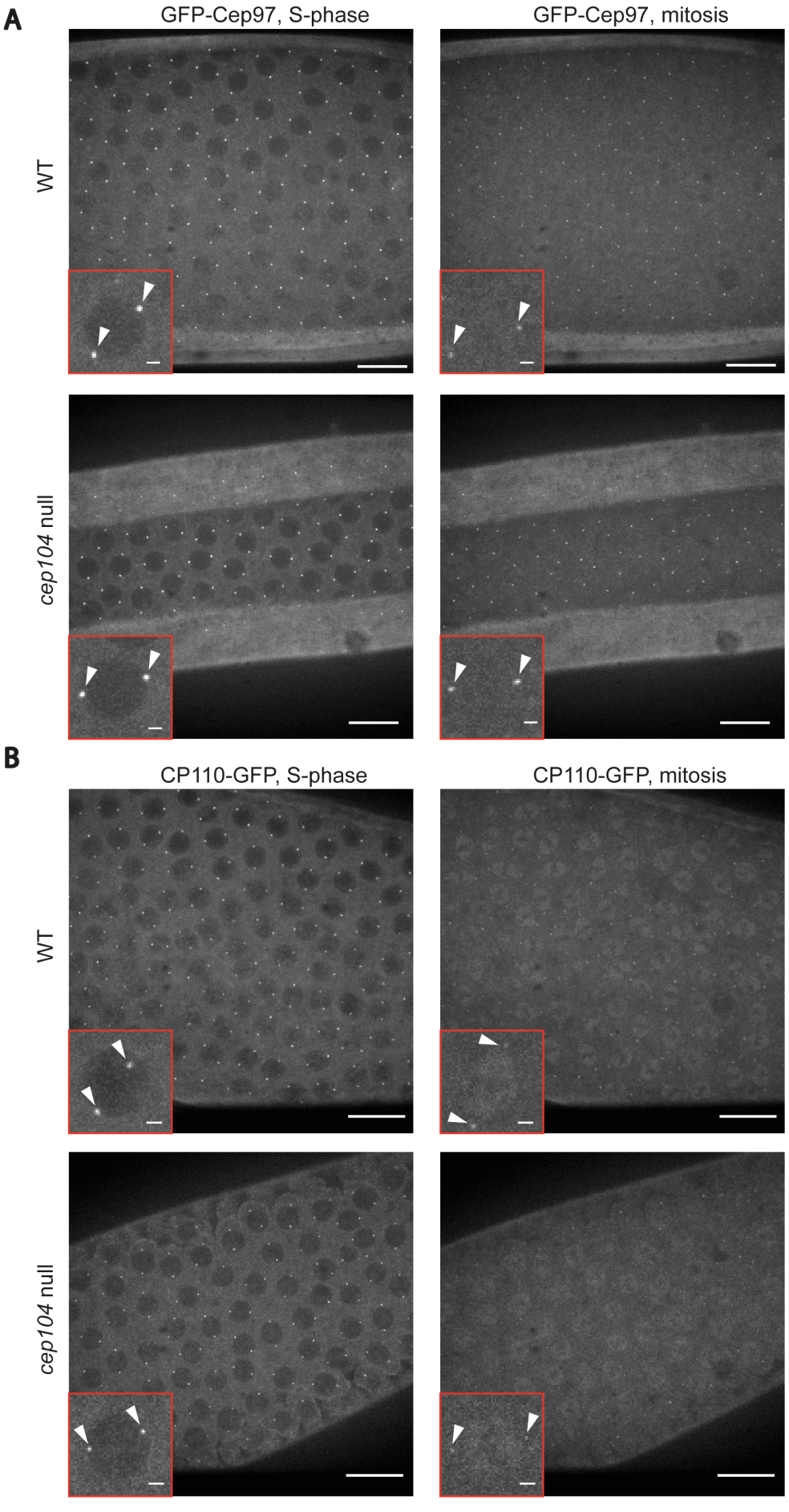

A
GFP-Cep97, S-phase
GFP-Cep97, mitosis
WT
cep104 null
B
CP110-GFP, S-phase
CP110-GFP, mitosis
WT
cep104 null

Figure 4.8 – Micrographs demonstrating that CP110/Cep97 localise to centrioles in the absence of full-length Cep104.

(A-B) Spinning disc confocal images showing Ubq-GFP-Cep97 (A) or Ubq-CP110-GFP (B) localisation during cycle 12 (mitosis and S-phase) in either WT or *cep104* null embryos. In both conditions, GFP-Cep97 and CP110-GFP can be seen localising to centrioles. Insets (orange borders) show a pair of centrosomes from each image in greater detail; arrowheads indicate the location of each centrosome. Scale bar in main image represents 20 μm, scale bar in inset represents 2 μm.

4.2.2.4. The parameters of the GFP-Cep97 oscillation are not perturbed by the absence of full-length Cep104

We have previously shown that CP110/Cep97 exhibit oscillatory behaviour when they are recruited to centrioles, peaking in S-phase and decreasing before the onset of mitosis (**Chapter 3**). I was therefore interested in whether this oscillatory behaviour might be affected by the absence of Cep104.

To investigate this, I used the fly lines where Ubq-GFP-Cep97 is expressed in either a WT or *cep104* mutant background to generate movies of nuclear cycle 12. I could then compare the parameters of the GFP-Cep97 oscillation in each case. In embryos lacking Cep104, I found that the Cep97 oscillation occurred just as it does in WT embryos. The intensity acquired (the change in fluorescence intensity from the initial measurement at the start of S-phase to the peak of the Cep97 oscillation profile) was not significantly different between *cep104* null and WT embryos (**Figure 4.9 A, C**). I also found that there was no significant change in the time taken to reach the peak of the oscillation (**Figure 4.9 A, C**). Since the peak of the Cep97 oscillation is strongly correlated with S-phase length, I also

checked whether correcting for differences in S-phase length between the two datasets had any impact on the position of the oscillation centre. Again, I saw that the S-phase-corrected centres were not significantly different between the WT and *cep104* null embryos (**Figure 4.9 E**).

I went on to investigate the GFP-Cep97 oscillation when Cep104 is overexpressed. To do this, I again generated movies of the oscillation in S-phase of cycle 12, this time using fly lines where Ubq-GFP-Cep97 was expressed alongside untagged Ubq-Cep104, in order to increase the dose of Cep104 in the embryo. Again, embryos expressing Ubq-GFP-Cep97 with WT levels of Cep104 expression acted as a control. I found that the GFP-Cep97 oscillation was unaffected by an increase in Cep104 dose (**Figure 4.9 B, D**). There was a slight increase in the time taken to reach the oscillation centre in Cep104 overexpression conditions when S-phase length corrections were applied (**Figure 4.9 F**). However, this effect is relatively small and is only observed when S-phase length corrections are performed. Overall, these results therefore suggest that Cep104 is not important for the appropriate recruitment of CP110/Cep97 in the developing fly embryo.

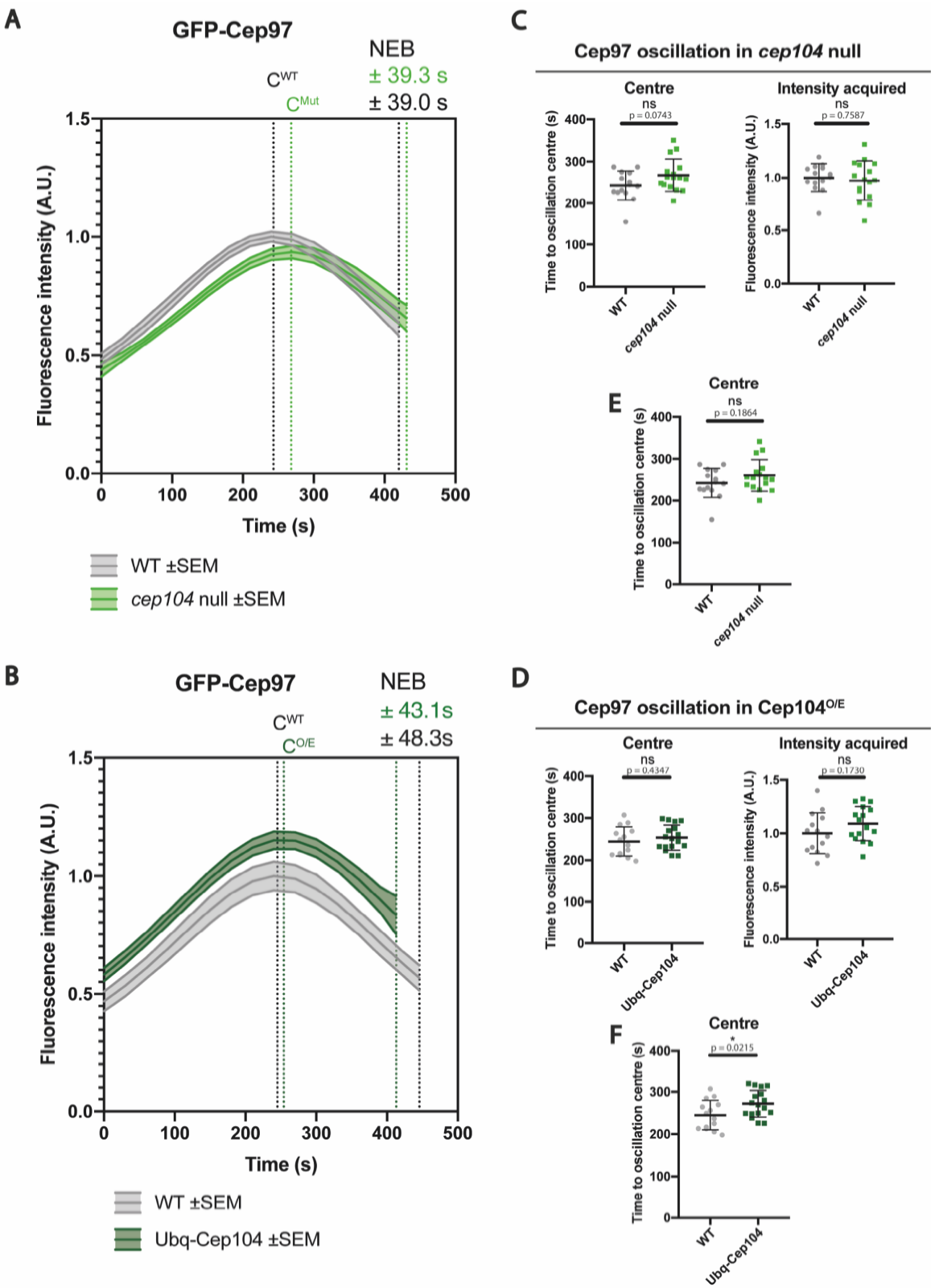

A
GFP-Cep97
NEB
C^WT ± 39.3 s
C^Mut ± 39.0 s
Fluorescence intensity (A.U.)
Time (s)
WT ±SEM
cep104 null ±SEM
B
GFP-Cep97
NEB
C^WT ± 43.1s
C^O/E ± 48.3s
Fluorescence intensity (A.U.)
Time (s)
WT ±SEM
Ubq-Cep104 ±SEM
C
Cep97 oscillation in cep104 null
Centre
ns
p = 0.0743
Intensity acquired
ns
p = 0.7587
Time to oscillation centre (s)
Fluorescence intensity (A.U.)
WT
cep104 null
E
Centre
ns
p = 0.1864
Time to oscillation centre (s)
WT
cep104 null
D
Cep97 oscillation in Cep104^O/E
Centre
ns
p = 0.4347
Intensity acquired
ns
p = 0.1730
Time to oscillation centre (s)
Fluorescence intensity (A.U.)
WT
Ubq-Cep104
F
Centre
*
p = 0.0215
Time to oscillation centre (s)
WT
Ubq-Cep104

Figure 4.9 – The parameters of the GFP-Cep97 oscillation in varying doses of Cep104.

(A-B) Average curves showing the Cep97 oscillation over S-phase of cycle 12 in WT and *cep104* null embryos (A) or WT and Ubq-Cep104 embryos (B). The peaks of the oscillation are indicated by C^{WT} (WT embryos), C^{Mut} (*cep104* null embryos) and $C^{O/E}$ (Ubq-Cep104). NEB = Nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(C-D)** Scatter plots showing quantification of time to the peak (or *Centre*) of the oscillation and of intensity acquired in either WT and *cep104* null (C) or WT and Ubq-Cep104 (D) embryos. **(E-F)** Centres from (C) and (D) corrected for S-phase length, to account for differences in S-phase length between the datasets. For (C-F), data are presented as mean ± SD; N ≥ 14 embryos, n ≥ 28 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

4.2.3. Like CP110/Cep97, Cep104 is recruited to centrioles in an oscillatory manner

4.2.3.1. Levels of Cep104-GFP oscillate at the centriole over the course of the cell cycle, peaking in S-phase

Since I had shown that Cep104 appeared to play no role in the proper recruitment of CP110/Cep97 to the centriole, I was curious as to the order in which Cep104, CP110 and Cep97 were recruited to centrioles. I was also interested in the dynamics of Cep104 recruitment in itself. To investigate this, I first followed the recruitment dynamics of Ubq-Cep104-GFP over multiple embryonic cycles (**Figure 4.10 A**). This recruitment appears to be rapid, peaking in S-phase of cycles 11, 12 and 13. This rapid rise to a peak is followed by a loss of Cep104-GFP from centrioles that begins before mitotic onset. This therefore represents an oscillatory mode of recruitment, reminiscent of CP110/Cep97 dynamics.

I next generated flies expressing both Ubq-Cep104-GFP and Ubq-RFP-Cep97, so that I could image the Cep104 oscillation and the Cep97 oscillation in the same embryo. Again, this imaging was carried out over cycles 11-13. Results from this initial experiment suggested that, in some embryos, Cep104 might be reaching the peak of its recruitment at a later time in S-phase than Cep97 does (**Figure 4.10 B**). By cycle 13, it is challenging to compare the peaks since the Cep97 peak seems to be much broader than the more distinct peak of the Cep104 signal. This sharp peak that I see in cycle 13 reflects the fact that the Cep104 amplitude also appears to grow higher as the nuclear cycles grow longer (**Figure 4.10 B**). This is particularly dramatic in nuclear cycle 13. By contrast, the amplitude of the Cep97 oscillation remains fairly constant between cycles, suggesting that the amount recruited to the centrioles is more consistent (**Figure 4.10 B**).

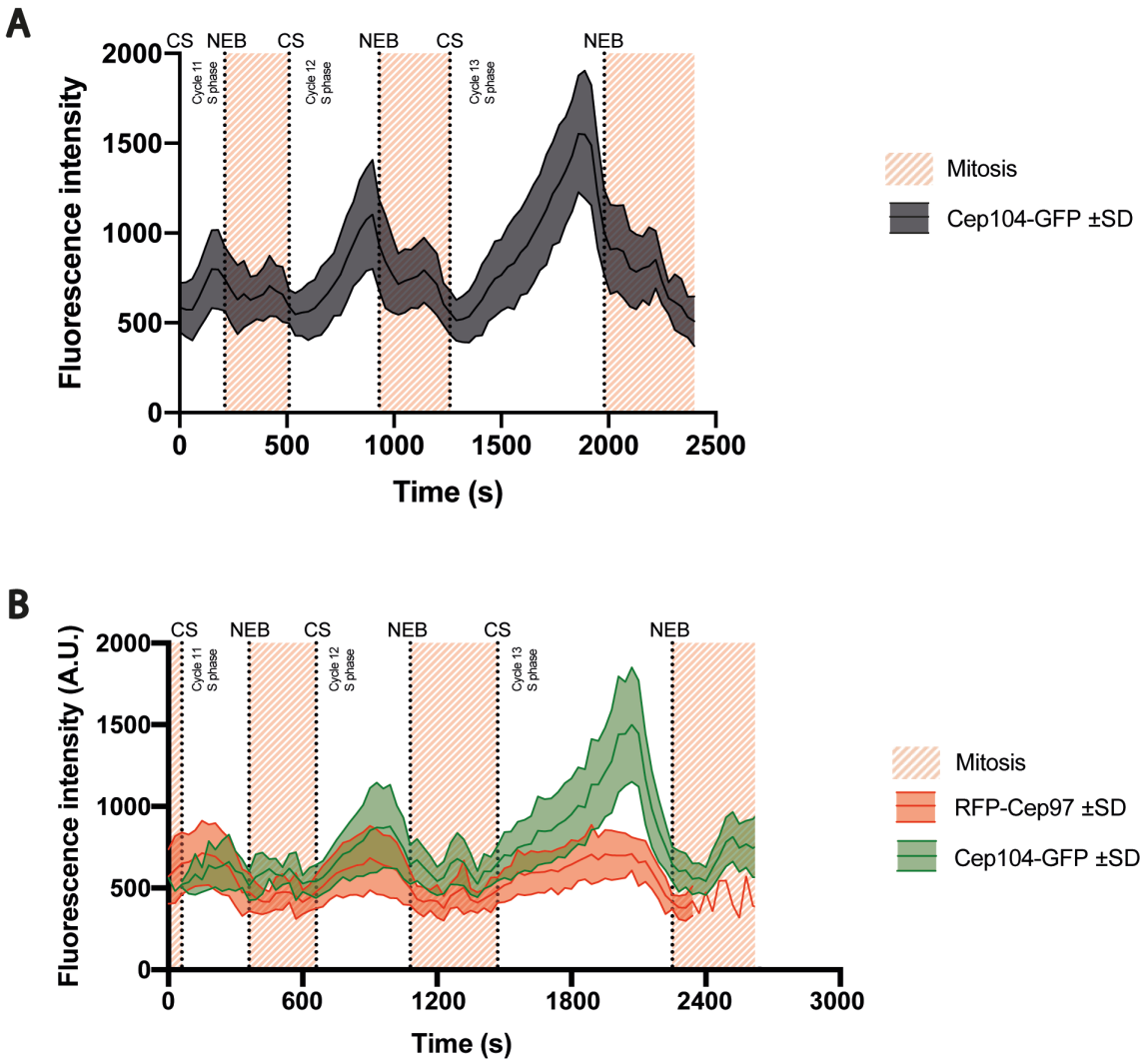


Figure 4.10 – Graphs showing the recruitment dynamics of Cep104-GFP over cycles 11-13.

(A-B) Graphs showing quantification of Ubq-Cep104-GFP (A) or Ubq-Cep104-GFP and Ubq-RFP-Cep97 (B) fluorescence intensity at centrosomes in a single embryo over nuclear cycles 11-13. Cep104 recruitment to centrioles appears to peak in late S-phase, decreasing before the onset of mitosis. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SD.

4.2.3.2. Cep104-GFP fluorescence peaks later in S-phase than RFP-Cep97

To further investigate the relative timing of the Cep104 and Cep97 peaks in a more quantitative manner, I used the same fly line as above to generate multiple movies of the Cep104 and Cep97 oscillations over the course of cycle 12. I chose to focus on cycle 12 because it was possible to capture a large enough number of movies (since this cycle is only ~12 minutes long; Foe and Alberts, 1983) with a large number of centriole pairs. Cycle 12 is also the cycle most routinely used by our lab, so our most detailed information about centriole protein dynamics concerns this cycle.

Once I had acquired these movies, I wanted to fit non-linear regressions to the curves from each embryo in order to determine the time at which each oscillation reaches a peak. Using these regressions would reduce any bias of determining when the intensity had peaked by eye. As discussed in **Chapter 3**, it is possible to use the Lorentzian equation to model the time to peak for Cep97 (**Figure 3.9**). However, I needed to determine a model that was a good fit for the Cep104 data, in order to also be able to identify the peak of Cep104.

To do this, I imaged embryos expressing just Ubq-Cep104-GFP over cycle 12 and then analysed these movies to extract fluorescence intensity data for each embryo. I then identified three non-linear regressions that appeared to fit this data well:

(1) Linear increase, linear decrease

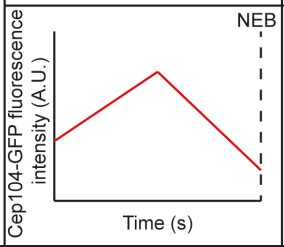
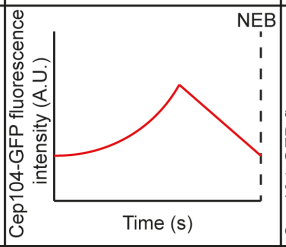
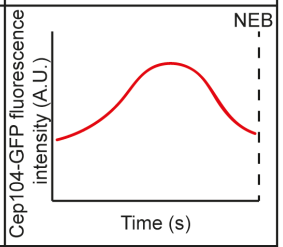
(2) Exponential increase, linear decrease

(3) Sum of two Gaussians

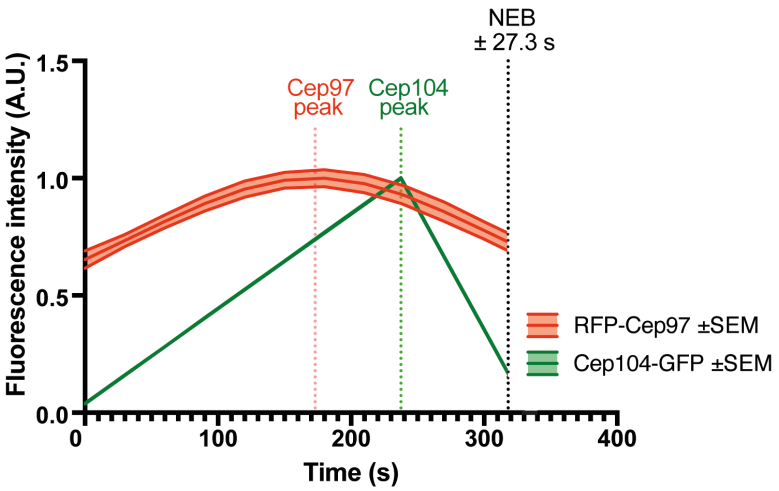
To test which of these would be the most appropriate fit, I used GraphPad Prism 8 to compare the goodness-of-fit of these three equations for each embryo from the dataset. I found that equations *(1)* and *(3)* gave very promising R^2 values of >0.95, suggesting a good fit to the underlying data. Overall, equation *(1)* was the preferred fit of 50% of the embryos that I analysed (**Figure 4.11 A**). I therefore decided to use this equation to fit the data from my two-colour imaging and compare the peaks of the Ubq-Cep104-GFP and Ubq-RFP-Cep97 data. In future, it would be beneficial to generate a larger dataset of Ubq-Cep104-GFP over cycle 12 to test these fits again, but for this analysis I was looking simply for an unbiased way of comparing the time at which peak fluorescence intensity is reached. Once I had fitted equation *(1)* to my Cep104-GFP data, I also performed a visual inspection to confirm that the inflection point of the fitted curve appeared to fall around the time of the apparent peak in fluorescence intensity.

I was now able to fit *Lorentzian* curves to the Cep97 data and *Linear increase, linear decrease* curves to the Cep104 data. Using the inflection points of the fitted curves, I was able to compare the time to oscillation peak for Cep104 and Cep97. I found that Cep104 peaks significantly later in S-phase than Cep97 (**Figure 4.11 B, C**).

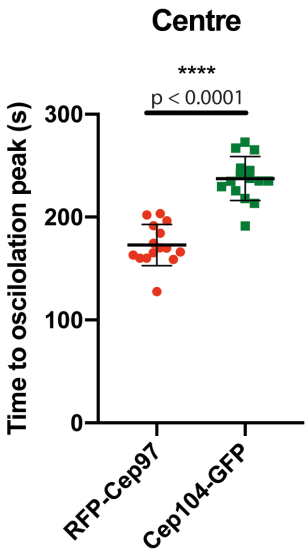
A

N = 8 embryos	Linear increase, linear decrease	Exponential increase, linear decrease	Sum of two Gaussians
			
Deg. of freedom	10	10	8
R^2 value	0.951±0.03	0.848±0.03	0.977±0.02
% of fit preference	50%	25%	25%

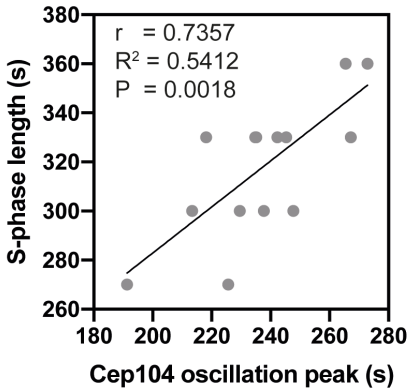
B



C



D



E

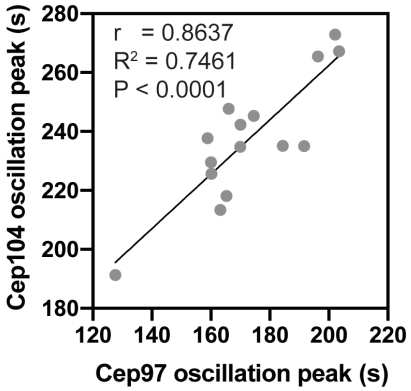


Figure 4.11 – Comparisons of the Cep104 and Cep97 peaks in cycle 12.

(A) To assess how to fit the Cep104 dynamic recruitment data, Ubq-Cep104-GFP data were fitted against three separate regression models (illustrated in red): *(1) Linear increase, linear decrease (2) Exponential increase, linear decrease (3) Sum of two Gaussians*. The majority of the graphs preferentially fitted *(1)*, so this model was selected to fit all Cep104 data. R^2 values were also assessed, since these give an indication of the goodness-of-fit. Further details of these models are provided in **Chapter 2**. **(B)** Average curves drawn from a population of embryos in cycle 12 to illustrate the mean peak times of Cep104 and Cep97. NEB = Nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(C)** Quantification of Ubq-Cep104-GFP peaks and Ubq-RFP-Cep97 peaks from a population of cycle 12 embryos. Data are presented as mean ± SD; N = 15 embryos, n ≥ 24 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs. **(D-E)** Graphs showing that there is a positive correlation between the Cep104 oscillation peak and S-phase length (D), and between the Cep104 and Cep97 oscillation peaks (E). Oscillation peak times are taken from (C). *r*, R^2 and *p*-values are shown on the graphs and linear regressions have been fitted.

4.2.3.3. The peak of the Cep104 oscillation correlates with S-phase length

Since I now had values for time to oscillation peak for both Cep97 and Cep104, I decided to check whether the peak of the Cep104 oscillation correlates with the length of S-phase. My correlation analysis revealed that the time at which the Cep104 peak falls does indeed correlate with the time taken to complete S-phase (**Figure 4.11 D**).

This is reminiscent of the strong correlation observed between S-phase length and time to CP110/Cep97 peak (**Chapter 3**). Given this, it is perhaps unsurprising that I found the time to Cep104's oscillation peak also correlated with the time to Cep97's oscillation peak (**Figure 4.11 E**). However, it should be noted that the correlation between the Cep104 peak and S-phase length (**Figure 4.11 D**) does not appear to be as strong as the correlation between the CP110/Cep97 peak and S-phase length (**Figure 3.9**).

4.2.4. Cep104 recruitment to centrioles is dramatically reduced in the absence of CP110/Cep97

Since the levels of Cep104 at the centriole appear to peak later in S-phase than Cep97 levels, with both proteins' recruitment appearing to be tightly coupled to the length of S-phase (and therefore to the cell cycle), it was of interest to determine whether the recruitment of Cep104 might in fact be dependent on CP110/Cep97. Mustafa Aydogan generated WT embryos expressing Ubq-Cep104-GFP and compared its localisation to that of Ubq-Cep104-GFP in embryos where CP110 or Cep97 had been genetically ablated. When CP110 or Cep97 were absent, the localisation of Ubq-Cep104-GFP to centrioles was dramatically reduced (**Figure 4.12**).

Overall, these results suggest that Cep104 acts at the centriole downstream of CP110/Cep97, and is in fact dependent on CP110/Cep97 for its recruitment or for its maintenance at centrioles.

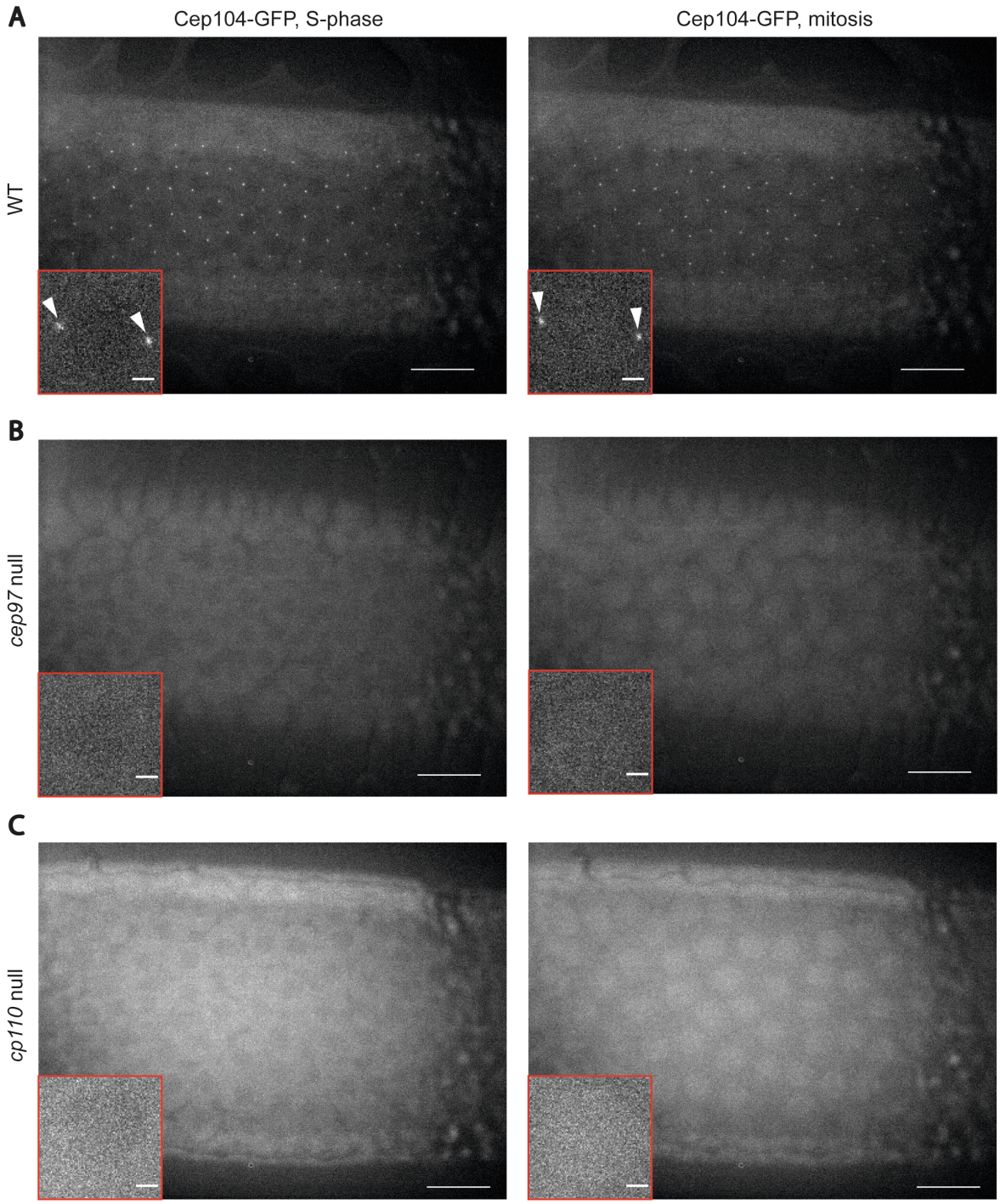


Figure 4.12 – Micrographs demonstrate that Cep104 recruitment is dramatically reduced in the absence of either CP110 or Cep97.

(A-C) Spinning disc confocal images showing Ubq-Cep104-GFP localisation during cycle 12 (mitosis and S-phase) in WT (A), *cep97* null (B) or *cp110* null (C) embryos. Recruitment of Cep104 to centrioles is dramatically reduced in both the *cep97* null and *cp110* null conditions. Insets (orange borders) show a pair of centrosomes from each image in greater detail; where centrosomes are visible, arrowheads indicate their location. Scale bar in main image represents 20 µm, scale bar in inset represents 2 µm. Images used to produce this figure are courtesy of Mustafa Aydogan.

4.2.5. Cep104 levels have little effect on cartwheel growth during centriole biogenesis

4.2.5.1. Centriole cartwheel growth is not perturbed by in *cep104* mutant heterozygotes

Since we were surprised to observe that altering the dose of CP110/Cep97 affected the growth of the cartwheel during daughter centriole biogenesis (**Chapter 3**), and because proper Cep104 recruitment to centrioles appears to be reliant on CP110/Cep97, I was curious as to whether Cep104 might also play a role in cartwheel growth control. In order to address this question, I first generated flies that were either WT or heterozygous for my *cep104* deletion and also expressed Sas-6-GFP under the endogenous *sas-6* promoter. I found that there was no significant difference in Sas-6-GFP incorporation over S-phase between WT embryos and those embryos that were heterozygous for the *cep104* deletion (**Figure 4.13**). The cartwheels also reached the same size (Sas-6-GFP intensity acquired) at the end of the growth phase for both datasets.

I then wanted to go on and follow Sas-6 incorporation in embryos that were homozygous for my *cep104* deletion, to investigate whether the total loss of full-length Cep104 had any impact on cartwheel growth. During this time, Thomas Steinacker (Raff Lab, University of Oxford) generated a CRISPR/Cas9 knock-in of Sas-6-mNeonGreen for the lab. The monomeric NeonGreen (mNG) tag is brighter than GFP and is also expected to fold faster in the embryo (Shaner *et al.*, 2013). I was therefore curious as to how the incorporation of Sas-6-mNG would compare to the Sas-6-GFP marker we had been using previously.

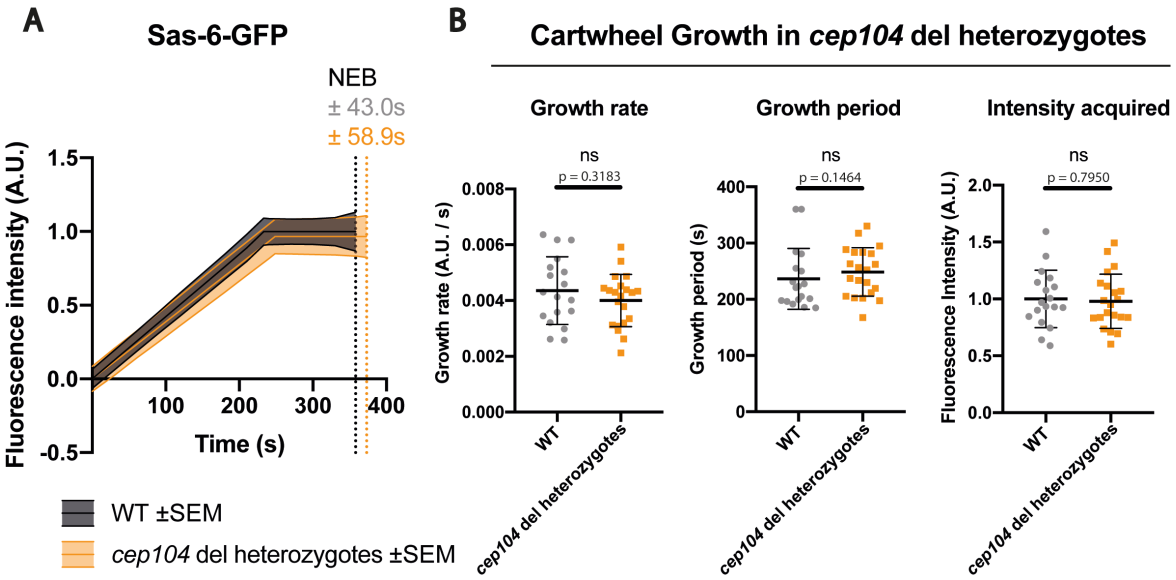


Figure 4.13 – Cartwheel growth dynamics in embryos heterozygous for the *cep104* deletion.

(A) Average curves showing the incorporation of Sas-6-GFP over S-phase of cycle 12 in WT embryos and embryos heterozygous for the *cep104* deletion. Sas-6-GFP fluorescence intensity is used here as a proxy for cartwheel size. NEB = Nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(B)** Scatter plots showing quantification of cartwheel growth rate, growth period and final cartwheel size (given by the proxy measurement of 'intensity acquired') in WT embryos or embryos heterozygous for the *cep104* deletion. Data are presented as mean ± SD; N ≥ 18 embryos, n ≥ 63 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

4.2.5.2. Cartwheel growth rate is faster in Sas-6-mNeonGreen cartwheels than Sas-6-GFP

To compare the incorporation of Sas-6 when tagged with mNG to its incorporation when tagged with GFP, I imaged embryos expressing one copy of Sas-6-mNG or Sas-6-GFP in cycle 12. Strikingly, the average growth rate of the Sas-6-mNG cartwheels was ~2x that of the Sas-6-GFP cartwheels (**Figure 4.14 A**). This difference could be due to the mNG tag being less disruptive than the GFP tag to the assembly of the Sas-6 dimers into the nine-

fold cartwheel structure. The mNG and GFP tags are a similar size, but perhaps the conformation of the mNG tag, or its distribution of hydrophobic patches, generates less steric hindrance than the GFP tag. Alternatively, since mNG folds faster than GFP (Shaner *et al.*, 2013), it could be that the available pool of Sas-6-NG with a fully-folded tag is larger than the pool of folded Sas-6-GFP. The cartwheel growth period was not significantly different between the two datasets (**Figure 4.14 B**).

A potential caveat here is that the S-phase lengths of the Sas-6-mNG dataset were slightly shorter than the Sas-6-GFP dataset (**Figure 4.14 C**). I therefore wanted to check whether the difference in growth rate that I observed was due to chance variabilities in S-phase length. At the time, I was carrying out some preliminary analysis to investigate whether a group of proteins implicated in the second phase of centriole elongation – Ana1 (Alvarez-Rodrigo *et al.*, 2021, Chang *et al.*, 2016), Ana3 and Rcd4 (Chen *et al.*, 2017, Sydor *et al.*, 2018), and Poc1 (Keller *et al.*, 2009, Meehl *et al.*, 2016, Pearson *et al.*, 2009) – might affect the growth of the cartwheel when their expression levels are reduced. I decided to carry out these experiments using both the Sas-6-GFP and Sas-6-mNG approaches, and found that, consistently, Sas-6-mNG cartwheels grew at a statistically significantly faster rate than Sas-6-GFP cartwheels (**Figure 4.14 D**). This was not always accompanied by significant differences in S-phase length (data not shown), which gives me more confidence that this increase in growth rate is due to the mNG tag.

However, it is important to note that there is also a difference between the growth rates of my two separate Sas-6-mNG control datasets (see the pale grey datasets in **Figure 4.14 D**). These data were collected from two experiments performed at different times of the year

(the faster growth rate was observed in a dataset from the summer, while a slower growth rate was observed for the control experiment performed in autumn). This should be considered when performing this growth assay, and highlights the importance of ensuring that experiments intended for comparison of growth rates and growth periods are carried out on the same day. This observation also suggests that it would be informative to image Sas-6-GFP and Sas-6-mNG incorporation side-by-side in the full Sas-6 mutant background, to properly test this phenomenon and to assess whether these possible differences in growth rates result in cartwheels acquiring different lengths at the end of S-phase.

Previous work by M. Aydogan has suggested that the GFP tag slows the incorporation of Sas-6. Since Sas-6-mNG appears to be incorporated at a faster rate, I reasoned that the use of this tag in the cartwheel growth assay might reveal more subtle differences in cartwheel growth parameters between different genotypes. I therefore resolved to carry out the experiments on cartwheel growth in Ubq-Cep104 and in *cep104* null embryos using the Sas-6-mNG fly line.

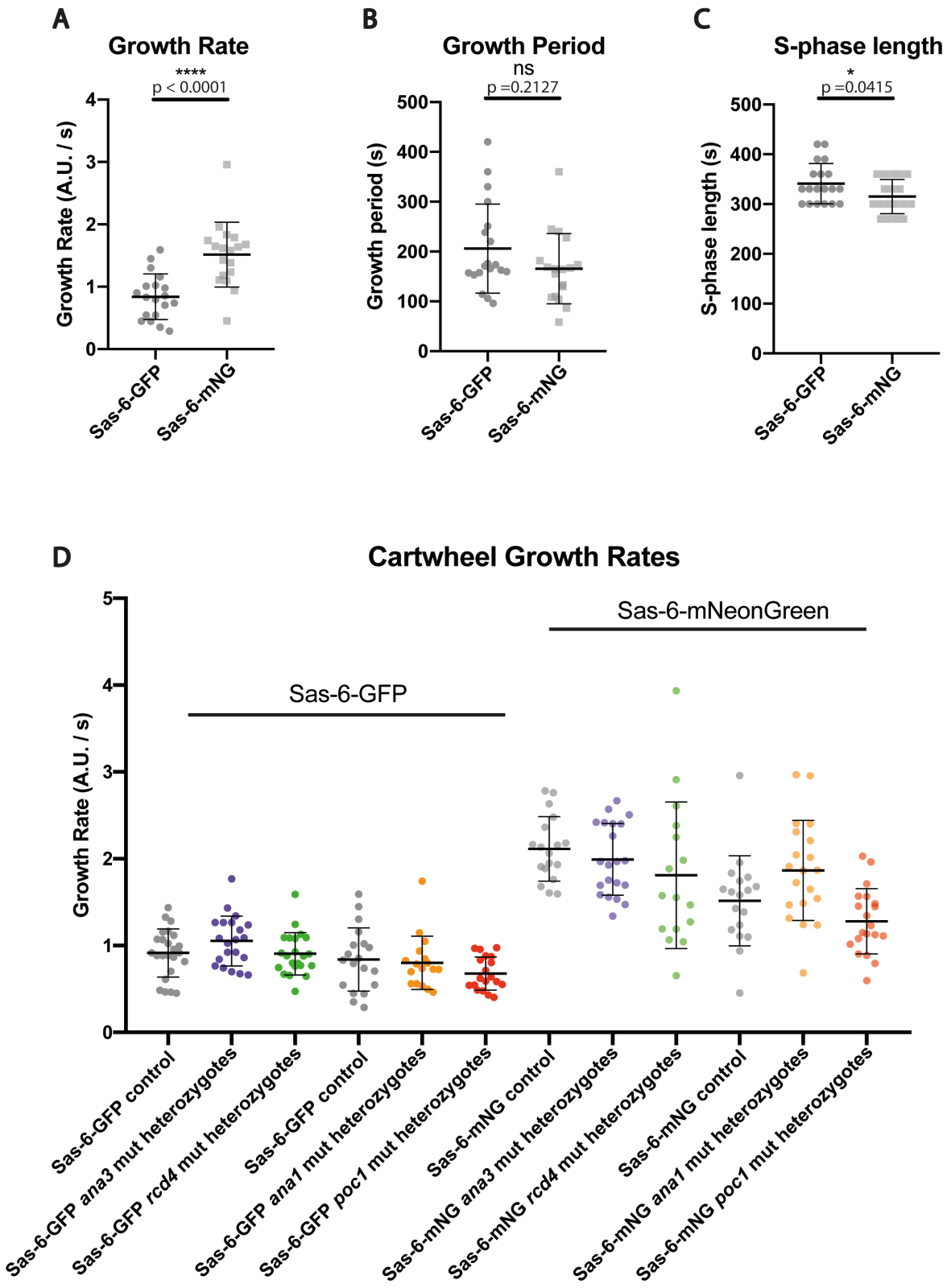

A
Growth Rate
****
p < 0.0001
Growth Rate (A.U. / s)
0
1
2
3
4
Sas-6-GFP
Sas-6-mNG
B
Growth Period
ns
p =0.2127
Growth period (s)
0
100
200
300
400
500
Sas-6-GFP
Sas-6-mNG
C
S-phase length
*
p =0.0415
S-phase length (s)
0
100
200
300
400
500
Sas-6-GFP
Sas-6-mNG
D
Cartwheel Growth Rates
Sas-6-GFP
Sas-6-mNeonGreen
Growth Rate (A.U. / s)
0
1
2
3
4
5
Sas-6-GFP control
Sas-6-GFP ana3 mut heterozygotes
Sas-6-GFP rcd4 mut heterozygotes
Sas-6-GFP control
Sas-6-GFP ana1 mut heterozygotes
Sas-6-GFP poc1 mut heterozygotes
Sas-6-mNG control
Sas-6-mNG ana3 mut heterozygotes
Sas-6-mNG rcd4 mut heterozygotes
Sas-6-mNG control
Sas-6-mNG ana1 mut heterozygotes
Sas-6-mNG poc1 mut heterozygotes

Figure 4.14 – Comparisons of Sas-6-GFP and Sas-6-mNG incorporation dynamics in cycle 12.

(A-C) Scatter plots showing quantification of cartwheel growth rate (A), growth period (B) and S-phase length (C) in Sas-6-mNG and Sas-6-GFP embryos. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs. **(D)** Scatter plots showing quantification of centriole cartwheel growth rate in further experiments employing either Sas-6-GFP (dark colours) or Sas-6-mNG (pale colours). Sas-6-mNG growth rates were consistently higher than the Sas-6-GFP rates ($p<0.05$ for all comparisons; statistics not shown on figure). This was not always accompanied by a decrease in S-phase length (data not shown), suggesting that the change in cartwheel growth rate could not simply be explained by discrepancies in cell cycle length. Data are presented as mean ± SD; N ≥ 16 embryos, n ≥ 41 centrioles per embryo.

4.2.5.3. Using Sas-6-mNeonGreen to assay cartwheel growth in different doses of Cep104

Having explored the mode of Sas-6-mNG incorporation in cycle 12, I generated flies expressing one copy of Sas-6-mNG in either a WT, *cep104* null or Ubq-Cep104 (i.e. a line overexpressing untagged Cep104) backgrounds. This allowed me to further investigate whether altering the dose of Cep104 affected the growth of the central cartwheel, by imaging Sas-6-mNG incorporation over the course of nuclear cycle 12 in each condition. Analysis revealed that altering the dose of Cep104 has no effect on the growth rate of the cartwheel (**Figure 4.15 A, B**). However, I did observe an increase in growth period in the *cep104* knockout flies (**Figure 4.15 A, B**). This increased growth period perhaps contributes to the final size of the *cep104* knockout cartwheels (given by Sas-6-mNG intensity acquired) being greater than the size reached in the WT controls. This is difficult to explain, since a change in growth period would suggest that the cartwheel somehow stops growing later in response to the altered levels of Cep104. However, I also observed that the *cep104* null embryos had significantly longer S-phase periods than the WT controls (**Figure 4.15 C**). This could be due to how I judge the start of S-phase, since I

look for the first frame in which the centrosomes of the embryo separate. It is possible that the *cep104* null embryos have a defect in centrosome separation; if their centrosomes separate earlier than the WT centrosomes, this could account for the longer S-phase of the nulls. This is perhaps a more likely explanation as to why I see a longer growth period for this genotype. It will therefore be important to repeat this experiment and see whether these differences in growth period and S-phase length are reproducible. Overall, I conclude that, in contrast to CP110/Cep97, there is little effect on cartwheel growth through changing the dose of Cep104, though further work is required to confirm this.

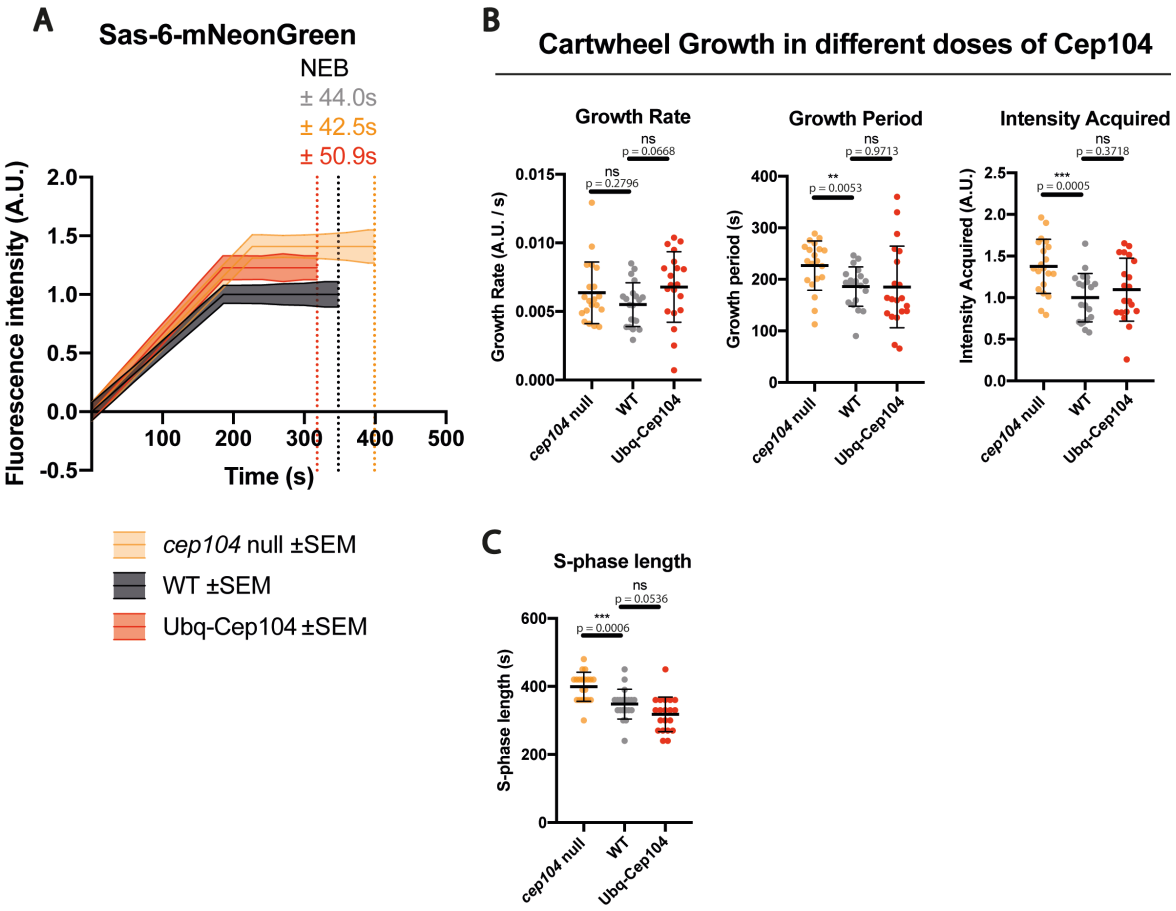


Figure 4.15 – Cartwheel growth dynamics in varying doses of Cep104.

(A) Average curves showing the incorporation of Sas-6-mNG over S-phase of cycle 12 in WT, *cep104* null and Ubq-Cep104 embryos. Sas-6-mNG fluorescence intensity is used here as a proxy for cartwheel size. NEB = Nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(B-C)** Scatter plots showing quantification of cartwheel growth rate, growth period and final cartwheel size (given by the proxy measurement of 'intensity acquired') in WT, *cep104* null and Ubq-Cep104 embryos (B). S-phase lengths are also compared (C). Data are presented as mean ± SD; N = 20 embryos, n ≥ 73 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

4.2.5.4. Preliminary experiments were carried out into the effect of Cep104 dose on centriole MT distancing

Finally, I carried out a preliminary experiment using the GFP-Cep97 distancing assay outlined in **Chapter 3**, to investigate whether the levels of Cep104 in the developing embryo might affect the growth rate and period of the centriole MTs. To do this, I generated fly lines that were either WT or heterozygous for my *cep104* deletion and also expressed Ubq-GFP-Cep97 (our marker for the distal-ends of the MTs). I then imaged embryos from these fly lines side-by-side using the Zeiss 880 LSM in order to resolve the mother and daughter centrioles. These images were used as input for the GFP-Cep97 distancing algorithm (**Chapter 3**).

However, we determined that the data generated by this assay was very noisy, suggesting that it is difficult to apply this assay to different genotypes. This could be because the GFP-Cep97 distancing algorithm requires retraining (such as that attempted for CP110-GFP distancing in **Chapter 3**) each time it is applied to a new experiment. Attempts to use the GFP-Cep97 distancing algorithm to analyse this data were therefore unsuccessful. It would be interesting to see if this assay can be improved to allow the analysis of this data, or alternatively to embark upon hand measurements of GFP-Cep97 distancing using the images that I have already collected.

4.3. Discussion

In this Chapter I have demonstrated that Cep104 is not required for the recruitment of CP110/Cep97 to centrioles in the *Drosophila* early embryo. This is inconsistent with the previous postulation that Cep104 acts as a potential recruiter of CP110/Cep97 to the growing ends of centriole MTs (Jiang *et al.*, 2012, Rezabkova *et al.*, 2016, Satish Tammana *et al.*, 2013).

Instead, it seems that, in *Drosophila* embryos at least, CP110/Cep97 are actually required for the recruitment or maintenance of Cep104 at centrioles. This conclusion is consistent with the recruitment dynamics of Cep104, since I demonstrate here that Cep104 peaks later in S-phase than Cep97. In addition, I show in this Chapter that Cep104 localises to both the mother and daughter centriole, forming a ring on the mother. This result complements studies published during my DPhil, which demonstrated for the first time that Cep104 also forms a ring on the daughter centriole (Frikstad *et al.*, 2019, Gilani, 2018). The ring is reminiscent of that formed by CP110/Cep97 at *Drosophila* (**Figure 3.1**) and human (Sonnen *et al.*, 2012) centrioles. Indeed, Frikstad *et al.* 2019 confirmed that the Cep104 ring co-localises with the CP110 ring, further supporting the idea that Cep104 is interacting with CP110/Cep97 at the centriole.

Cep104 does not appear to play a major role in daughter centriole growth control in *Drosophila* early embryos, with different doses of this protein having little discernible effect on the growth parameters of the central cartwheel. This is in stark contrast to Cep104's interaction partners CP110/Cep97, which play a surprising role in regulating the

growth rate and growth period of this structure (**Chapter 3**). This suggests that Cep104 is recruited to (or maintained at) the centriole by CP110/Cep97, but performs a distinct function. I did see an increase in cartwheel growth period in *cep104* null embryos, although this may be a result of the increased S-phase length I also observe in these embryos. It will therefore be important to repeat this experiment looking at cartwheel growth dynamics to check whether the differences in S-phase length reported here are reproducible. Additionally, it may be possible to directly assay whether Cep104 is affecting centrosome separation by repeating this experiment with a tagged version of PCNA. PCNA can be used as a marker for S-phase (Kisielewska *et al.*, 2005), so it would be interesting to see whether the length of S-phase in these *cep104* null embryos as measured via PCNA is consistent with that measured from centrosome separation (CS) and nuclear envelope breakdown (NEB).

When considering what role Cep104 might be playing at the centriole, is important to note that ciliogenesis does not occur in fly embryos. Given its postulated role in cilia formation, it is possible that Cep104 serves a more important function in ciliated cells. *Drosophila* only have two cell types that are ciliated: spermatocytes, during their development into mature sperm (Gottardo *et al.*, 2013, Riparbelli *et al.*, 2012, Tates, 1971), and sensory neurons (Kernan, 2007). It would therefore be interesting to investigate whether cilia form normally in these tissues when full-length Cep104 is absent. This would also be important for validating whether the *cep104* deletion described in this Chapter represents a loss-of-function mutation. Flies that are homozygous for the *cep104* deletion appear grossly normal in terms of their fertility and coordination, but more detailed behavioural studies will be informative to explore whether they exhibit any subtle defects.

It is possible that Cep104 acts downstream of CP110/Cep97 to allow MT elongation in centrioles that go through a second phase of elongation, since it has been suggested that Cep104's tubulin-binding TOG domain could facilitate the addition of new tubulin subunits, thus facilitating MT growth (Al-Jassar *et al.*, 2017, Louka *et al.*, 2018, Rezabkova *et al.*, 2016, Yamazoe *et al.*, 2020). It will therefore be interesting to investigate the effect of the *cep104* deletion on centriole length in cells that go through a second centriole elongation phase. Loss of CP110 or Cep97 is also known to cause dramatic elongation of centriole MTs in *Drosophila* wing disc cells (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013). I am currently collaborating with Saroj Saurya (Raff Lab, University of Oxford) to test if centrioles in *cp110* null wing discs can still form centriole MT extensions when Cep104 is absent, using electron microscopy (EM) to assess the length of the centriole MTs in *cep104* mutant, and *cp110/cep104* double mutant, wing discs. As described in this Chapter, I also attempted to assay the centriole MT growth of *cep104* mutants dynamically, using the GFP-Cep97 distancing assay developed in **Chapter 3**, but more work is required to generalise this assay to different genetic backgrounds. It is also possible to attempt measurements of the distance between the GFP-Cep97 markers by hand. This would involve me going through each image, marking the centres of each mother and daughter signal, and measuring the distance between them. A similar approach has successfully been applied in the laboratory previously, though it is highly labour-intensive (Aydogan *et al.*, 2018). This analysis, whether automated or done by hand, would ultimately provide an insight into both the growth dynamics and the final approximate length of the centriole MTs in these mutant embryos.

Finally, it is worth highlighting that although I cannot detect any full-length Cep104 protein in my mutant line, some Cep104 coding sequences remain, so there may be some truncated version of the Cep104 protein that is still expressed in my knockout flies. To further investigate this, it would also be interesting to carry out qRT-PCR, since this would allow me to assess whether the transcript itself is being produced. An alternative strategy to validate the *cep104* mutant described in this Chapter would be to ablate the whole *cep104* gene and hope that the overlapping gene has no effect. This could not be completed due to time constraints, but would allow me to compare whether I see similar results from this full gene knockout. If interesting phenotypes were observed, then I could test whether this is due to the effect from the deleted, overlapping gene by seeing if I could rescue these phenotypes by inducing the expression of WT Cep104.

It is also conceivable that Cep104 plays different roles in humans and flies. To explore whether the domains of human Cep104 are likely to be conserved in *Drosophila,* I first aligned the human and fly protein sequences using NCBI BLAST's Needleman-Wunsch algorithm (NCBI-Resource-Coordinators, 2018). Overall sequence conservation was low (around 30.4%) but the human TOG domain region specifically appeared to be well-conserved in *Drosophila* (**Figure 4.16**). I then used PSIPRED to carry out secondary structure predictions on the *Drosophila* Cep104 homologue (Buchan and Jones, 2019, Jones, 1999, Jones and Cozzetto, 2015). The predicted helices (shown in green in **Figure 4.16**) in *Drosophila* Cep104 appear to coincide with the region of this protein that aligns well with the human TOG domain (shown in orange in **Figure 4.16**). Other predicted helices also correspond well to the first coiled-coil domain of the human protein (shown in yellow in **Figure 4.16**). Excitingly, tertiary structural predictions from trRosetta (Yang *et al.*, 2020) predicted a series of six helix pairs (**Figure 4.17**), whose structure aligned well

with the structure of the TOG domain from both human and chicken Cep104 (Al-Jassar *et al.*, 2017, Rezabkova *et al.*, 2016). Overall, this suggests that the function of the TOG domain in human Cep104 is likely to be well-conserved in its *Drosophila* homologue. This also implies that I have disrupted the most functionally relevant region in my Cep104 mutant flies, since I have certainly ablated the part of the *Drosophila* homologue that aligns with the human TOG domain. It will therefore be interesting to investigate whether cilia are able to form normally in these mutants, given the speculation surrounding this domain's potential importance in facilitating axoneme extension.

The above results suggest that I have deleted the region of Cep104 that would, by analogy to studies of the human homologue, be predicted to contribute to MT polymerisation (the TOG domain). Since I do not see evidence for a shortened form of Cep104 being expressed in the *cep104* mutant fly line, I tentatively conclude that Cep104 does not play a role in CP110/Cep97 recruitment in flies. This conclusion is further supported by the fact that Cep104 recruitment depends on CP110/Cep97, and by the observation that the Cep104 peak occurs later in S-phase than the Cep97 peak.

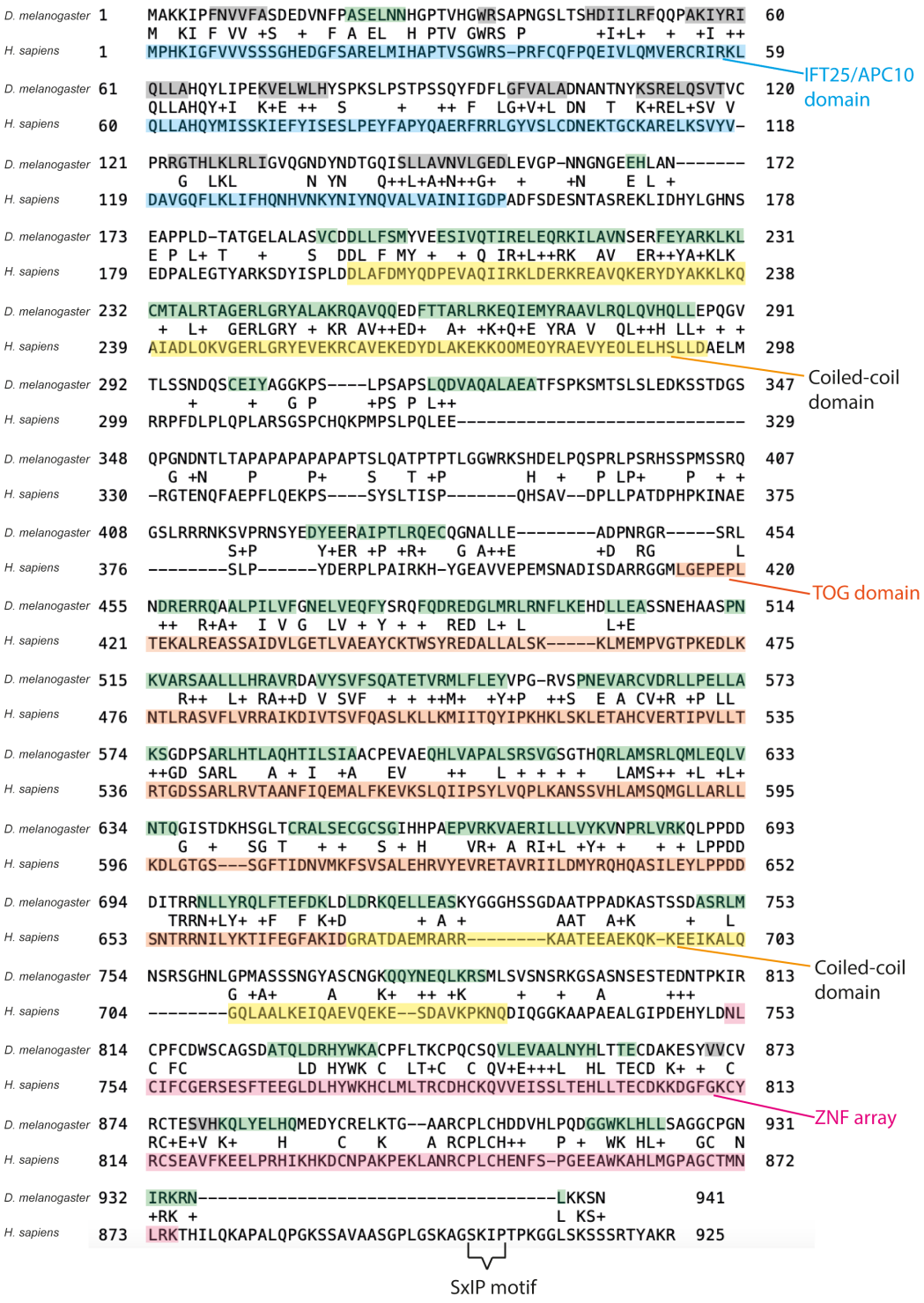


Figure 4.16 – Alignment of *D. melanogaster* and *H. sapiens* Cep104 protein sequences.

Alignment of the *D. melanogaster* and *H. sapiens* Cep104 sequences produced using BLAST (NCBI-Resource-Coordinators, 2018). The domains of the human protein are indicated as follows: IFT25/APC10 domain in blue, coiled-coil domains in yellow, TOG domain in orange, ZNF array in pink. The SxIP motif is also indicated. The *Drosophila* protein sequence is annotated to indicate predicted helices (green) and strands (grey). These secondary predictions are from PSIPRED (Buchan and Jones, 2019, Jones, 1999, Jones and Cozzetto, 2015).

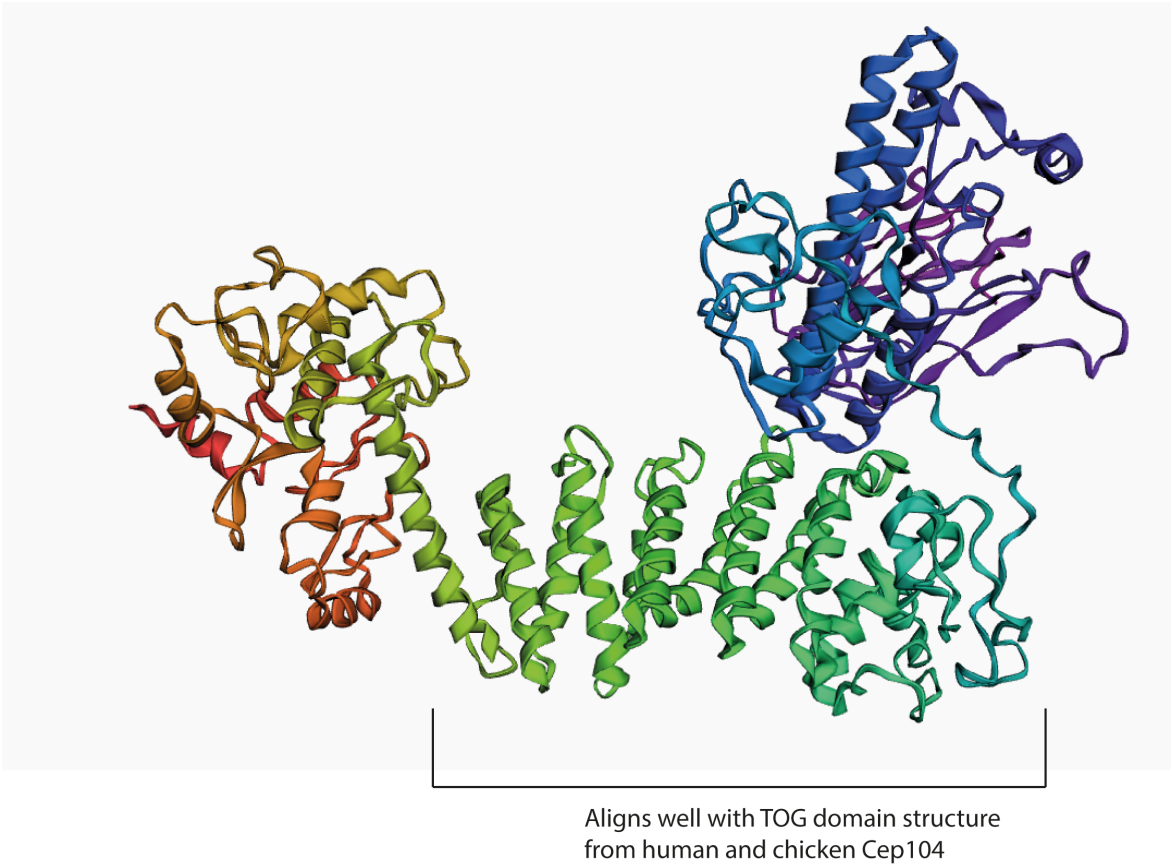


Figure 4.17 – Predicted tertiary structure of *D. melanogaster* Cep104.

Schematic showing the predicted tertiary structure of *Drosophila* Cep104, generated using trRosetta (Yang *et al.*, 2020). The bracket on the figure indicates the part of this structure that aligns well with the structure of the TOG domain from both human and chicken Cep104 (Al-Jassar *et al.*, 2017, Rezabkova *et al.*, 2016).

Chapter 5: Exploring the CP110/Cep97 Interaction Network – Klp10A

5.1. Introduction

5.1.1. CP110/Cep97 interact with the MT depolymerase Klp10A

CP110 has been shown to co-localise and interact with Klp10A, an MT depolymerase of the kinesin-13 family that also appears to play a role in the control of centriole length (Chen *et al.*, 2016, Delgehyr *et al.*, 2012, Franz *et al.*, 2013, Gottardo *et al.*, 2016). This interaction is independent of Klp10A's MT depolymerase activity (Delgehyr *et al.*, 2012). Like other kinesin-13 family members, Klp10A's motor domain has been co-opted for MT depolymerisation and does not appear to play a role in facilitating Klp10A's movement towards the ends of MTs (Moores and Milligan, 2006). Despite this, Klp10A decorates the plus-ends of cytoplasmic MTs during interphase, where it has been shown to localise via an interaction with EB1, a protein that tracks the growing ends of MTs (Goshima and Vale, 2005, Mennella *et al.*, 2005). How it localises to centrioles is less clear, and immunofluorescence experiments have not detected EB1 at centrioles in *Drosophila* (Saroj Saurya, personal communication). It has been suggested that CP110/Cep97 recruit Klp10A to their 'capping complex' at the distal-end of the centriole MTs, where it might sit and trim back any MT overgrowth (Franz *et al.*, 2013).

Another kinesin-13, Kif24, interacts with CP110/Cep97 and has been shown to stabilise CP110 at the centriole in mammals (Kobayashi *et al.*, 2011). It could be interesting to investigate whether Klp10A might play a similar role in stabilising or recruiting CP110/Cep97 to the distal-end, since I showed in **Chapter 4** that Cep104 does not appear to be important for this.

Taken together, these data suggest the possibility of some interesting recruitment dynamics, with Klp10A potentially recruiting CP110/Cep97, or vice versa. I was therefore interested in first addressing whether disruption of these proteins might affect the recruitment of the other members of this putative distal-end interaction network.

5.1.2. Klp10A is implicated in centriole length control

As well as interacting with CP110, Klp10A appears to play a role in controlling centriole length; in *Drosophila* wing discs, reduced expression of this kinesin results in dramatically over-elongated centriole MTs (Franz *et al.*, 2013). This is reminiscent of the phenotype observed in *cp110* and *cep97* null flies, though in their case the MT extensions are composed of singlet MTs whereas the extensions in the *klp10A* mutant are doublet MTs (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013). In *Drosophila* spermatocytes and in cell culture, Klp10A depletion appears to result in longer mother centrioles but shorter daughters, possibly because these elongated structures are unstable and this effect is most severely pronounced in the immature centrioles, leading to daughter centriole fragmentation (Delgehyr *et al.*, 2012, Gottardo *et al.*, 2016). Klp10A is also enriched at interphase centrosomes in *Drosophila* male germline stem cells, where its depletion causes

centriole elongation (Chen *et al.*, 2016). However, unlike in spermatocytes, this elongation effect is restricted to the centriole pair that comprise the older centrosome, leading to asymmetries in mitotic spindle formation (Chen *et al.*, 2016).

As mentioned above, it has been suggested that Klp10A acts to depolymerise any MTs that extend beyond the CP110/Cep97 cap, thus maintaining correct centriole MT length (Franz *et al.*, 2013). This could explain why depletion of this kinesin results in over-elongated centrioles. Klp10A is not essential for centriole duplication (Delgehyr *et al.*, 2012). However, it may also play some role in cartwheel size regulation, since it is possible that the cartwheels in *klp10A* mutant wing discs are longer than in WT (Franz *et al.*, 2013). Although it is challenging to confirm whether the electron-dense 'centriole' region that is measured in this study truly represents the region of the centriole containing the cartwheel, I showed in **Chapter 3** that CP110/Cep97 protein levels can influence the dynamic growth of the cartwheel. I was therefore interested in whether Klp10A might also play a role in centriole growth control during the process of centriole biogenesis, or whether it simply acts to maintain the appropriate length after the centriole growth process is complete. Consequently, I decided to investigate the effect of different Klp10A expression levels on the dynamic properties of daughter cartwheel and MT growth during S-phase.

5.1.3. Key findings in this chapter

In this Chapter, I present preliminary work showing that a hypomorphic *klp10A* mutant which is expected to exhibit reduced Klp10A expression levels displays normal recruitment of Cep97 to centrioles. I attempt to investigate Klp10A localisation at the

centriole, to test whether CP110/Cep97 might be required to recruit it to the distal-ends of the centriole MTs instead. Finally, I demonstrate that Klp10A does not play a major role in regulating the growth of the centriole cartwheel. Overall, it seems likely that Klp10A acts to maintain centriole MT length once the centrioles have finished growing to an appropriate size, but does not perform a regulatory function during centriole biogenesis.

5.2. Results

5.2.1. The Cep97 oscillation is not perturbed in embryos heterozygous for a *klp10A* mutation

I was interested in whether Klp10A might play a role in recruiting CP110/Cep97 to the centriole, or maintaining the complex there, in a similar manner to Kif24 (Kobayashi *et al.*, 2011). This might explain how CP110/Cep97 arrive at the distal-end, since Klp10A is known to co-localise here with CP110 (Delgehyr *et al.*, 2012).

To test this, I decided to use a hypomorphic *klp10A* mutant fly line in which expression of Klp10A is reduced, and males that are homozygous for this mutation are sterile due to sperm flagella defects (Delgehyr *et al.*, 2012). I used this strain, rather than attempting to generate a Klp10A knockout, because Klp10A's important roles in the mitotic spindle mean that some endogenous Klp10A expression is required to avoid lethality (Radford *et al.*, 2012). Flies homozygous for the *klp10A* mutation are uncoordinated (Delgehyr *et al.*, 2012), so I generated flies heterozygous for the *klp10A* mutation that also express one copy of Ubq-GFP-Cep97. This should still result in a reduced level of Klp10A expression in the embryo, although it should be noted that this remains to be validated. I also generated WT flies expressing one copy of Ubq-GFP-Cep97 to use as a control. I then imaged these embryos over cycle 12 and compared the parameters of the GFP-Cep97 oscillation in each case, to see whether reducing the expression levels of Klp10A affects the recruitment of Cep97.

I observed that the time taken to reach the peak of the Cep97 oscillation was not perturbed in the *klp10A* mutant heterozygotes (**Figure 5.1**). The amount of Cep97 recruited over the course of the cycle (as indicated by the intensity acquired) also appeared normal; however, there was a slight, if not significant, decrease in intensity acquired (**Figure 5.1**). It would be interesting to see if this is also the case for CP110 recruitment. I could also use a deficiency line to check if I see the same effect. Overall, my results suggest that Klp10A does not play an important role in recruiting Cep97 to the centriole. However, it should be noted that the level of Klp10A expression in *klp10A* mutant heterozygous embryos has not been quantified (see Discussion). In addition, Klp10A expression is known to be high in the syncytial embryo (Larkin *et al.*, 2020), so it could be that any reduction in Klp10A levels in flies heterozygous for the hypomorphic mutation is insufficient to perturb the Cep97 oscillation. I think it is likely that I would also see normal CP110 recruitment dynamics in *klp10A* mutant heterozygotes, since my experiments in **Chapter 3** showed that CP110/Cep97 rely on each other for their recruitment and are sensitive to changes in the dose of the other protein. Thus, I would expect that normal Cep97 recruitment is indicative of normal CP110 recruitment, though this remains to be tested.

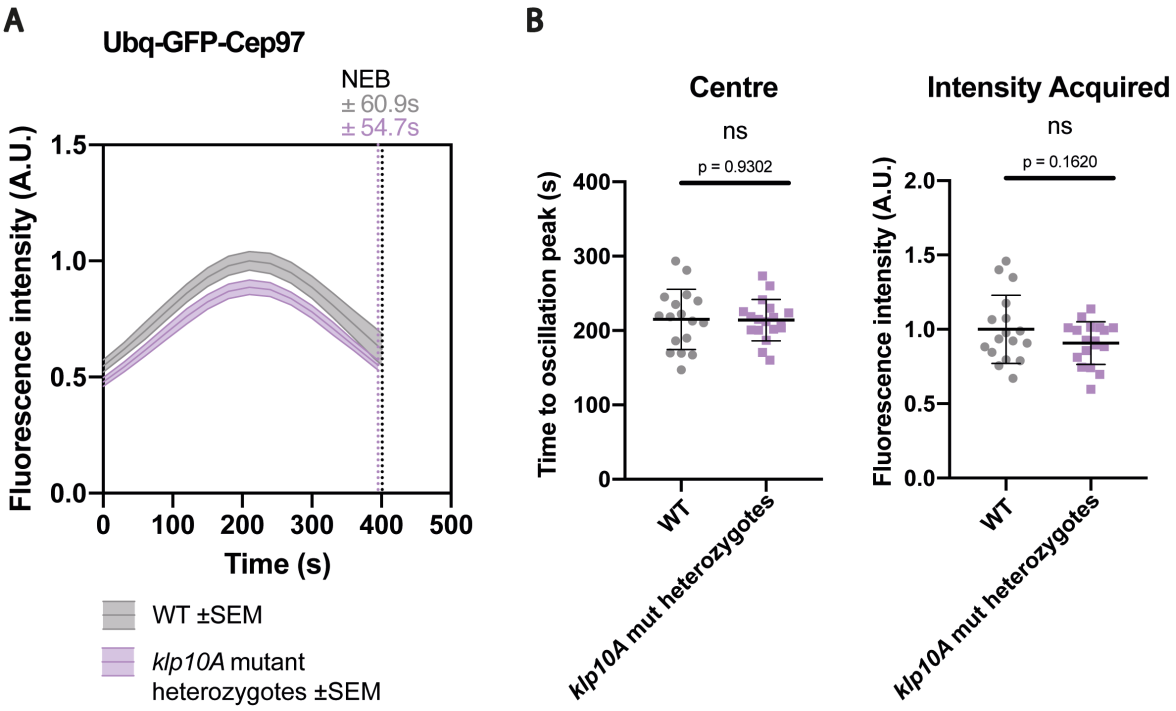


Figure 5.1 – The parameters of the GFP-Cep97 oscillation are not significantly perturbed in embryos heterozygous for the *klp10A* mutation.

(A) Average curves showing the GFP-Cep97 oscillation over S-phase of cycle 12 in WT embryos and embryos heterozygous for the *klp10A* mutation. NEB = Nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(B)** Scatter plots showing quantification of time to the peak (or *Centre*) of the oscillation and of intensity acquired in either WT embryos or embryos heterozygous for the *klp10A* mutation. Data are presented as mean ± SD; N ≥ 17 embryos, n ≥ 32 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

I next wanted to explore whether CP110/Cep97 play a role in recruiting Klp10A to the centriole, as postulated by Franz *et al.* 2013. It is challenging to visualise Klp10A in the embryo, since it is strongly recruited to centrosomes where it acts to help depolymerise the minus-ends of the MTs in the spindle (Rogers *et al.*, 2004). Indeed, the centriolar signal from fluorescently tagged versions of Klp10A in *Drosophila* S2 cells appears to be obscured by the strong signal on the centrosomes and astral MTs during mitosis, and by

the signal on the cytoplasmic MTs during interphase (Goshima and Vale, 2005, Mennella *et al.*, 2005, Rogers *et al.*, 2004). I therefore decided to use wing discs from third-instar *Drosophila* larvae, since centriole ultrastructure can be readily investigated via immunostaining in this tissue (Gartenmann *et al.*, 2017, Roque *et al.*, 2012). If successful, this experiment would allow me to probe the centriolar localisation of Klp10A in more detail, as well as allowing me to test whether this localisation is perturbed in the *cp110* or *cep97* null mutant lines. It would also allow me to investigate whether Klp10A forms a ring at the centriole in a similar manner to CP110/Cep97. I had successfully used this technique in **Chapter 4** to demonstrate that Cep104 also localises to centrioles as a ring.

I therefore examined wing discs from larvae expressing Ubq-Klp10A-GFP. I stained these wing discs using an α-GFP nanobody booster and an α-Asl antibody, before imaging them using 3D-SIM super resolution microscopy. Unfortunately, the Klp10A-GFP signal was highly dispersed (**Figure 5.2**), and did not clearly co-localise with Asl, a marker for mother centrioles (Novak *et al.*, 2014). This likely reflects the high levels of Klp10A expression in the cytoplasm and on cytoplasmic MTs.

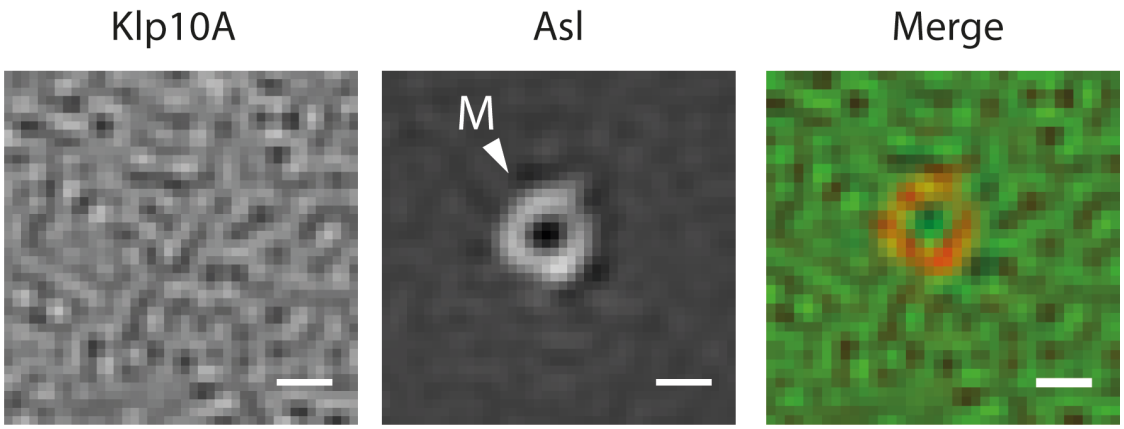


Figure 5.2 – Attempts to visualise Klp10A-GFP at wing disc centrioles.

3D-SIM image showing centrioles in fixed wing disc preparations stained for Asl (a marker of the mother centriole; Novak *et al.* 2014) and Klp10A-GFP. The first panel shows that the Klp10A-GFP signal is highly dispersed. The second panel shows the mother centriole (M), labelled by Asl. Scale bar represents 250 nm.

5.2.2. Altering the dose of Klp10A does not affect centriole cartwheel growth

To determine whether Klp10A levels might influence the growth of the centriole cartwheel, I looked at Sas-6-GFP incorporation in embryos expected to express different levels of Klp10A. I carried out two experiments: in the first, I compared cartwheel growth in WT embryos with that in embryos heterozygous for the *klp10A* mutation (Delgehyr *et al.*, 2012). I observed no significant differences in cartwheel growth rate or growth period between these embryos (**Figure 5.3 A, C**). The final size of the cartwheels at the end of the growth phase (given by the measurement of Sas-6-GFP intensity acquired) was also not significantly different between the WT embryos and those embryos that were heterozygous for the *klp10A* mutation.

I went on to also investigate the effect of overexpressing Klp10A on the growth of the cartwheel. To do this, I generated flies overexpressing an untagged version of Klp10A (Ubq-Klp10A) alongside one copy of Sas-6-GFP. I also generated WT flies expressing Sas-6-GFP as a control. I then imaged Sas-6-GFP incorporation during S-phase of cycle 12 as before, and analysed the cartwheel growth parameters in each condition. Once again, I saw no difference in the growth rate, growth period, or final size of the cartwheel (**Figure 5.3 B, D**). I therefore conclude that Klp10A does not play a role in regulating the growth of the daughter centriole cartwheel. It will be important, however, to confirm via western blot that these perturbations actually significantly alter the levels of Klp10A in these embryos (see Discussion).

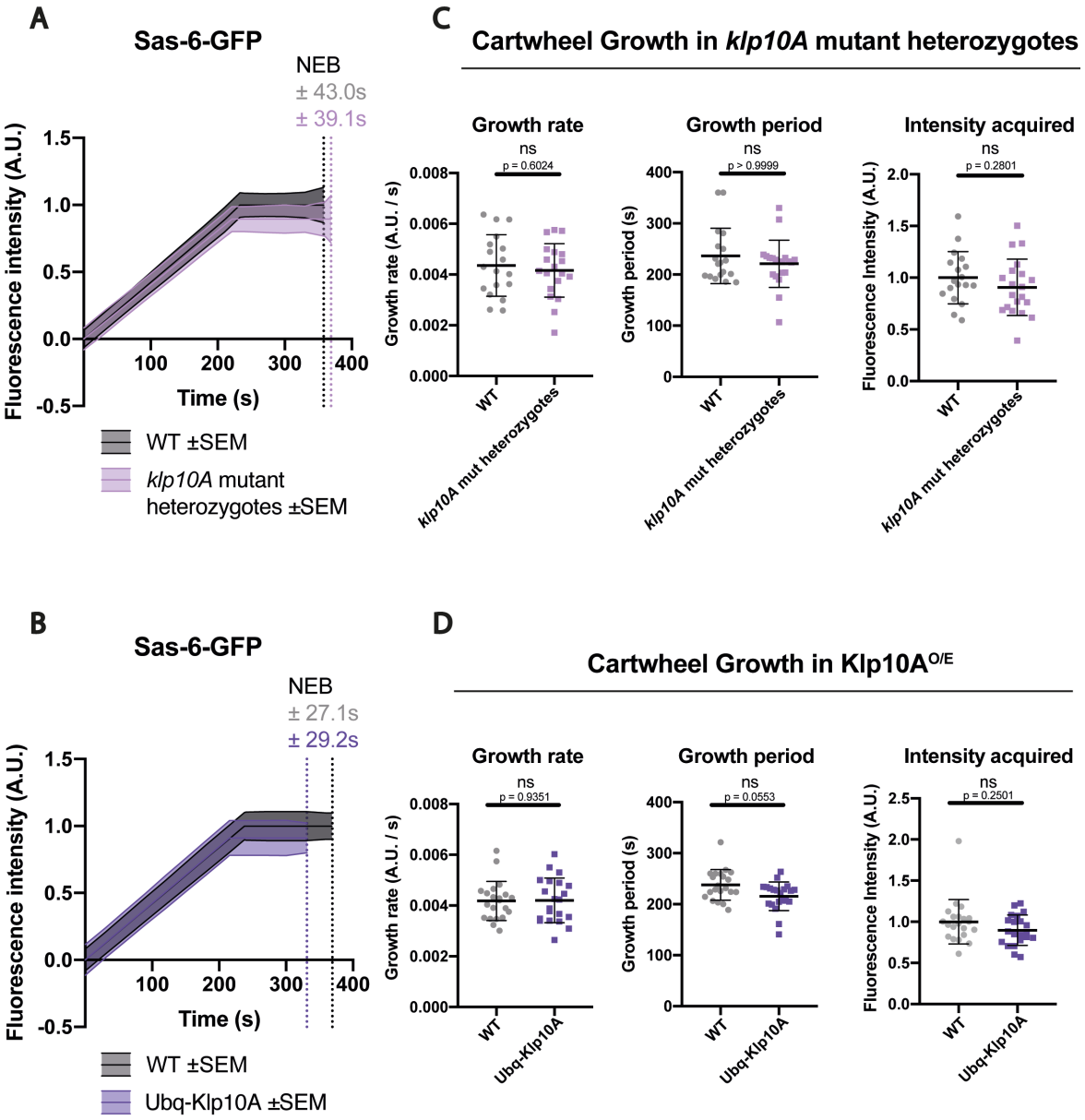


Figure 5.3 – Cartwheel growth dynamics under various expression levels of Klp10A.

(A-B) Average curves showing the incorporation of Sas-6-GFP over the course of cycle 12 in WT embryos and embryos heterozygous for the *klp10A* mutation (A) or in WT and Ubq-Klp10A (B) embryos. Sas-6-GFP fluorescence intensity is used here as a proxy for cartwheel size. NEB = Nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(C-D)** Scatter plots showing quantification of cartwheel growth rate, growth period and final cartwheel size (given by the proxy measurement of 'intensity acquired') in WT embryos and embryos heterozygous for the *klp10A* mutation (C) or in WT and Ubq-Klp10A (D) embryos. Data are presented as mean ± SD; N ≥ 18 embryos, n ≥ 51 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs.

5.2.3. Preliminary experiments were carried out on the effect of Klp10A dose on centriole MT distancing

Finally, I carried out a preliminary experiment using the GFP-Cep97 distancing assay outlined in **Chapter 3**, to investigate whether altering the dose of Klp10A in the developing embryo might affect the growth dynamics of the centriole MTs. To do this, I generated fly lines expressing Ubq-GFP-Cep97 (our marker for the distal-ends of the MTs) in either WT, *klp10A* mutant heterozygote, or Ubq-Klp10A embryos. I then imaged embryos from these fly lines side-by-side using the Zeiss 880 LSM in order to resolve the mother and daughter centrioles.

However, as discussed in **Chapter 4**, I determined that it is difficult to apply this assay to different genotypes. It would be interesting to see if this assay can be improved to allow the analysis of this data, or alternatively to embark upon hand measurements of GFP-Cep97 distancing using the images that I have already collected. This distancing analysis will be important for testing whether Klp10A exerts its effects on centriole MTs during their growth, or whether it exclusively acts to maintain centriole MT length once the daughters have reached the appropriate size.

5.3. Discussion

This Chapter represents some preliminary experiments into the role of Klp10A at the centriole, which first aimed to investigate whether Klp10A participates in CP110/Cep97 recruitment. Using a hypomorph for *klp10A,* I have shown that embryos where Klp10A expression is expected to be reduced exhibit normal recruitment of Cep97 to the centriole. It will be interesting to check whether CP110 recruitment is also normal in these mutants, though the co-dependency of these proteins for each other's proper recruitment (**Chapter 3**) suggests that this is likely to be the case. This result suggests that Klp10A is not important for the recruitment of CP110/Cep97. However, it is possible that any reduction in Klp10A expression here is not sufficient to produce a noticeable phenotype, since expression of this kinesin in the embryo is very high (Larkin *et al.*, 2020).

It would therefore be interesting to assay how reduced Klp10A expression is in the embryos of the *klp10A* hypomorphs via western blotting. Due to the coronavirus pandemic, this was not possible; our laboratory was shut down and the *klp10A* mutant fly line in our stocks died over this period. However, it will be important to order this line again and carry out this experiment, since different tissues may be affected by the mutation to different extents. For example, the reduction in Klp10A expression in these mutants is known to be more dramatic in brains than in testes (Delgehyr *et al.*, 2012). Another way to reduce the levels of Klp10A in the embryo would be depletion via RNAi, through which it may be possible to elicit a more significant reduction in expression. This has been applied successfully in other insect cells to study this protein (Chen *et al.*, 2016, Delgehyr *et al.*, 2012). It might also be possible to use a degron system (Housden *et al.*, 2017). These approaches would be particularly valuable in the case of Klp10A, since loss-of-function

mutants are lethal due to the important role of this protein during mitosis (Radford *et al.*, 2012).

I also wanted to investigate whether CP110/Cep97 might be recruiting Klp10A to the centriole, as postulated by Franz *et al.* 2013. This involved attempting to visualise Klp10A localisation specifically at the centriole MTs, which is challenging. My attempt to do this in wing discs, a technique I had successfully applied in **Chapter 4** for the study of Cep104, was unsuccessful here. This might be due to high background signal, or it could simply be due to the dispersed nature of Klp10A-GFP in the embryo. Attempts to optimise the staining protocol, in concert with the Cep104-GFP wing discs, might improve this. Alternatively, it may be fruitful to explore using the centrioles of the *Drosophila* testes, since these are larger than the wing disc centrioles and Klp10A has successfully been visualised here previously (Delgehyr *et al.*, 2012). Intriguingly, staining of these centrioles for CP110 has revealed that CP110 may localise at both the proximal- and distal-ends of the centriole, though the latter signal is more intense (Franz *et al.*, 2013). Klp10A has also been reported to localise at both ends of the centriole (Delgehyr *et al.*, 2012). Ideally, successful staining of Klp10A in wing discs would have allowed me to test whether Klp10A localises to centrioles as a ring in a similar manner to CP110, Cep97 and Cep104. However, successful staining in spermatocytes would still allow me to address whether Klp10A localises to centrioles in the absence of CP110/Cep97.

I also investigated whether Klp10A plays a role in centriole cartwheel growth by measuring Sas-6-GFP incorporation in different doses of Klp10A expression. I was curious about this because *klp10A* mutant flies possibly exhibit slightly elongated

cartwheel structures, representing a similar phenotype to that observed in *cp110* null flies (Franz *et al.*, 2013). However, I saw no significant effect on cartwheel growth dynamics from perturbing the levels of Klp10A in the embryo. It is important to note though, as explained above, that Klp10A is already expressed at high levels in the embryo, so decreasing its expression levels may simply not be sufficient to affect the growth of the cartwheel. Equally, increasing the expression of Klp10A further may have little effect if its levels are already high. Overall, however, my data indicates that Klp10A does not play a major part in regulating the initial growth of the daughter centriole cartwheel. This is a significant result, as it suggests that Klp10A largely acts to maintain the length of centrioles that are already fully-grown.

It would be interesting to assess whether the elongated centrioles observed in *klp10A* mutant wing discs by Franz *et al.* 2013 represent mothers, daughters, or both; this could be assayed using correlative light and electron microscopy (CLEM) for markers of the mother centriole. If only the mother centrioles exhibit centriole MT elongation, this would support the idea that Klp10A only affects the length of mature centrioles, not growing daughters. Indeed, one study suggested that mother centrioles, not daughter centrioles, are elongated in *klp10A* mutants (Gottardo *et al.*, 2016). However, it has been postulated that this is because Klp10A promotes elongation in both mothers and daughters, with this elongation destabilising the daughters more readily than the mothers, thus resulting in daughter fragmentation (Delgehyr *et al.*, 2012, Gottardo *et al.*, 2016). If this is the case, this fragmentation does not appear to occur during the growth of the daughters in S-phase, since I found that the amount of Sas-6-GFP recruited to WT or *klp10A* mutant daughters seems consistent.

It could be that Klp10A does not influence cartwheel growth but does affect the growth of the centriole MTs, and that this growth is not coupled to the growth of the cartwheel, as postulated in **Chapter 3**. To test this possibility, it will be important to optimise the GFP-Cep97 distancing algorithm to enable assaying of centriole MT dynamics. This would be key to determining if the MT extensions caused by depleting Klp10A (Franz *et al.*, 2013, Gottardo *et al.*, 2016) are formed during centriole biogenesis or whether they extend after the centriole has finished growing. Overall, given the results from this Chapter, and our data suggesting that MT growth dynamics track the cartwheel (**Chapter 3**), I suspect the most likely explanation is that Klp10A does not influence centriole MT growth during the process of centriole biogenesis, but does play a role in maintaining proper centriole length once centrioles reach their full size.

Chapter 6: CP110/Cep97 and the Cell Cycle

6.1. Introduction

6.1.1. CP110/Cep97 levels and recruitment dynamics vary across different cell cycle stages

In human cells, CP110 levels are regulated in a cell cycle-dependent manner; levels are low during G_1 and begin to increase at the onset of S-phase (Chen *et al.*, 2002). CP110 is a target of the $SCF^{CyclinF}$ ubiquitin ligase complex, which ubiquitylates CP110 in G_2 and thus targets it for degradation (**Figure 6.1**; D'Angiolella *et al.*, 2010). This effect is counteracted by USP33, a deubiquitylase that associates with centrosomes during S- and G_2-phase to remove ubiquitin groups from CP110, thus protecting CP110 from the proteasome degradation pathway (**Figure 6.1**; Li *et al.*, 2013). Less is known about the regulation of Cep97, though it also appears to be the target of ubiquitin ligases (Nagai *et al.*, 2018). Since CP110 and Cep97 are dependent on each other for recruitment to the centriole and appear to affect each other's stability (Spektor *et al.*, 2007, Zibrat, 2019), it is also possible that any effects on CP110 would in turn affect Cep97.

Our own work has also shown that the peak of the CP110/Cep97 oscillation correlates strongly with S-phase length (**Figure 3.9**), so it seems reasonable to suggest that the cell cycle influences the recruitment of these proteins to centrioles. This might be mediated

through unknown kinases and phosphatases that are active at different points in the cell cycle, and their effect on the phosphorylation state of CP110/Cep97. Alternatively, the oscillatory recruitment profile that we observe could be an indirect consequence of a change in cytoplasmic levels of CP110/Cep97 over the course of the cell cycle. However, this seems unlikely given the short duration of the cell cycles in the early *Drosophila* embryo.

CP110 itself has also been implicated in cell and centrosome cycle regulation. CP110 was initially identified as a CDK substrate (**Figure 6.1**; Chen *et al.*, 2002), and two of its predicted CDK phosphorylation sites have subsequently been shown to play important roles in centrosome clustering to prevent anaphase catastrophe (Hu *et al.*, 2015a, Hu *et al.*, 2015b). CP110 does not seem essential for centriole biogenesis, since knock-outs in flies and mice still have centrioles (Franz *et al.*, 2013, Yadav *et al.*, 2016). However, depletion of CP110 in human cells can inhibit centriole over-duplication induced by S-phase arrest (Chen *et al.*, 2002) or Plk4 overexpression (Kleylein-Sohn *et al.*, 2007). Curiously, overexpression of CP110 can promote the formation of supernumerary centrioles (Bijnsdorp *et al.*, 2016, D'Angiolella *et al.*, 2010, Li *et al.*, 2013). These studies further support the idea that regulation of CP110 levels is important for proper cell function, and indicate a potential complex relationship between CP110 and the core centriole duplication machinery.

6.1.2. CP110/Cep97 have been identified as putative PLK1 interactors

Recently, GFP-PLK1 pulldown assays combined with mass spectrometry have identified CP110 and Cep97 as potential Polo-like kinase 1 (PLK1) binding partners in human cells (Legg, 2019). PLK1 is homologous to *Drosophila* Polo, a mitotic kinase that associates with centrosomes throughout the cell cycle in syncytial fly embryos (Llamazares *et al.*, 1991, Moutinho-Santos *et al.*, 1999). Polo/PLK1 has several important roles during the cell division cycle and also specifically at centrosomes during the centrosome cycle (reviewed in Pintard and Archambault, 2018). Most notably, it contributes to centrosome maturation, which is the expansion of the pericentriolar material (PCM, which nucleates spindle MTs) at the onset of mitosis (Alvarez-Rodrigo *et al.*, 2019, Conduit *et al.*, 2014a, Dobbelaere *et al.*, 2008, Feng *et al.*, 2017, Goshima *et al.*, 2007, Lane and Nigg, 1996). PLK1 is also required for centriole disengagement and for daughter centriole maturation (Kim *et al.*, 2015a, Kong *et al.*, 2014, Loncarek *et al.*, 2010, Novak *et al.*, 2016, Shukla *et al.*, 2015, Tsou *et al.*, 2009, Wang *et al.*, 2011).

Intriguingly, human PLK1 is required for centriole elongation during G_2 (Kong *et al.*, 2020), and it has recently been shown in the Raff laboratory that the centriole protein Ana1 helps to recruit Polo to the distal-end of centrioles to promote this second phase of centriole elongation (Alvarez-Rodrigo *et al.*, 2021). I wondered therefore whether Polo might also interact with CP110/Cep97 to help recruit it to the centriole distal-ends. Alternatively, Polo binding to CP110/Cep97 might indicate that CP110/Cep97 are being phosphorylated by Polo, and a change in their phosphorylation state might mediate their oscillatory recruitment to, and dissociation from, centrioles. Polo phosphorylation sites are

challenging to identify in candidate Polo substrates since a range of motifs appear to be targeted (Santamaria *et al.*, 2011). However, Polo/PLK1 kinases have a conserved Polo-box (PB) domain; this domain is thought to target this family of kinases to specific regions of the mitotic cell by binding to proteins that have been phosphorylated at Ser-pSer/pThr motifs (Elia *et al.*, 2003a, Elia *et al.*, 2003b, Lee *et al.*, 1998, Liu *et al.*, 2004, Reynolds and Ohkura, 2003, Seong *et al.*, 2002, Song *et al.*, 2000). In previous studies from our laboratory, it was possible to essentially block the interactions of Spd-2 and Ana1 with Polo by substituting all their potential S-S/T motifs to T-S/T (Alvarez-Rodrigo *et al.*, 2019, Alvarez-Rodrigo *et al.*, 2021). These substitutions did not appear to perturb other aspects of protein function. I therefore decided to take a similar approach to assess whether the ability of CP110 and/or Cep97 to recruit Polo via its PB domain was important for their function.

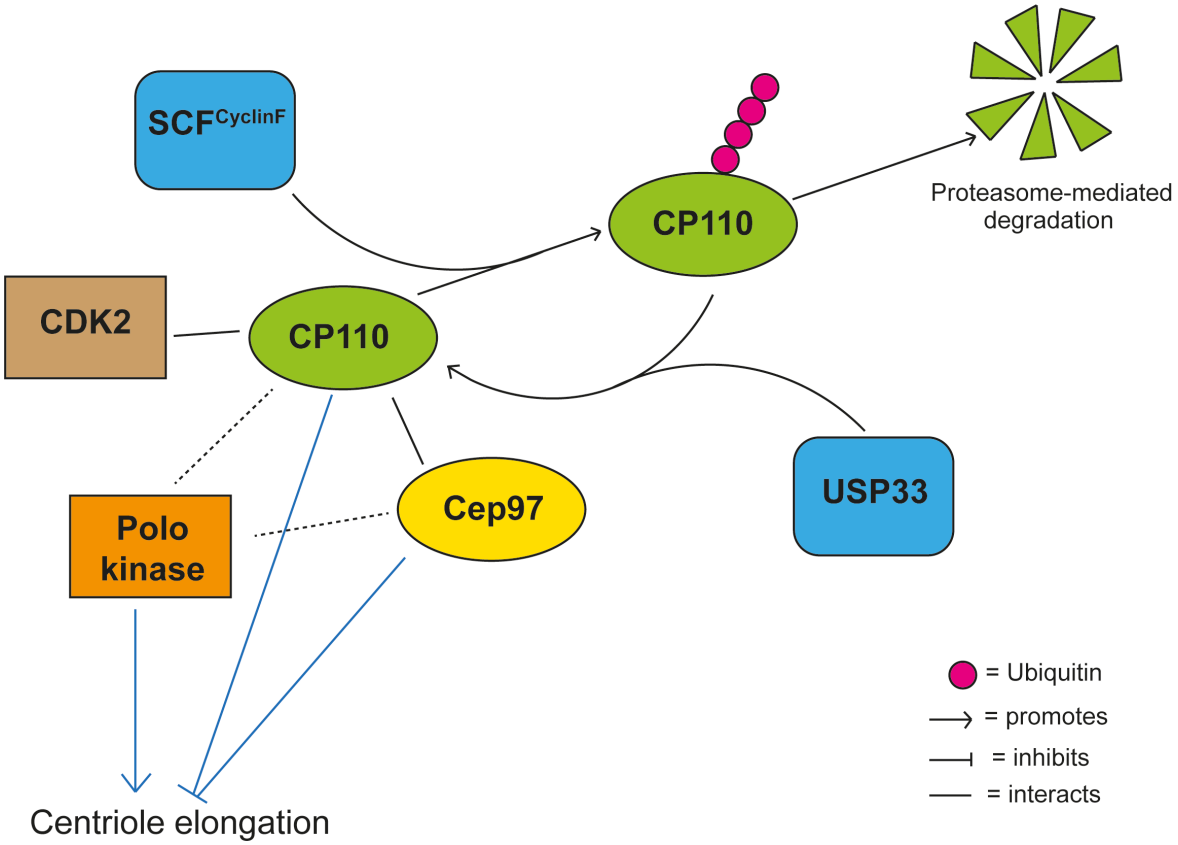


Figure 6.1 – Schematic summarising the CP110/Cep97 regulatory network.

Schematic showing the putative CP110/Cep97 interaction network as described in the above sections. Confirmed interactions are indicated with straight lines, potential interactions with dashed lines. CP110 interacts with Cep97, and both these proteins act to inhibit over-elongation of centrioles (Franz *et al.*, 2013, Kohlmaier *et al.*, 2009, Schmidt *et al.*, 2009, Spektor *et al.*, 2007, Tang *et al.*, 2009). CP110/Cep97 have recently been identified as potential binding partners of Polo (Legg, 2019). Polo recruitment to the centriole is required for centriole elongation in G_2 (Alvarez-Rodrigo *et al.*, 2021), and its human homologue (PLK1) can promote centriole elongation when overexpressed in its active form (Kong *et al.*, 2020). CP110 is also a known substrate of another kinase, CDK2 (Chen *et al.*, 2002, Hu *et al.*, 2015a, Hu *et al.*, 2015b). Finally, CP110 levels in the cell are regulated by $SCF^{CyclinF}$, which adds ubiquitin chains to CP110 to target it for degradation (D'Angiolella *et al.*, 2010), and USP33, which counteracts this by deubiquitylating CP110 and thus rescuing it from degradation (Li *et al.*, 2013).

6.1.3. Key findings in this chapter

In this Chapter, I demonstrate that inhibiting Polo binding to CP110 has no effect on the recruitment dynamics of these proteins at the centriole. In collaboration with Saroj Saurya (Raff Lab, University of Oxford), we demonstrate that Polo binding is also not required for CP110's capping function at the ends of centriole MTs, or for its surprising role in influencing cartwheel growth, which we uncovered in **Chapter 3**. I go on to investigate whether phosphorylation in general is likely to play a role in the CP110 centriolar recruitment process by deleting two regions of CP110 that contain a high density of potential Polo phosphorylation sites. Injection of this mutant construct showed that, despite these large disruptions to the protein sequence, CP110 was still recruited to centrioles in an oscillatory manner. These results lead me to suggest that any interaction between Polo and CP110/Cep97 does not appear to be important for the proper recruitment of these proteins. Indeed, I also show that the Cep97 oscillation persists under S-phase arrest conditions, suggesting that the influence of cell cycle factors is not crucial for these dynamics (though they probably play a role in the entrainment of the oscillation). Finally, I show that the CP110-GFP signal is lost from centrioles under mitotic arrest induced by colchicine injection, suggesting that CP110 cannot bind to centrioles during mitosis.

6.2. Results

6.2.1. Polo binding is not important for the recruitment of CP110/Cep97

6.2.1.1. Generation of fly lines expressing CP110/Cep97 mutants that cannot bind Polo

Ines Alvarez Rodrigo, a former DPhil student in the Raff Lab, had recently developed a successful approach for perturbing Polo binding to Spd-2 and Ana1 by mutating all the potential Polo-box binding motifs (S-S/T) to T-S/T (Alvarez-Rodrigo *et al.*, 2019, Alvarez-Rodrigo *et al.*, 2021). She had also designed CP110 and Cep97 constructs where all these motifs were similarly mutated (**Figure 6.2 A**). To initially test if CP110/Cep97 dynamics might be disrupted if they cannot bind Polo, I cloned these mutant CP110/Cep97 constructs into plasmids which incorporated a GFP C-terminal tag. I then sent these vectors to the Cambridge Fly Facility for injection into flies and transgenic integration via P-element insertion.

Initial imaging of these constructs in WT embryos revealed that these mutant fusion proteins were still recruited to centrioles (**Figure 6.2 B, C**) in an apparently oscillatory manner, with signal intensity at centrioles increasing during S-phase before decreasing as embryos entered mitosis (data not shown, see below).

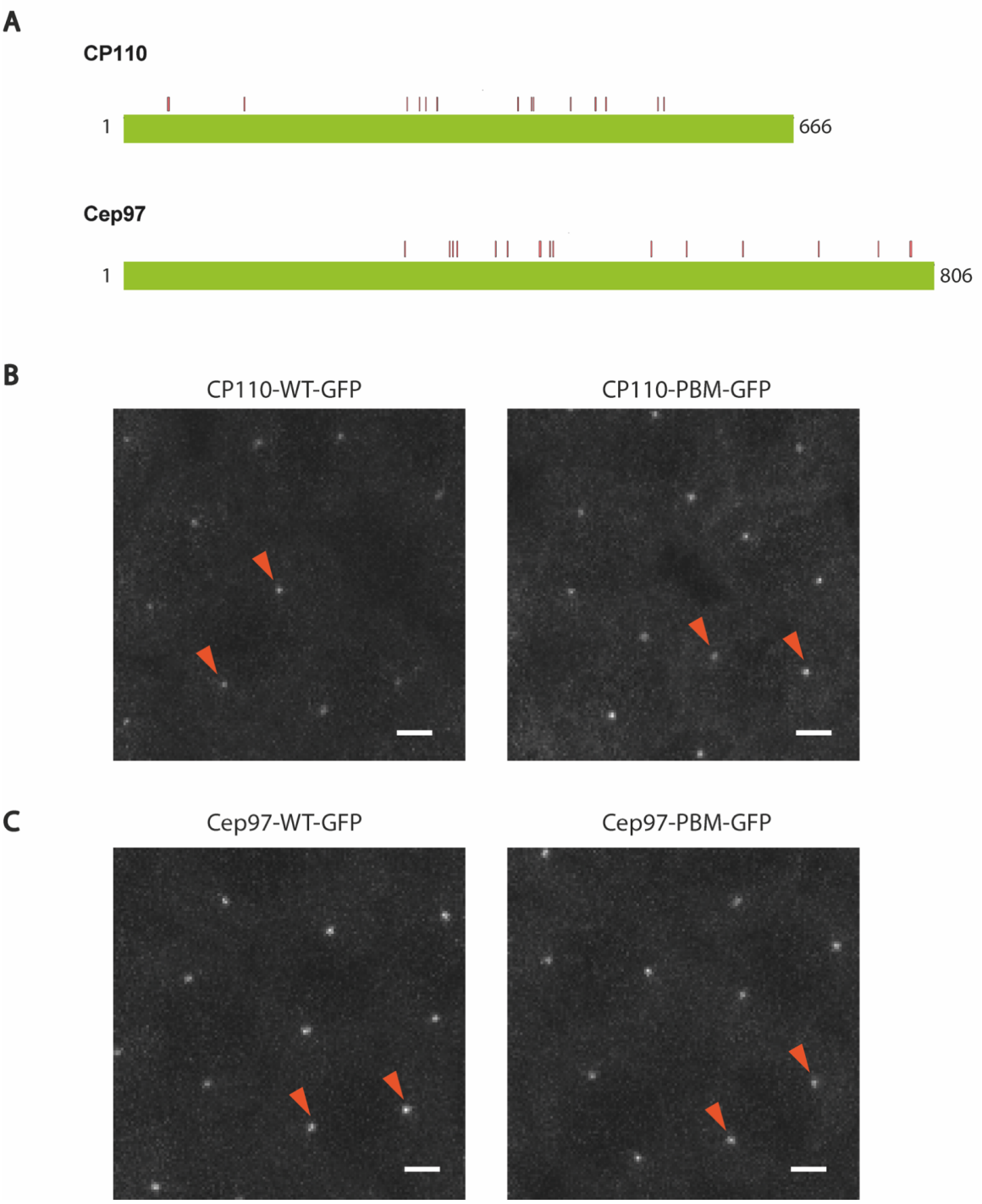


Figure 6.2 – Localisation of CP110 and Cep97 mutants that cannot bind Polo.

(A) Schematic showing centriolar proteins CP110 and Cep97 (green boxes; numbers indicate the length of the amino acid sequences) with the locations of potential Polo binding sites (S-S/T motifs) indicated by orange boxes. These sites were mutated (to T-S/T) to generate tagged mutants of CP110 (CP110-PBM-GFP) and Cep97 (Cep97-PBM-GFP) that cannot bind Polo. **(B-C)** Spinning disc confocal images showing CP110-WT-GFP and CP110-PBM-GFP (B) or Cep97-WT-GFP and Cep97-PBM-GFP (C) localisation during cycle 12 in WT embryos. Note that these embryos also express endogenous CP110/Cep97. In both cases, the GFP signal is observed at centrioles. Example centriole pairs are indicated with orange arrowheads. Scale bar represents 3 μm.

To investigate their recruitment in more detail, I decided to image embryos expressing these tagged Polo-binding mutants in cycle 12, and compare their recruitment dynamics to those of WT CP110 and Cep97. Due to time constraints with regards to *Drosophila* crossing schemes, I decided to focus on CP110. From here on, I will refer to the WT Ubq-CP110-GFP line as '***CP110-WT-GFP'*** and to the tagged **P**olo-**b**inding **m**utant line as ***'CP110-PBM-GFP'***.

First, I confirmed via western blot that the expression levels of CP110-WT-GFP and CP110-PBM-GFP were similar (**Figure 6.3**). This was important to test, to rule out the possibility that any phenotypes I might observe are a result of differences in expression level. This is particularly relevant in the case of CP110, since our work in **Chapter 3** revealed that altering the global levels of this protein affects the recruitment dynamics of Cep97, centriole cartwheel growth dynamics and the centriolar recruitment of Plk4.

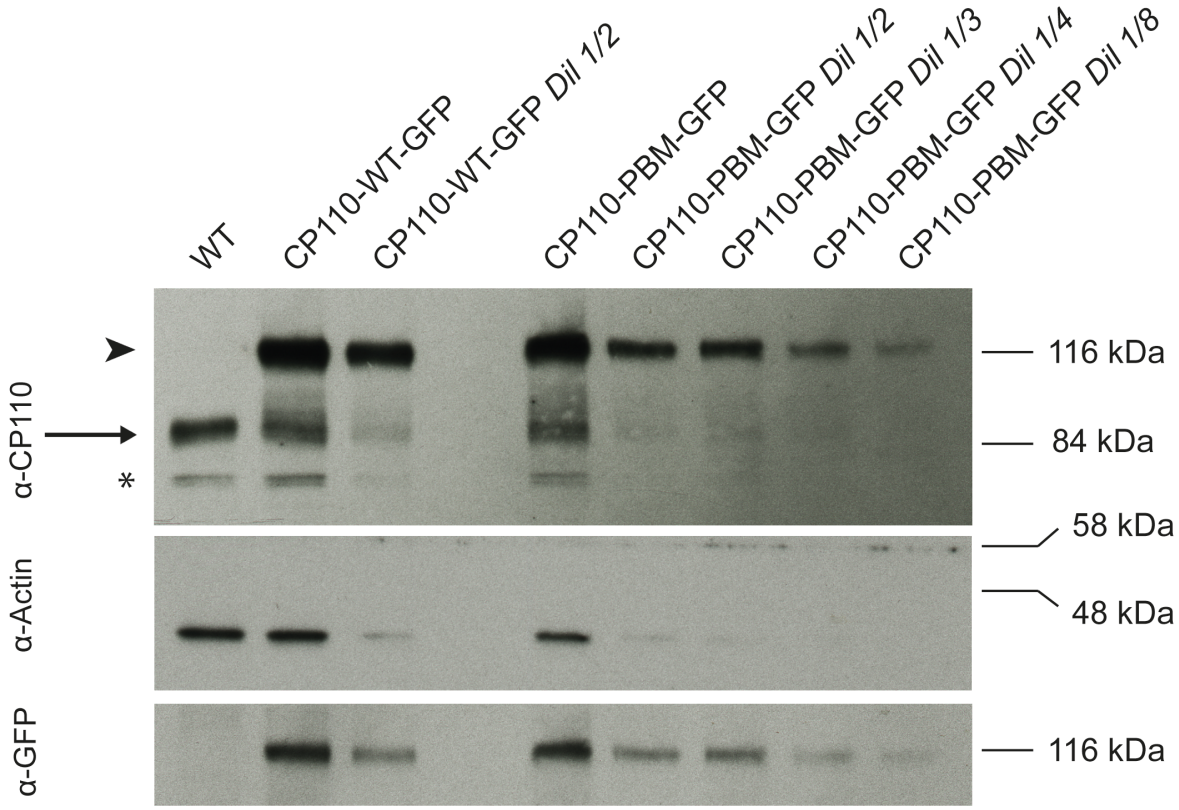


Figure 6.3 – Western blot to compare expression levels of CP110-WT-GFP and CP110-PBM-GFP.

Western blot comparing the level of CP110 expression in WT embryos, embryos expressing one copy of CP110-WT-GFP, and embryos expressing one copy of CP110-PBM-GFP. The endogenous CP110 band is indicated with an arrow. The band corresponding to the GFP-tagged protein is indicated with an arrowhead. Non-specific bands are indicated with *. The results demonstrate that expression of both endogenous CP110 and CP110-GFP is similar in both the CP110-WT-GFP and the CP110-PBM-GFP lines. Actin is shown as a loading control. N = 2 technical repeats.

6.2.1.2. CP110-GFP mutants that cannot bind Polo are still recruited to the centriole, where they exhibit normal dynamics

I went on to image WT embryos expressing one copy of either CP110-WT-GFP or CP110-PBM-GFP over the course of cycle 12. I was then able to follow the change in GFP fluorescence intensity at the centriole over S-phase. Both CP110-WT-GFP and CP110-

PBM-GFP exhibited oscillatory dynamics at the centriole, with levels peaking in S-phase. Fitting Lorentzian non-linear regressions (as described in **Chapter 3**) to the data from each embryo allowed me to quantify the time to peak and the GFP intensity acquired for each dataset. I observed that the parameters of the CP110-PBM-GFP oscillation were not significantly different from those of the CP110-WT-GFP oscillation (**Figure 6.4 A, B**). I also carried out a correlation analysis, to test whether the CP110-PBM-GFP oscillation peak still correlated with S-phase length. We had shown this was consistently the case with WT versions of CP110/Cep97 (**Figure 3.9**). This was important to check, since it was possible that an interaction with cell cycle regulators such as Polo is not required for CP110/Cep97's recruitment, but that it might play a role in entraining the period of the CP110 oscillation to the cell cycle. However, I found that the peaks of both the CP110-PBM-GFP (**Figure 6.4 C**) and the CP110-WT-GFP (data not shown) oscillations correlated strongly with S-phase length. For completeness, I also corrected the time to oscillation peak to account for differences in S-phase length between the two datasets (as discussed in **Chapter 3**) since the peaks in both datasets were correlated to S-phase length (**Figure 6.4 D**).

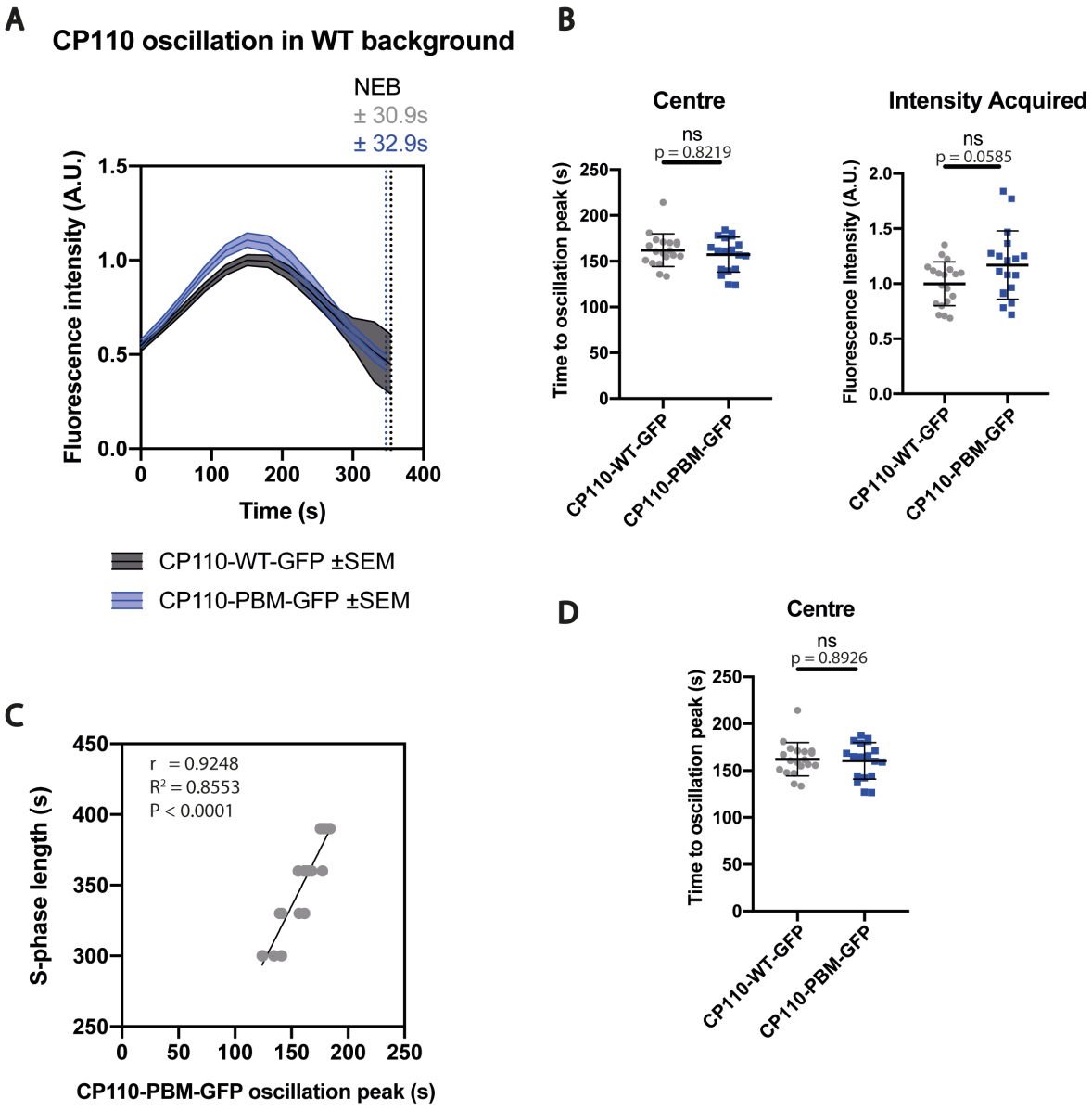


Figure 6.4 – The CP110-PBM-GFP oscillation in WT embryos.

(A) Average curves showing the CP110-GFP oscillation over S-phase of cycle 12 in CP110-WT-GFP and CP110-PBM-GFP embryos. NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(B)** Scatter plots showing quantification of time to the peak (or *Centre*) of the oscillation and of intensity acquired in CP110-WT-GFP and CP110-PBM-GFP embryos. Data are presented as mean ± SD; N ≥ 18 embryos, n ≥ 30 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs. **(C)** Graph showing that there is a positive correlation between the time of the CP110-PBM-GFP peak and S-phase length in cycle 12. *r*, R^2 and *p*-values are shown on the graph and a linear regression is plotted. **(D)** Centres from (B) corrected for S-phase length, to account for differences in S-phase length between the datasets.

It was possible that CP110-PBM-GFP might be interacting with endogenous, untagged CP110 in the embryo, thus facilitating its recruitment to the centriole. To exclude this possibility, I generated flies homozygous for a genetic ablation of CP110 (Franz *et al.*, 2013) that express one copy of either CP110-WT-GFP or CP110-PBM-GFP. Embryos from these flies should therefore lack endogenous WT CP110. I imaged the centriolar recruitment of these proteins over the course of cycle 12. Again, I observed no significant difference between the oscillation parameters of CP110-WT-GFP and CP110-PBM-GFP, even in the absence of endogenous CP110 (**Figure 6.5 A, B, D**). Furthermore, the peak of the CP110-PBM-GFP oscillation correlated with S-phase length (**Figure 6.5 C**).

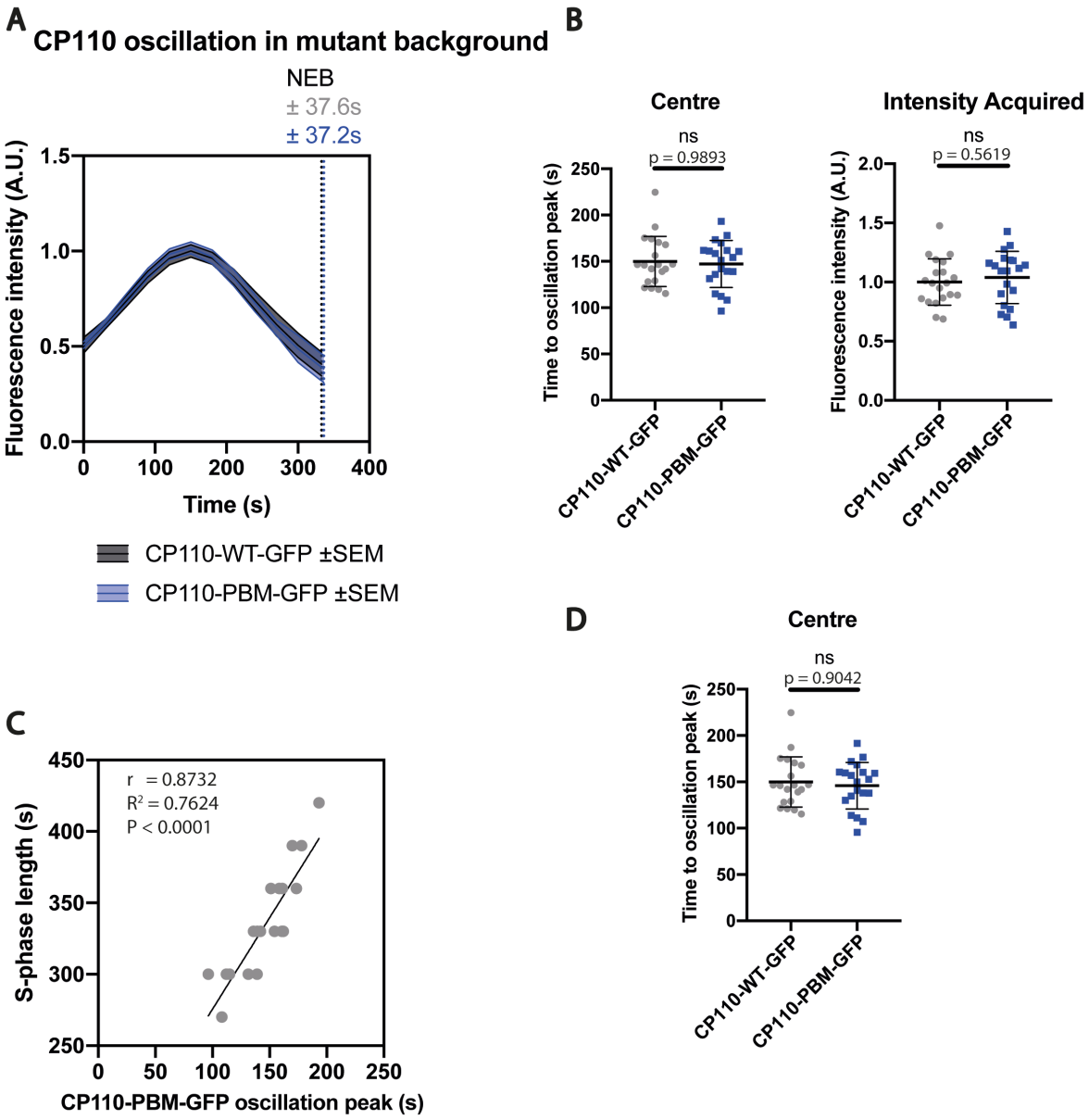


Figure 6.5 – The CP110-PBM-GFP oscillation in embryos lacking endogenous CP110.

(A) Average curves showing the CP110-GFP oscillation over S-phase of cycle 12 in CP110-WT-GFP and CP110-PBM-GFP embryos that do not express endogenous CP110. NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SEM. **(B)** Scatter plots showing quantification of time to the peak (or *Centre*) of the oscillation and of intensity acquired in CP110-WT-GFP and CP110-PBM-GFP embryos that do not express endogenous CP110. Data are presented as mean ± SD; N ≥ 18 embryos, n ≥ 51 centrioles per embryo. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs. **(C)** Graph showing that there is a positive correlation between the time of the CP110-PBM-GFP peak and S-phase length in cycle 12, even in the absence of endogenous CP110. *r*, R^2 and *p*-values are shown on the graph and a linear regression is plotted. **(D)** Centres from (B) corrected for S-phase length, to account for differences in S-phase length between the datasets.

To confirm that these transgenes were expressed in the absence of any endogenous WT CP110, I collected embryos from the fly lines I had imaged in **Figure 6.5** and carried out a western blot to check for endogenous CP110 expression. As expected, no endogenous CP110 expression was observed in these fly lines (**Figure 6.6**), thus excluding the possibility that the recruitment of the Polo-binding mutants was being facilitated by dimerisation with WT CP110. A band running below the GFP band was observed in these lines, which likely corresponds to a degradation product resulting from the overexpression of these proteins.

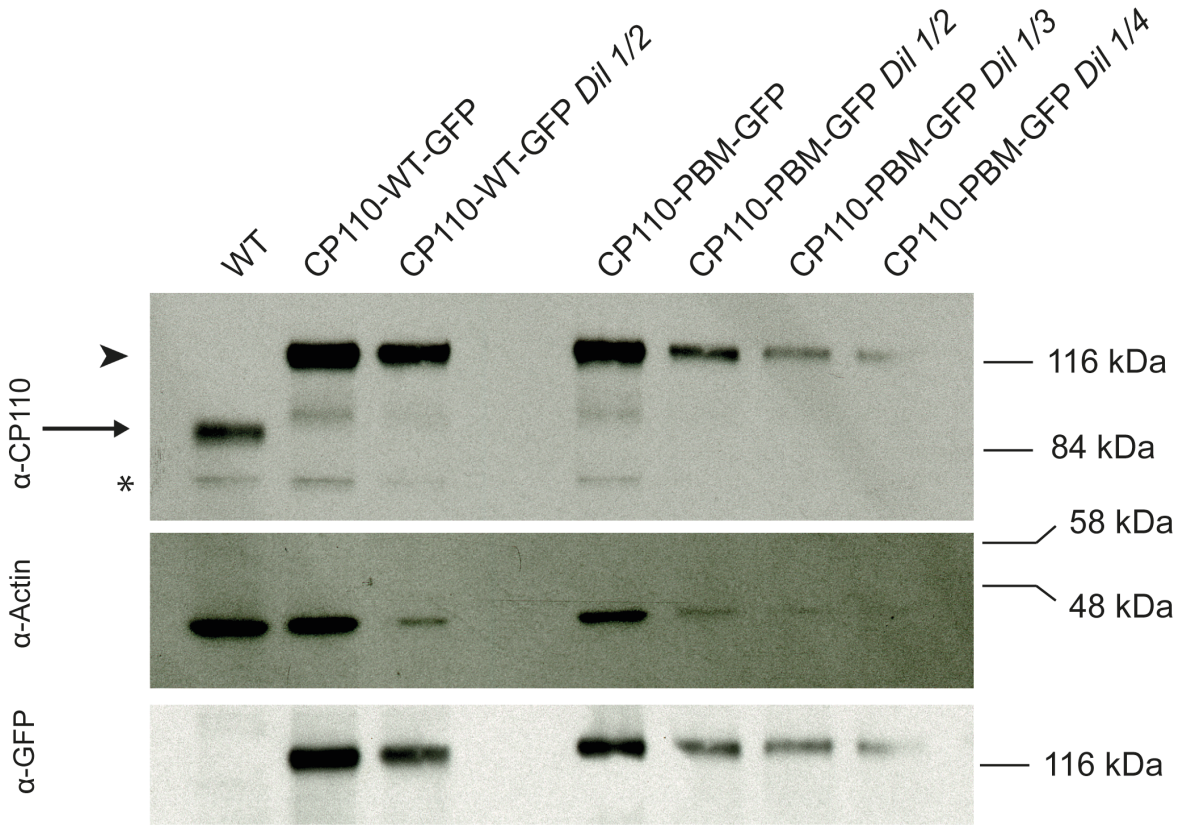


Figure 6.6 – Western blot to examine CP110 expression in embryos expressing CP110-WT-GFP or CP110-PBM-GFP in the full CP110 mutant background.

Western blot comparing the level of CP110 expression in WT embryos, *cp110* null embryos expressing one copy of CP110-WT-GFP, and *cp110* null embryos expressing one copy of CP110-PBM-GFP. The endogenous CP110 band is indicated with an arrow. The band corresponding to the GFP-tagged protein is indicated with an arrowhead. A degradation product is visible below the GFP bands. Non-specific bands are indicated with *. The results confirm that the embryos expressing CP110-WT-GFP or CP110-PBM-GFP do not express any detectable level of endogenous CP110. Actin is shown as a loading control. N = 2 technical repeats.

Overall, my results suggest that Polo binding to CP110 via its PB domain is not important for CP110's proper recruitment to centrioles. It will be important to also inspect the recruitment of Cep97-PBM-GFP in the Cep97 mutant background, since it is possible that any effect might be mediated through Cep97 rather than CP110.

6.2.1.3. Design of CP110/Cep97 with mutations to putative CDK phosphorylation sites

CP110 is also known to be a CDK substrate (Chen *et al.*, 2002). It is therefore possible that this cell cycle factor could play some role in CP110's recruitment or function. I have therefore also designed CP110 and Cep97 mutants that cannot be phosphorylated by CDK, or that potentially 'mimic' constitutive phosphorylation. Since it is established that Ser-Pro or Thr-Pro (SP/TP) motifs are the preferred phosphorylation sites of CDK, I identified these motifs in the CP110 and Cep97 sequence and designed constructs where the S or T residues were mutated to Ala (A). This would produce mutants that cannot be phosphorylated by CDK. I also designed constructs where I mutated the S or T residues to Asp (D) or Glu (E), respectively, in order to potentially mimic a constitutively phosphorylated protein. The locations of the putative CDK sites are indicated in the schematic in **Figure 6.7**. At the time of writing, I have produced vectors containing these constructs for injection into flies.

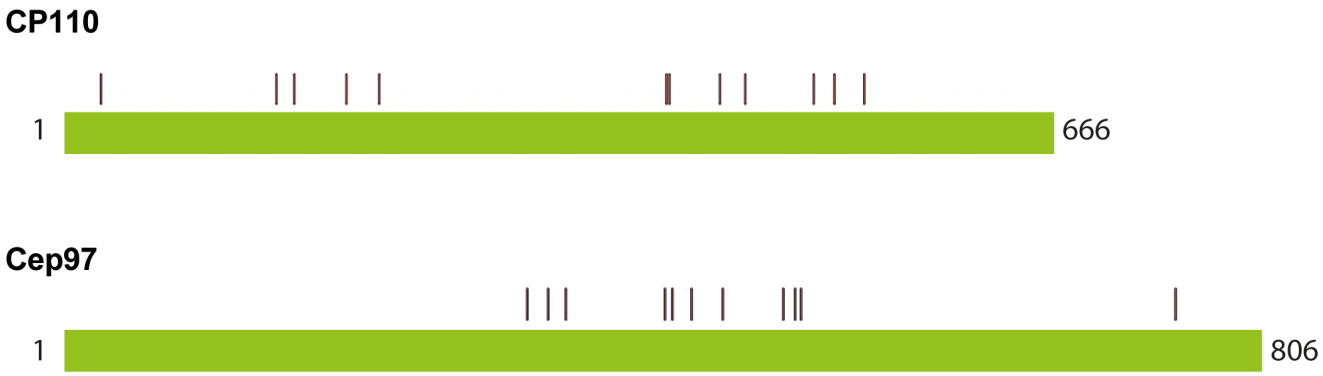


Figure 6.7 – Schematic showing predicted CDK phosphorylation sites in CP110/Cep97.

Schematic showing centriolar proteins CP110 and Cep97 (green boxes; numbers indicate the length of the amino acid sequences) with the locations of predicted CDK phosphorylation sites indicated by orange boxes.

6.2.1.4. CP110 recruitment to the centriole is not perturbed by deletion of regions that contain a high density of putative phosphorylation sites

Thus far, my results had indicated that binding between CP110 and Polo via Polo's PB domain is not important for CP110's recruitment to the centriole. To further explore the possibility that the phosphorylation state of CP110/Cep97 affects the recruitment of this complex to centrioles, I decided to take a broader approach, using alignments of these proteins in different *Drosophila* species to identify clusters of Ser/Thr (S/T) residues. Our laboratory has successfully employed this approach in previous work to identify important phosphorylation sites (for example, see Conduit *et al.*, 2014a). I found that Cep97 had a high density of S/T residues along the entire length of the protein, which made it challenging to identify specific clusters that might be phosphorylated *in vivo*. However,

two defined regions of CP110 exhibited clusters of S/T residues; these were found in residues 270-312 and 480-533 (these clusters are shown in **Figure 6.8 A**; see **Chapter 2** for more details).

To test if these regions of CP110 were required for its localisation at centrioles, I generated mRNA constructs where the sequences coding for these two regions had been deleted (**Figure 6.8 A**). I shall refer to this as ***'CP110-del-GFP'***. I also produced an mRNA construct encoding the WT CP110 sequence to act as a control. As before, I shall refer to this as ***'CP110-WT-GFP'***. Both mRNAs also encoded a C-terminal GFP tag, so that I could observe the localisation of the translated proteins. I then injected either CP110-del-GFP or CP110-WT-GFP into *cp110* null embryos expressing Asl-mCh, to mark the centrioles (**Figure 6.8 B**).

I found that, as with the CP110-WT-GFP construct (**Figure 6.8 C**), protein translated from the CP110-del-GFP construct was still recruited to centrioles, despite these deletions to the protein sequence (**Figure 6.8 D**). It appears that the background signal may be higher when the CP110-del-GFP construct is injected, and that this construct associates more strongly with the actin furrows than CP110-WT-GFP; it would therefore be interesting to generate transgenic flies expressing these constructs in order to quantify the amount of CP110 recruited to centrioles in each case. It is important to note that the injected embryos do not express any endogenous CP110, so it is not possible for CP110-del-GFP to be 'hitchhiking' to the centriole by dimerising with WT CP110. It is of course possible that there are other residues that might be phosphorylated to mediate this recruitment, though it

is interesting that CP110 is still recruited despite such relatively large sections of the protein being deleted.

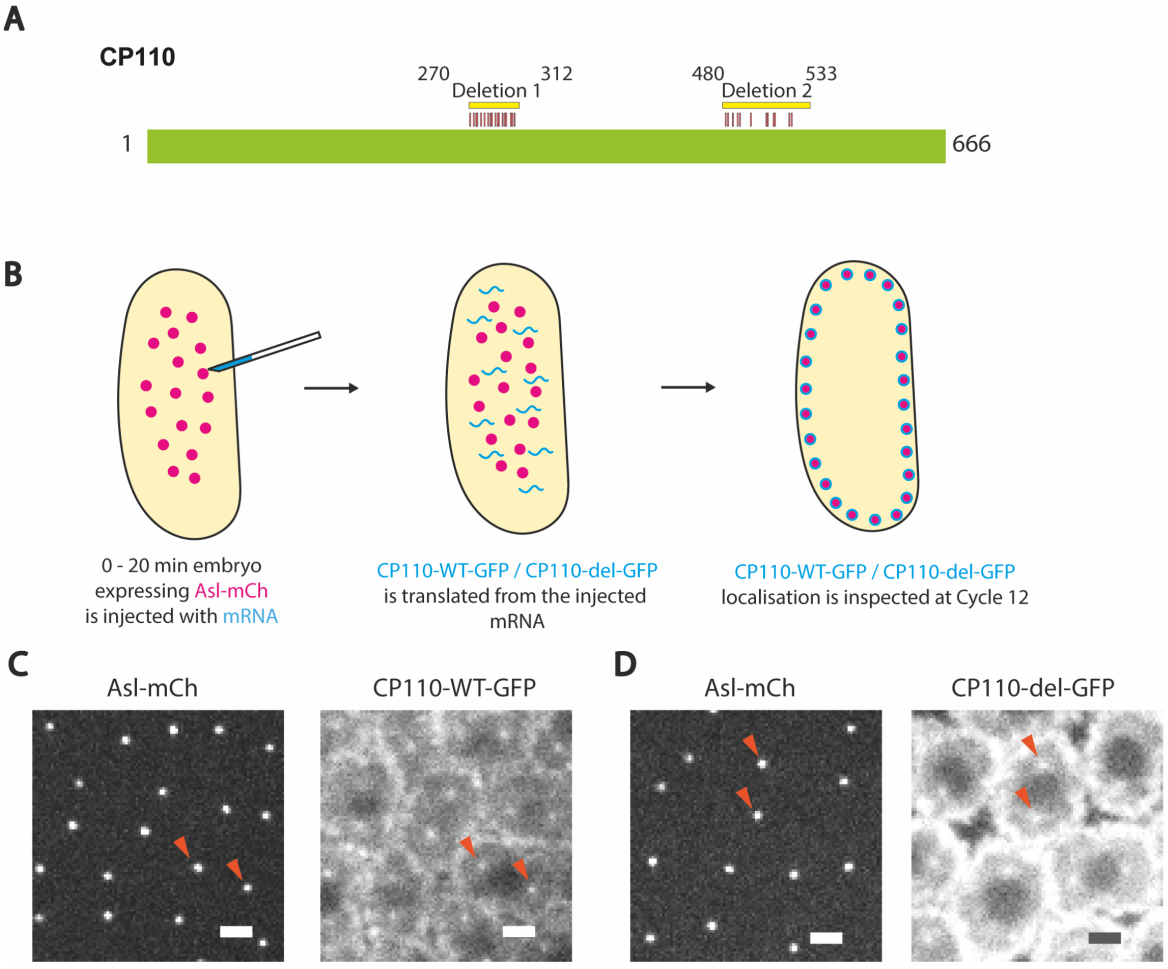


Figure 6.8 – Design and injection of the CP110-del-GFP construct into *Drosophila* embryos.

(A) Schematic showing centriolar proteins CP110 and Cep97 (green boxes; numbers indicate the length of the amino acid sequences) with the locations of the two deletions indicated by yellow boxes. The individual S/T residues are indicated by the vertical orange lines. **(B)** Schematic showing the process of the mRNA injection experiment. 0-20 min old embryos were injected with mRNA encoding either CP110-WT-GFP or CP110-del-GFP. The embryos were then left to develop for two hours at 25°C, during which time the injected mRNA was translated. Embryos were imaged in cycle 12, to investigate the localisation of the translated constructs. **(C-D)** Spinning disc confocal images showing Asl-mCherry (shown as a centrosome marker) and CP110-WT-GFP (C) or CP110-del-GFP (D) localisation in *cp110* null embryos. Example centrosome pairs are indicated with orange arrowheads. Scale bars represent 3 µm. These micrographs are presented as best-focussed slices rather than maximum intensity projections, in order to minimise the high background signal.

6.2.2. Polo binding is not important for CP110's MT capping function in wing discs

My results so far suggest that Polo binding is not important for the proper recruitment of CP110 to centrioles. It was therefore not clear what the significance of a potential CP110-Polo interaction might be. To test whether this interaction might be important for CP110's role in capping centriole MTs, I collaborated with Saroj Saurya to investigate centriole structure in *Drosophila* wing discs using electron microscopy (EM).

Franz *et al.* 2013 demonstrated that centrioles in *cp110* null mutant wing discs have dramatically extended centriole MTs, as well as a slightly longer cartwheel structure. This phenotype can be rescued by overexpression of the long isoform of CP110 (Franz *et al.*, 2013). To test whether this phenotype can also be rescued by the CP110 Polo-binding mutant, I generated larvae lacking endogenous CP110 and expressing one copy of either CP110-WT-GFP or CP110-PBM-GFP. I also grew up WT larvae as an additional control. S. Saurya dissected wing discs from these third-instar larvae and prepared them for EM. After imaging and measuring the lengths of centrioles in these wing discs (in a blinded analysis), she did not observe MT extensions in either the CP110-WT-GFP or CP110-PBM-GFP wing discs (**Figure 6.9 A**). This suggests that both forms of CP110 are able to rescue the *cp110* null phenotype, so Polo binding does not seem important for CP110's MT-capping function.

S. Saurya also found that centrioles in *cp110* null wing discs expressing either CP110-WT-GFP or CP110-PBM-GFP resulted in centrioles that were slightly, but significantly, shorter than those in the WT wing discs (**Figure 6.9 B**). A previous study from our laboratory has demonstrated that overexpressing WT CP110 in *cp110* null wing discs results in shortened centrioles when compared with WT wing discs (Franz *et al.*, 2013). Furthermore, we recently demonstrated that cartwheels grow to shorter lengths when CP110/Cep97 are overexpressed (**Chapter 3**). Hence it seems likely that the effect on centriole length in **Figure 6.9 B** is a result of the CP110-WT-GFP and CP110-PBM-GFP transgenes being overexpressed compared with the endogenous CP110 levels in the WT wing discs. Interestingly, while centrioles in CP110-WT-GFP or CP110-PBM-GFP were shorter than WT centrioles, they were the same length as each other. This implies that Polo binding is also not important for this effect of CP110 overexpression on centriole length.

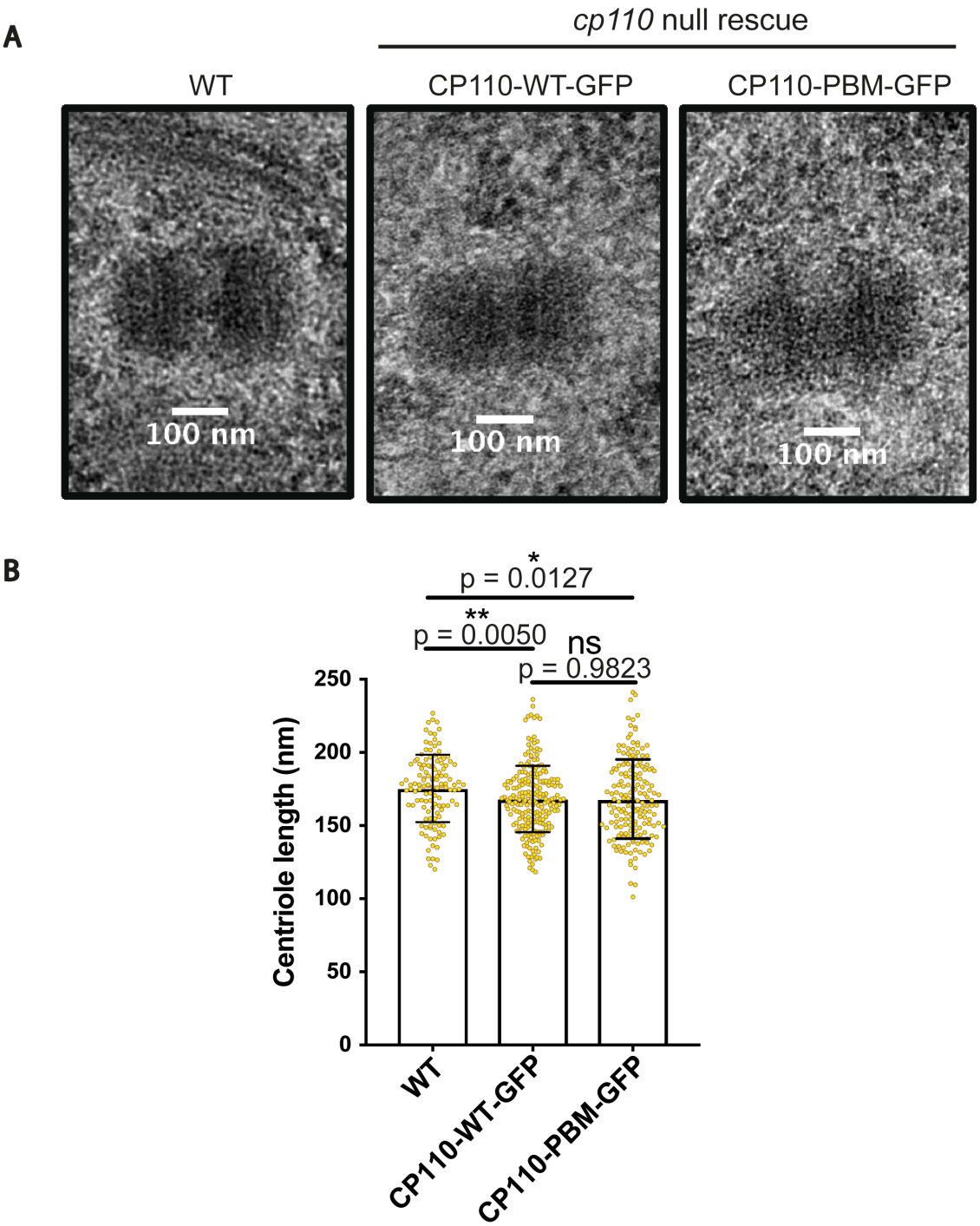


Figure 6.9 – Centriole length in *cp110* null wing discs overexpressing CP110-WT-GFP or CP110-PBM-GFP.

(A) Transmission EM micrographs showing representative centrioles in longitudinal section from WT wing discs, *cp110* null wing discs overexpressing CP110-WT-GFP, or *cp110* null wing discs overexpressing CP110-PBM-GFP. No centriole MT extensions were observed in any of the samples. **(B)** Bar charts showing quantifications of centriole length from the samples described in (A). Data are presented as mean ± SD; N ≥ 125 centrioles. T-tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values are shown on the graphs. Experiment and analysis carried out by Saroj Saurya. Images used to produce figure are courtesy of S. Saurya.

6.2.3. The Cep97 oscillation persists under S-phase arrest

In **Chapter 3**, I showed that the peak of the CP110/Cep97 oscillation strongly correlates with S-phase length, suggesting that the oscillation is in some way influenced by the cell cycle machinery. However, none of the perturbations to CP110 attempted in this Chapter thus far (disrupting Polo-binding, deletion of potential phosphorylation sites) have given an insight into how the cell cycle might influence the oscillation. I therefore attempted to employ blunter tools, such as arresting embryos at various stages of the cell cycle and then seeing how this affected the oscillation.

Mohammad Mofatteh (Raff Lab, University of Oxford) had recently generated double stranded RNAs (dsRNAs) against the *Drosophila* mitotic cyclins. Injection of these dsRNAs into embryos causes S-phase arrest, though centrioles are still able to continue duplicating, suggesting that the Cdk/cyclin oscillator (CCO) is perturbed (McCleland and O'Farrell, 2008). This approach has previously been used by our laboratory to demonstrate that the Plk4 oscillation is also able to continue under these conditions, and that it appears to drive repeated rounds of stochastic centriole duplication (Aydogan *et al.*, 2020). To test how the CP110/Cep97 oscillation is affected by this treatment, I injected the dsRNAs synthesised by M. Mofatteh into embryos expressing Ubq-RFP-Cep97 alongside Plk4-GFP. Intriguingly, under cell cycle arrest, levels of Cep97 still oscillated at the centriole (**Figure 6.10**). Like Plk4, there was one Cep97 oscillation peak per centriole duplication event. Cep97 appears to oscillate out of phase with Plk4 as in normal cell cycle conditions (**Figure 3.8 D**); this is most obvious in **Figure 6.10 A**, however it is unclear whether the normal degree of phase difference is maintained.

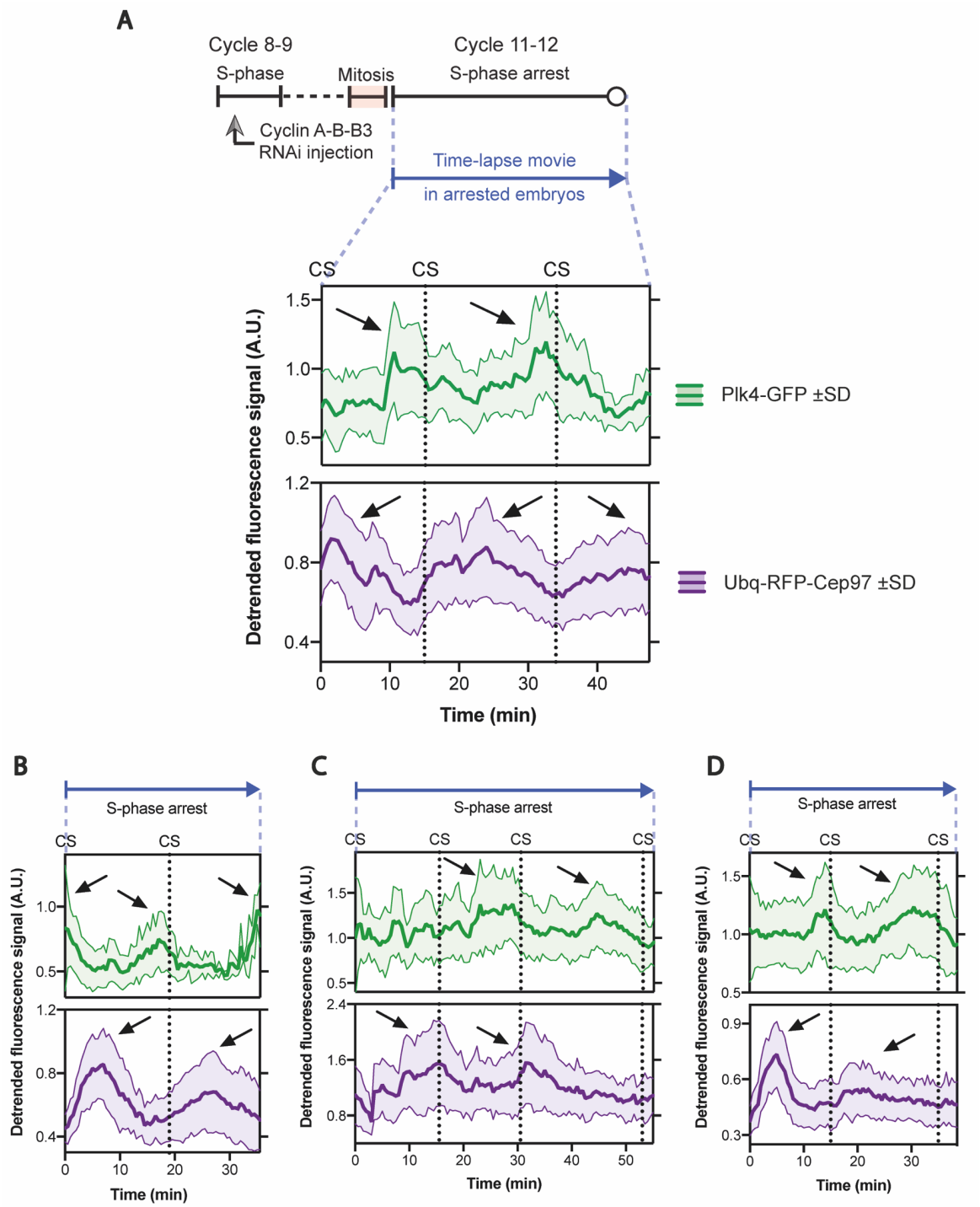


Figure 6.10 – The Cep97 oscillation continues under S-phase arrest.

(A) Graph showing quantification of Ubq-RFP-Cep97 and Plk4-GFP fluorescence intensity at centrosomes in a single embryo injected with dsRNA against the *Drosophila* mitotic cyclins. A schematic above the graph details the experimental methodology and timings. CS = centrosome separation (indicating centriole duplication events). Black arrows indicate oscillation peaks. Data are presented as mean ± SD. Further examples are shown in (B-D). Analysis was performed in collaboration with Mustafa Aydogan (Raff Lab, University of Oxford) and Felix Zhou (Lu Lab, University of Oxford).

6.2.4. CP110 is lost from centrioles under mitotic arrest

I was also curious about how CP110/Cep97 behaved in centrioles under mitotic arrest. To investigate this, I injected embryos expressing one copy of Ubq-CP110-GFP with colchicine. Colchicine is a potent MT depolymerising drug, which arrests embryos in mitosis. I consistently found that, in mitotic arrest, the CP110 signal at the centrioles decreased (as it normally would during mitosis) but was then completely lost from centrioles (**Figure 6.11**). This did not appear to be due to the centrioles sinking deeper into the embryo and out of the imaging plane, since I investigated whether they had moved in *z* and could not locate them. This loss of CP110 could be due to decreasing levels of CP110 in the cell during mitosis (Chen *et al.*, 2002). It would therefore be interesting to test the levels of CP110-GFP in the colchicine-arrested embryos. Alternatively, the loss of signal could be due to some effect of colchicine on the ends of the centriole MTs. Centriole MTs are well known to be more stable than cytoplasmic MTs, partly due to posttranslational modifications (reviewed in Janke and Magiera, 2020). Centriole MTs are therefore known to resist depolymerisation by colchicine and, unlike cytoplasmic MTs, are still visible with EM in cells that have been treated with this drug (Lagunoff and Chi, 1976). However, if CP110 sits at the far distal-ends of the centriole MTs, it is possible that some mild depolymerisation occurs there. It will therefore be important to arrest cells in mitosis using another technique, such as taxol injection (which stabilises MTs) or blocking the activation of the APC/C (Oliveira *et al.*, 2010) to test whether this loss of CP110 from centriole MTs is consistently observed.

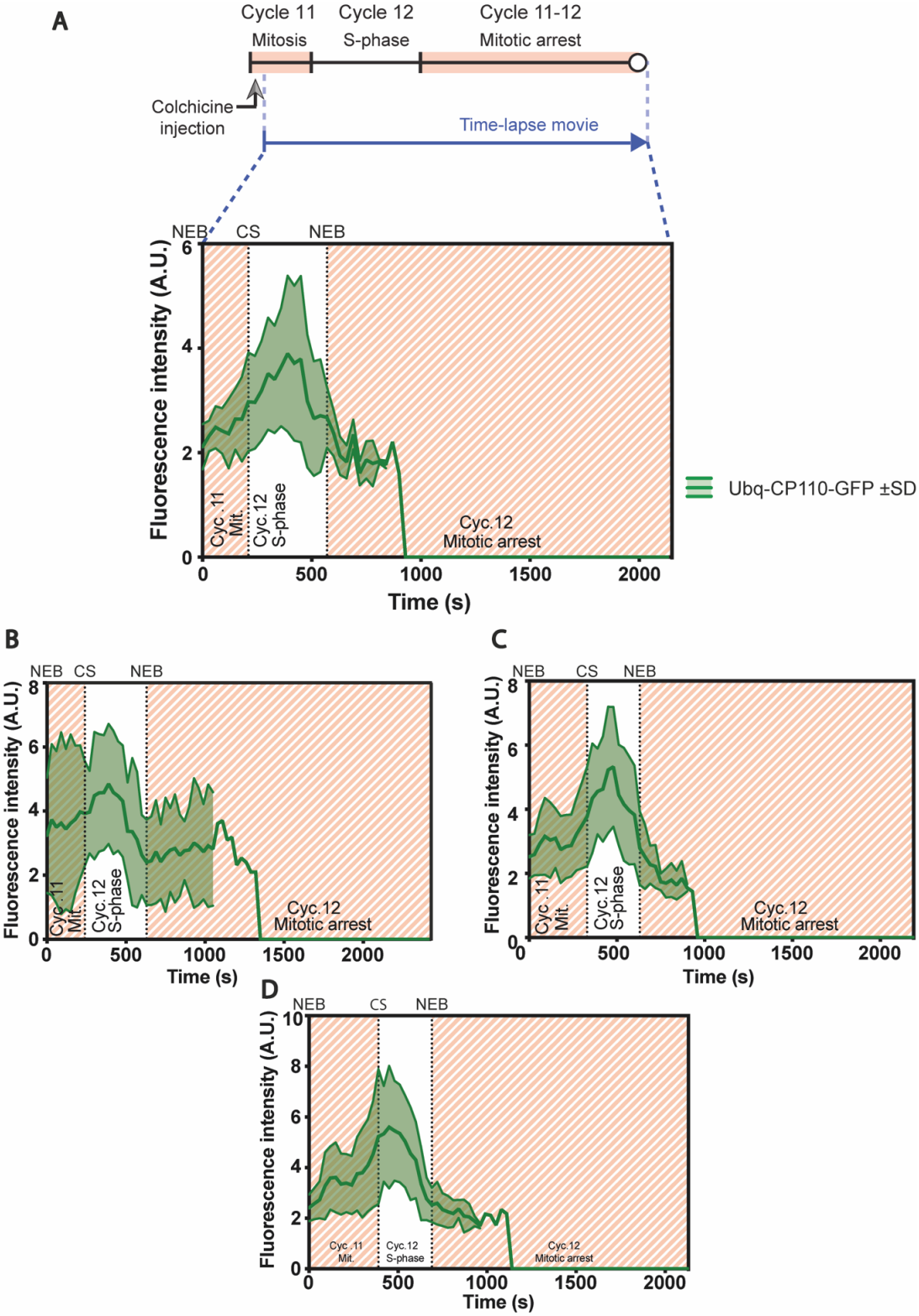


Figure 6.11 – The CP110-GFP signal is lost from centrioles under mitotic arrest.

(A) Graph showing quantification of Ubq-CP110-GFP fluorescence intensity at centrosomes in a single embryo injected with colchicine. A schematic above the graph details the experimental methodology and timings. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). Data are presented as mean ± SD. Further examples are shown in (B-D).

6.3. Discussion

In this Chapter I have shown that the ability of CP110 to interact with Polo via Polo's PB domain is not important for the proper recruitment of CP110 to the centriole. Furthermore, a form of CP110 in which I deleted two regions that contained a large number of potential phosphorylation sites still localised to centrioles. Surprisingly, I also found that Cep97 is able to oscillate at the centriole even under S-phase arrest conditions. In **Chapter 3**, I showed that the peak of the CP110/Cep97 oscillation correlates well with S-phase length. Together, these results suggest that, although the CP110/Cep97 oscillation is influenced by the Cdk/cyclin oscillator (CCO) that drives the rapid nuclear divisions in these embryos, the CCO may not be required to drive the oscillation *per se*. This is further supported by results from Chen *et al.* 2002, which demonstrated that CP110 mutants that could not be phosphorylated by CDK were still recruited to centrosomes (although it is unclear whether the dynamics of their recruitment there are normal). Intriguingly, I found that levels of CP110 at the centrioles drop dramatically in embryos that have been arrested in mitosis with the MT-depolymerising drug colchicine. This might suggest that the mitotic state is necessary to remove CP110 from centrioles, but I note that this mechanism of CP110 removal is unlikely to explain why levels of CP110/Cep97 normally start to decline at centrioles during S-phase.

Although I saw no effect of Polo-binding mutations on CP110's recruitment dynamics, it is possible that any Polo binding to the CP110/Cep97 complex might be mediated through Cep97 (or a redundant combination of CP110 and Cep97). CP110 and Cep97 are known interactors (Spektor *et al.*, 2007), which could explain why both proteins were identified as PLK1 interactors in pulldown experiments (Legg, 2019). It will therefore be important to

also assess the recruitment dynamics of Cep97-PBM-GFP and compare these to Cep97-WT-GFP. Preparations are currently being made for this experiment, though due to delays as a result of the coronavirus pandemic it is not complete at the time of writing.

It was possible that the ability of CP110/Cep97 to bind Polo was important for their ability to influence centriole MT length, rather than their recruitment to centrioles. In collaboration with Saroj Saurya, we found that overexpression of CP110-WT-GFP in *cp110* null wing discs caused a small, but significant, reduction in centriole length. This is consistent with previously published EM results from the laboratory (Franz *et al.*, 2013) and with our finding that CP110 overexpression leads to a small but significant decrease in centriole cartwheel length in living embryos (**Chapter 3**). However, the ability of CP110 to bind Polo does not seem to be important for this function, since overexpressing CP110-PBM-GFP in wing discs induced a similar reduction in centriole length as overexpressing CP110-WT-GFP. Furthermore, over-elongated MTs were not observed in *cp110* null wing expressing either CP110-WT-GFP or CP110-PBM-GFP. This suggests that, even if CP110's Polo-binding sites are disrupted, it is still able to suppress the dramatic elongation of the centriole MTs in wing discs. However, it should be noted that the Polo-binding mutants have not been tested for their ability to bind Polo, so it is possible that they are still able to interact; indeed, there are examples of proteins interacting with Polo via non-canonical domains (Archambault *et al.*, 2008). It will therefore be important to employ interaction assays in order to confirm whether the mutations to CP110's Polo-binding sites have successfully inhibited its interaction with Polo.

Finally, I also investigated the effect of mitotic arrest on CP110 recruitment. Unlike S-phase arrest, where Cep97 was able to continue oscillating, mitotic arrest resulted in the CP110-GFP signal being lost from centrioles. This could be because CP110 levels are decreasing in mitosis, which in human cells is a result of its ubiquitylation by the $SCF^{CyclinF}$ complex (Chen *et al.*, 2002, D'Angiolella *et al.*, 2010), so there is not a sufficient pool of CP110-GFP to replenish the signal at the centriole and maintain the oscillation. However, it is unknown whether CP110 levels decrease in mitotic *Drosophila* embryos. Alternatively, it is possible that the ends of the centriole MTs are being depolymerised, thus removing the CP110 signal. It will therefore be important to investigate CP110 dynamics in embryos arrested in mitosis via other approaches.

To conclude, it seems that the cell cycle is important for the entrainment of the CP110/Cep97 oscillation. However, my preliminary results regarding Polo, coupled with my finding that Cep97 is able to oscillate even under S-phase arrest, suggest that cell cycle regulation is not essential to mediate CP110/Cep97's oscillatory recruitment. Instead, I propose that the growth of the centriole MTs themselves (whose dynamics I suggest in **Chapter 3** are largely mediated by the growth rate and period of the underlying cartwheel structure) promotes the recruitment of CP110/Cep97 to the distal-ends of centrioles. This could be because CP110/Cep97 have an affinity for the growing ends of MTs which would lead them to accumulate at the growing tip. It is known that EB proteins accumulate at the growing ends of MTs, and that the state of the MT appears to affect the number of available EB binding sites (Akhmanova and Steinmetz, 2015); a similar phenomenon might explain why CP110/Cep97 begin to dissociate from the centriole in late S-phase, once the centriole MTs reach the end of their growth period. This might also explain why the CP110-GFP signal was lost from centrioles arrested in mitosis, since colchicine might

be depolymerising or otherwise affecting the state of the centriole MT ends (perhaps by suppressing growth, so the proteins completely dissociate). However, we know that CP110/Cep97 are recruited specifically to the centriole MTs, and not to cytoplasmic MTs. It therefore seems likely that, while CP110/Cep97 might have some affinity for growing MTs, additional factors are required (possibly centriolar proteins) to ensure that CP110/Cep97 are stabilised specifically at centriole MTs.

Chapter 7: General Discussion

7.1. Summary of Key Findings

In this thesis, I have investigated the role of CP110/Cep97 and its interaction network in centriole size control. I now summarise my key findings, as discussed in the preceding Chapters. Together with my collaborators, I found that centriole MTs grow slower and to a shorter length when CP110/Cep97 are overexpressed. They may also grow for a longer period (**Figure 3.5**). Meanwhile, CP110/Cep97 levels oscillate at the centriole, peaking in mid-S-phase (**Figure 3.8**). This was an exciting observation, because Plk4 had recently been found to oscillate at the centriole and thus regulate the growth rate and growth period of the cartwheel structure (Aydogan *et al.*, 2020). However, I have shown that the CP110/Cep97 oscillation does not appear to directly influence the growth parameters of the centriole MTs (**Figures 3.13-3.15**). Instead, it seems that modulating CP110/Cep97 levels has a surprising impact on the cartwheel growth parameters, possibly via an interaction with Plk4 (**Figures 3.16** and **3.17**). Interestingly, when CP110/Cep97 are overexpressed, the cartwheel grows slower and for a longer period. This is reminiscent of the effect of the same perturbation on centriole MTs.

I have also studied the interaction partners of CP110/Cep97 at the distal-end. I found that Cep104, a protein previously posited to recruit CP110/Cep97 to centrioles (Jiang *et al.*, 2012, Rezabkova *et al.*, 2016, Satish Tammana *et al.*, 2013) is not important for mediating CP110/Cep97 recruitment dynamics in fly embryos (**Figures 4.8** and **4.9**). Cep104 also oscillates at the centriole, but peaks later in S-phase than Cep97 and appears to act

downstream of CP110/Cep97 (**Figures 4.10-4.12**). These results are consistent with a study published towards the end of my DPhil which suggested Cep104 functions independently of CP110, after CP110 has been removed from mother centrioles (Yamazoe *et al.*, 2020). My preliminary studies into the MT depolymerase Klp10A have also suggested that this protein is not important for CP110/Cep97 recruitment (**Figure 5.1**). Finally, I explored a recent report that CP110/Cep97 interact with the mitotic kinase PLK1 in human cells (Legg, 2019). Since the CP110/Cep97 oscillation seems entrained by the cell cycle, I was curious as to whether Polo (the *Drosophila* homologue of PLK1) might be important for mediating this recruitment. Instead, I found that perturbation of Polo binding to CP110 had no effect on CP110's recruitment dynamics (**Figure 6.4** and **6.5**) or, indeed, on its ability to cap centriole MTs in *Drosophila* wing discs (**Figure 6.9**). Indeed, the Cep97 oscillation is able to persist even under S-phase arrest, suggesting that the cell cycle is not required to mediate CP110/Cep97's recruitment to, or dissociation from, centrioles (**Figure 6.10**).

7.2. Proposed model for daughter centriole growth

The above findings have led me to propose a '***cartwheel-driven***' model of centriole biogenesis. It is widely suspected in the field that the cartwheel structure must assemble in order for the surrounding centriole MTs to be nucleated. Under the cartwheel-driven model, these MTs assemble as the cartwheel forms, resulting in the growth rate and period of these centriole MTs tracking that of the underlying cartwheel structure. This would mean that MT growth is passively regulated by the growth of the cartwheel, with the

centriole MTs growing to the same length as the cartwheel. This suggestion is supported by previous work, which showed that reducing Plk4 expression levels in the fly embryo causes a decrease in the growth rate of both the cartwheel and the centriole MTs (Aydogan *et al.*, 2018). We also found that increasing the expression levels of CP110/Cep97 resulted in a decrease in growth rate for both the cartwheel (**Figure 3.16**) and the centriole MTs (**Figure 3.5**), possibly mediated by an interaction between CP110/Cep97 and Plk4 (**Figure 3.17**).

However, this model is not as straightforward as it first appears. For example, it suggests that CP110/Cep97 interact with Plk4 to somehow influence the growth rate and period of the cartwheel. Our work has suggested that they interact at the centriole, rather than in the cytoplasm, but, as illustrated in **Figure 7.1 A**, CP110/Cep97 and Plk4 localise to opposite ends of the centriole (Kleylein-Sohn *et al.*, 2007). However, perhaps they interact at the onset of biogenesis, when they are more likely to be close together on the nascent daughter. Recruitment of CP110/Cep97 and Plk4 to centrioles certainly begins early in S-phase (**Figure 3.8**), so this explanation is plausible.

Furthermore, it is widely believed in the field that the centriole MTs grow from the distal-end of the daughter, whereas Sas-6 is known to be incorporated at the proximal-end (Aydogan *et al.*, 2018). If the centriole MTs assemble as the cartwheel does, this suggests that the growing cartwheel would have to 'slide' along the surrounding centriole MTs, rather like a train on a track (**Figure 7.1 B**). Alternatively, it is possible that the MTs are able to grow from both ends of the centriole. Indeed, centrioles have been observed to elongate in spermatocytes, despite their distal-ends being incorporated in the CLR

(Riparbelli *et al.*, 2012), which suggests incorporation of subunits at the proximal-end is possible. Electron microscopy studies in *C. elegans* have shown that centriole MTs seem to first nucleate randomly along the length of the cartwheel as they begin to assemble, implying that they can either slide, or can grow from either the proximal- or distal-ends of centrioles (though the distal-end appears slightly more favoured; Pelletier *et al.*, 2006).

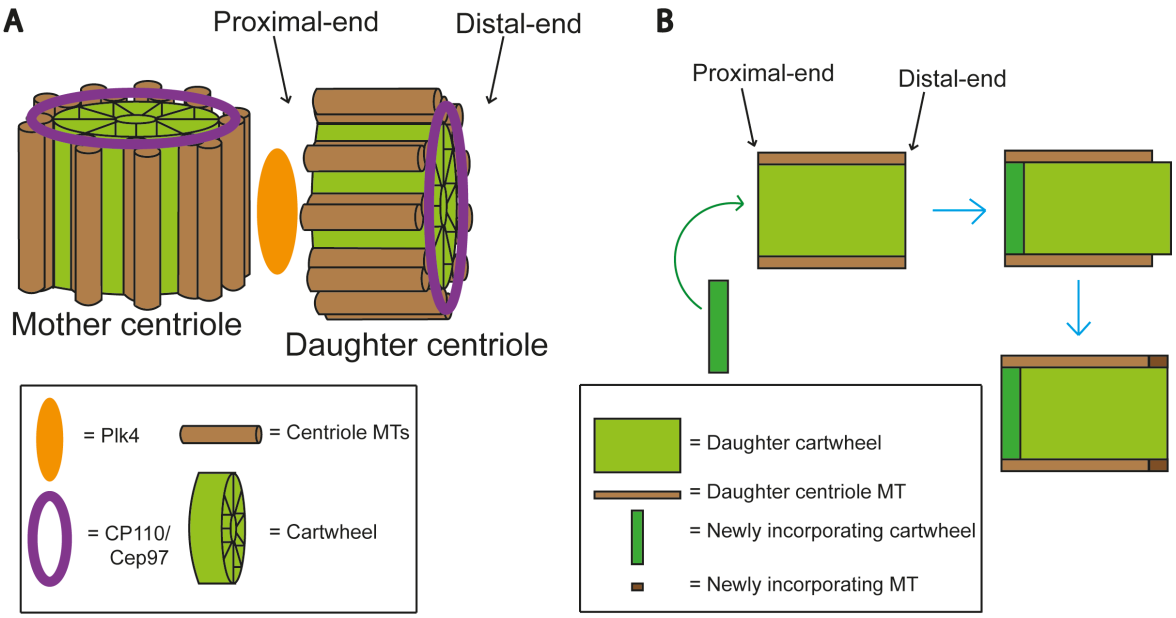


Figure 7.1 – Schematic illustrating important considerations for the '***cartwheel-driven***' model of daughter centriole biogenesis.

(A) Schematic showing the localisation of Plk4 (orange dot) at the proximal-end of the daughter centriole and of CP110/Cep97 (purple ring) at the distal-ends of the mother and daughter centrioles. **(B)** Schematic illustrating how the daughter centriole might grow if cartwheel (pale green) and MTs (pale brown) are assembled at the same time. If new cartwheel units (dark green) are incorporated at the proximal end (Aydogan *et al.*, 2018) while MTs are assembling at the distal-end of the cartwheel, it seems necessary for the elongating cartwheel to slide along the MTs. MTs could then be extended (dark brown) at the distal-end.

This concept of the cartwheel sliding on the centriole MTs raises another possibility, which I term the '***mechanical coupling***' model. Under this scenario, loss of CP110/Cep97 from the distal-end might promote over-elongation of the centriole MTs, thus providing

more ‘track’ upon which to slide the cartwheel structure (**Figure 7.2**). This could reduce steric hindrance for Sas-6 incorporation at the proximal-end, thus explaining why cartwheels are able to reach a longer length in *cp110*/*cep97* null embryos (**Figure 3.16**). It has previously been shown that the distribution of Sas-4 and Plk4 extends in centrioles with disrupted cartwheels, which raises the possibility that the cartwheel can physically inhibit the incorporation of proteins (Gottardo *et al.*, 2016). However, given that I did not find strong evidence for CP110/Cep97 directly influencing MT growth during daughter biogenesis (**Figures 3.13-3.15**), this seems less likely than the cartwheel-driven model. The mechanical coupling model also does not explain how CP110/Cep97 are able to influence the parameters of the Plk4 oscillation.

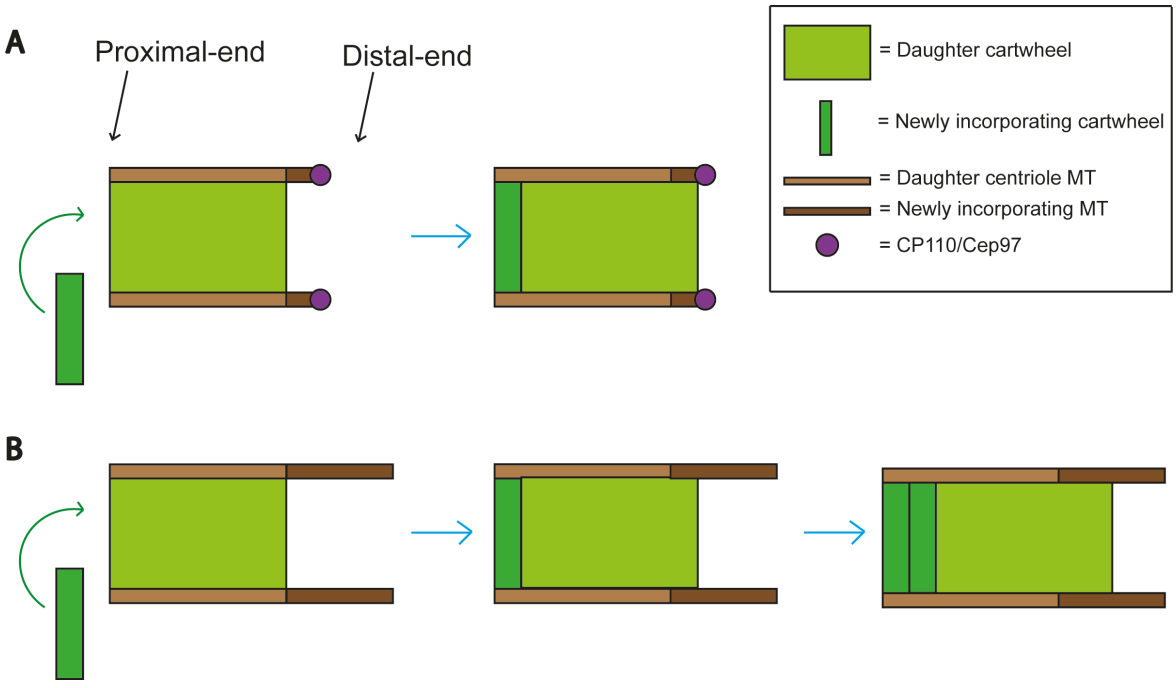


Figure 7.2 – Schematic illustrating the concept of the ‘***mechanical model***’ of daughter centriole biogenesis.

(A) Schematic illustrating how centrioles might grow in WT embryos under the assumptions of the ‘***mechanical model***’. Here, CP110/Cep97 bind to the distal-ends of the centriole MTs, restricting their over-elongation. Sas-6 is incorporated at the proximal-end; the cartwheel ‘slides’ along the centriole MTs to allow this incorporation. **(B)** Schematic illustrating how cartwheels might grow to a longer final length in *cp110/cep97* null embryos. Here, if CP110/Cep97 are absent, and if this allows the centriole MTs to grow in a less restricted manner, there might be more MT ‘track’, thus allowing the cartwheel to slide further along the length of the centriole MTs and, in turn, allowing more Sas-6 incorporation at the proximal-end.

Given my findings in **Chapter 6**, I propose that CP110/Cep97 are partly recruited to centriole MTs via their affinity for growing MTs. This is because I have shown that Cep97 still exhibits oscillatory recruitment under S-phase arrest, suggesting that there is an intrinsic ability to recruit these proteins to growing centrioles independent of cell cycle regulation. Centriole biogenesis has been shown to continue in embryos under S-phase arrest by mitotic cyclin RNAi (Aydogan *et al.*, 2020), suggesting that these centrioles, and thus probably their MTs, are indeed growing.

I suggest that the recruitment of CP110/Cep97 to MTs is somewhat analogous to the recruitment of EB proteins (Akhmanova and Steinmetz, 2015), although I suspect another centriolar protein is involved to facilitate recruitment of CP110/Cep97 specifically to the centriole MTs, and not to cytoplasmic MTs. Although cell cycle regulation is not required for CP110/Cep97's recruitment to, and dissociation from, centriole MTs, it likely plays a role in entraining the oscillation, since we found that the peak of the CP110/Cep97 oscillation consistently correlates strongly to S-phase length. I have shown that this correlation is maintained even when CP110 cannot bind Polo (**Figure 6.4** and **6.5**), suggesting that this entrainment is mediated by some other cell cycle factor, possibly a CDK.

The above model suggests that dissociation of CP110/Cep97 from centriole MTs might be triggered by the end of the growth phase of centriole MTs. However, this is unlikely to be the full explanation, since I showed in **Chapter 3** that the plateau of the cartwheel growth does not correlate with the peak of the Cep97 oscillation (**Figure 3.9**). This is also consistent with Aydogan *et al.* 2018, where it was reported that the plateau of cartwheel

growth does not correlate with S-phase length. If the centriole MTs stop growing at the same time as the cartwheel, it therefore seems unlikely that this is what prompts the release of CP110/Cep97. It is possible that the cell cycle is the most important driver of CP110/Cep97 dynamics in cycling cells, whilst their affinity for growing MTs might facilitate oscillatory recruitment dynamics under cell cycle arrest conditions in a more passive manner.

7.3. Centriole size maintenance

Once the daughter has grown, it must also be maintained at the correct size. For example, mother centrioles can over-elongate when CPAP is overexpressed, even though they are fully-grown (Kohlmaier *et al.*, 2009, Schmidt *et al.*, 2009, Tang *et al.*, 2009). In addition to their surprising role in influencing cartwheel growth dynamics (**Figure 3.16**), CP110/Cep97 also appear to function in 'capping' the centriole MTs to inhibit over-elongation (Dobbelaere *et al.*, 2020, Franz *et al.*, 2013). As discussed in **Chapter 3**, I suggest that this role might be particularly important when cells enter G_2-phase, since there is a second period of daughter centriole growth at this stage. In flies, this is most clearly observed in the spermatocyte centrioles (Tates, 1971), and CP110 appears to be removed from these centrioles at the onset of this second phase of growth (Franz *et al.*, 2013, Shoda *et al.*, 2021). Perhaps CP110/Cep97 normally act to limit this growth in other tissues, such as the wing discs. As discussed in **Chapter 3**, two protein isoforms of CP110 exist in flies, with differences in their expression patterns suggesting that they might perform different roles in different tissues (Franz *et al.*, 2013).

Unlike CP110/Cep97, I found no evidence for Klp10A playing a role in cartwheel size regulation during the growth of the daughter centriole (**Figure 5.3**). These results are consistent with Klp10A acting after the daughter centriole has reached its final size. Klp10A depletion results in over-elongated centriole MTs, and this kinesin has been suggested to trim back any centriole MTs that overgrow (Delgehyr *et al.*, 2012, Franz *et al.*, 2013, Gottardo *et al.*, 2016). It has also been suggested that CP110 protects centriole MTs from Klp10A (Delgehyr *et al.*, 2012), although this seems unlikely since loss of CP110 can promote formation of dramatic centriole MT extensions even in the presence of Klp10A (Franz *et al.*, 2013). It is possible that Klp10A is recruited to the distal-ends of the centriole by CP110 (Franz *et al.*, 2013); my results suggest that Klp10A does not recruit CP110/Cep97 (**Figure 5.1**), but further work is needed to be able to explore the effect of CP110/Cep97 loss on the sub-cellular localisation of Klp10A.

Cep104 also appears to have little involvement in regulating the growth of the daughter centriole, and likely functions downstream of CP110/Cep97 (**Chapter 4**). This suggests that it plays a role after the daughter centriole has finished growing. Unlike Klp10A, however, I suggest that Cep104 is not involved in centriole size maintenance and is more important in cells that form cilia. This is because Cep104's TOG domain has been suggested to mediate axoneme extension (Al-Jassar *et al.*, 2017, Louka *et al.*, 2018, Rezabkova *et al.*, 2016, Yamazoe *et al.*, 2020). The predicted tertiary structure of *Drosophila* Cep104 is similar to known TOG domain structures (**Figure 4.17**) which suggests that Cep104 could perform similar roles in flies and humans. However, any effect of *cep104* deletion on cilia length or structure is likely to be subtle, since I showed that *cep104* null flies are viable and fertile (**Figure 4.7**) despite these mutants lacking almost the entire predicted TOG domain. Cep104 recruitment to centrioles in *Drosophila* embryos

appears to be mediated by CP110/Cep97 (**Figure 4.12**), and it will be interesting to see whether this recruitment is also affected in *cp110*/*cep97* null ciliated cells.

Results from a recent publication suggest that an MT polymerase called Orbit/CLASP, which has previously been shown to antagonise Klp10A in the context of the mitotic spindle (Laycock *et al.*, 2006), is important for centriole MT extension (Shoda *et al.*, 2021). It is intriguing that proteins involved in spindle size maintenance can also play analogous roles in centriole MT size control, and this may form an interesting area of research going forward.

7.4. Future directions

As alluded to throughout this thesis, the context of cell-type is important to consider. I believe that this is particularly important for understanding the role of Cep104. Given my results and the existing literature, it seems likely that Cep104 is important for ciliogenesis but not for centriole growth regulation/centriole size maintenance. However, I cannot rule out the possibility that Cep104 contributes to centriole MT polymerisation in both ciliated and non-ciliated cells. Indeed, I did observe an increase in cartwheel growth period in *cep104* null embryos; this is likely due to the increased S-phase length in these embryos compared with controls, though it will be important to repeat this experiment. To further test whether Cep104 affects centriole size, I am collaborating with Saroj Saurya to examine whether centriole MTs in *cep104* null *Drosophila* wing discs are shortened, and whether the centriole MT extensions that form in the absence of CP110 (Franz *et al.*, 2013) can still assemble when Cep104 is also absent. It will also be important to examine

ciliated cells in the *cep104* mutant flies and determine whether their cilia form normally, as is implied by my finding that these mutant flies are not uncoordinated.

My studies have also revealed that the rate of Sas-6 incorporation into growing daughter cartwheels is influenced by the choice of fluorescent tag, with Sas-6-mNG incorporating into growing daughter centrioles at a faster rate than Sas-6-GFP. This work therefore identifies Sas-6-mNG as a useful tool for the laboratory moving forwards, since this faster incorporation has the potential to reveal more subtle phenotypes. It will be important to compare the growth rate of cartwheels in embryos lacking endogenous Sas-6 and expressing either Sas-6-mNG or Sas-6-GFP. It will also be interesting to see whether these cartwheels grow to reach the same size, or whether the faster growth rate exhibited in Sas-6-mNG cartwheels translates to an increased final length. This is important to test, since it is also possible that a molecular ruler might regulate centriole size; if the cartwheels in these embryos reach different sizes, this would provide strong evidence that such a mechanism does not operate in *Drosophila* embryo centrioles.

Finally, it will be important to test the assumptions of the 'cartwheel-driven' model of centriole biogenesis proposed above. This could include assaying whether the centriole MTs and the cartwheel grow in the same way, by simultaneously imaging GFP-Cep97 and Sas-6-mCherry in the same embryo. This would require some modifications to the existing GFP-Cep97 distancing algorithm, and possibly retraining of the convolutional neural network (CNN). To further test the model, it would also be interesting to investigate whether mutations that perturb the growth of the cartwheel perturb the centriole MT dynamics in the same way, and vice versa. Such experiments would address long-held

questions in the field about how the cartwheel and centriole MTs co-ordinately assemble during centriole biogenesis.

References

AKHMANOVA, A. & STEINMETZ, M. O. 2015. Control of microtubule organization and dynamics: two ends in the limelight. *Nat Rev Mol Cell Biol,* 16**,** 711-26.

AL-BASSAM, J. & CHANG, F. 2011. Regulation of microtubule dynamics by TOG-domain proteins XMAP215/Dis1 and CLASP. *Trends Cell Biol,* 21**,** 604-14.

AL-JASSAR, C., ANDREEVA, A., BARNABAS, D. D., MCLAUGHLIN, S. H., JOHNSON, C. M., YU, M. & VAN BREUGEL, M. 2017. The Ciliopathy-Associated Cep104 Protein Interacts with Tubulin and Nek1 Kinase. *Structure,* 25**,** 146-156.

ALVAREZ-RODRIGO, I., STEINACKER, T. L., SAURYA, S., CONDUIT, P. T., BAUMBACH, J., NOVAK, Z. A., AYDOGAN, M. G., WAINMAN, A. & RAFF, J. W. 2019. Evidence that a positive feedback loop drives centrosome maturation in fly embryos. *Elife,* 8**,** e50130.

ALVAREZ-RODRIGO, I., WAINMAN, A., SAURYA, S. & RAFF, J. W. 2021. Ana1/Cep295 helps recruit Polo/PLK1 to centrioles to promote mitotic PCM assembly and centriole elongation. *J Cell Sci*.

ANDERSON, R. G. 1972. The three-dimensional structure of the basal body from the rhesus monkey oviduct. *J Cell Biol,* 54**,** 246-65.

ARCHAMBAULT, V., D'AVINO, P. P., DEERY, M. J., LILLEY, K. S. & GLOVER, D. M. 2008. Sequestration of Polo kinase to microtubules by phosphopriming-independent binding to Map205 is relieved by phosphorylation at a CDK site in mitosis. *Genes Dev,* 22**,** 2707-20.

ARQUINT, C., GABRYJONCZYK, A. M., IMSENG, S., BOHM, R., SAUER, E., HILLER, S., NIGG, E. A. & MAIER, T. 2015. STIL binding to Polo-box 3 of PLK4 regulates centriole duplication. *Elife,* 4.

ARQUINT, C. & NIGG, E. A. 2014. STIL microcephaly mutations interfere with APC/C-mediated degradation and cause centriole amplification. *Curr Biol,* 24**,** 351-60.

AVIDOR-REISS, T., MAER, A. M., KOUNDAKJIAN, E., POLYANOVSKY, A., KEIL, T., SUBRAMANIAM, S. & ZUKER, C. S. 2004. Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. *Cell,* 117**,** 527-39.

AYDOGAN, M. G., STEINACKER, T. L., MOFATTEH, M., WILMOTT, Z. M., ZHOU, F. Y., GARTENMANN, L., WAINMAN, A., SAURYA, S., NOVAK, Z. A., WONG, S. S., GORIELY, A., BOEMO, M. A. & RAFF, J. W. 2020. An Autonomous Oscillation Times and Executes Centriole Biogenesis. *Cell,* 181**,** 1566-1581 e27.

AYDOGAN, M. G., WAINMAN, A., SAURYA, S., STEINACKER, T. L., CABALLE, A., NOVAK, Z. A., BAUMBACH, J., MUSCHALIK, N. & RAFF, J. W. 2018. A homeostatic clock sets daughter centriole size in flies. *J Cell Biol,* 217**,** 1233-1248.

AZIMZADEH, J., HERGERT, P., DELOUVEE, A., EUTENEUER, U., FORMSTECHER, E., KHODJAKOV, A. & BORNENS, M. 2009. hPOC5 is a centrin-binding protein required for assembly of full-length centrioles. *J Cell Biol,* 185**,** 101-14.

BANGS, F. K., SCHRODE, N., HADJANTONAKIS, A. K. & ANDERSON, K. V. 2015. Lineage specificity of primary cilia in the mouse embryo. *Nat Cell Biol,* 17**,** 113-22.

BASSETT, A. R., TIBBIT, C., PONTING, C. P. & LIU, J. L. 2014. Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System. *Cell Rep,* 6**,** 1178-1179.

BASTO, R., LAU, J., VINOGRADOVA, T., GARDIOL, A., WOODS, C. G., KHODJAKOV, A. & RAFF, J. W. 2006. Flies without centrioles. *Cell,* 125**,** 1375-86.

BAYLESS, B. A., GIDDINGS, T. H., JR., WINEY, M. & PEARSON, C. G. 2012. Bld10/Cep135 stabilizes basal bodies to resist cilia-generated forces. *Mol Biol Cell,* 23**,** 4820-32.

BENOIT, M., ASENJO, A. B. & SOSA, H. 2018. Cryo-EM reveals the structural basis of microtubule depolymerization by kinesin-13s. *Nat Commun,* 9**,** 1662.

BETTENCOURT-DIAS, M. & CARVALHO-SANTOS, Z. 2008. Double life of centrioles: CP110 in the spotlight. *Trends Cell Biol,* 18**,** 8-11.

BETTENCOURT-DIAS, M., RODRIGUES-MARTINS, A., CARPENTER, L., RIPARBELLI, M., LEHMANN, L., GATT, M. K., CARMO, N., BALLOUX, F., CALLAINI, G. & GLOVER, D. M. 2005. SAK/PLK4 is required for centriole duplication and flagella development. *Curr Biol,* 15**,** 2199-207.

BIJNSDORP, I. V., HODZIC, J., LAGERWEIJ, T., WESTERMAN, B., KRIJGSMAN, O., BROEKE, J., VERWEIJ, F., NILSSON, R. J., ROZENDAAL, L., VAN BEUSECHEM, V. W., VAN MOORSELAAR, J. A., WURDINGER, T. & GELDOF, A. A. 2016. miR-129-3p controls centrosome number in metastatic prostate cancer cells by repressing CP110. *Oncotarget,* 7**,** 16676-87.

BLACHON, S., GOPALAKRISHNAN, J., OMORI, Y., POLYANOVSKY, A., CHURCH, A., NICASTRO, D., MALICKI, J. & AVIDOR-REISS, T. 2008. Drosophila asterless and vertebrate Cep152 Are orthologs essential for centriole duplication. *Genetics,* 180**,** 2081-94.

BOBINNEC, Y., KHODJAKOV, A., MIR, L. M., RIEDER, C. L., EDDE, B. & BORNENS, M. 1998. Centriole disassembly in vivo and its effect on centrosome structure and function in vertebrate cells. *J Cell Biol,* 143**,** 1575-89.

BOESE, C. J., NYE, J., BUSTER, D. W., MCLAMARRAH, T. A., BYRNES, A. E., SLEP, K. C., RUSAN, N. M. & ROGERS, G. C. 2018. Asterless is a Polo-like kinase 4 substrate that both activates and inhibits kinase activity depending on its phosphorylation state. *Mol Biol Cell,* 29**,** 2874-2886.

BRESLOW, D. K. & HOLLAND, A. J. 2019. Mechanism and Regulation of Centriole and Cilium Biogenesis. *Annu Rev Biochem,* 88**,** 691-724.

BUCHAN, D. W. A. & JONES, D. T. 2019. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res,* 47**,** W402-W407.

CAJANEK, L. & NIGG, E. A. 2014. Cep164 triggers ciliogenesis by recruiting Tau tubulin kinase 2 to the mother centriole. *Proc Natl Acad Sci U S A,* 111**,** E2841-50.

CALLAINI, G. & RIPARBELLI, M. G. 1990. Centriole and centrosome cycle in the early Drosophila embryo. *J Cell Sci,* 97 (Pt 3)**,** 539-43.

CALLAINI, G., WHITFIELD, W. G. & RIPARBELLI, M. G. 1997. Centriole and centrosome dynamics during the embryonic cell cycles that follow the formation of the cellular blastoderm in Drosophila. *Exp Cell Res,* 234**,** 183-90.

CANTWELL, H. & NURSE, P. 2019. Unravelling nuclear size control. *Curr Genet,* 65**,** 1281-1285.

CAO, J., SHEN, Y., ZHU, L., XU, Y., ZHOU, Y., WU, Z., LI, Y., YAN, X. & ZHU, X. 2012. miR-129-3p controls cilia assembly by regulating CP110 and actin dynamics. *Nat Cell Biol,* 14**,** 697-706.

CARVALHO-SANTOS, Z., AZIMZADEH, J., PEREIRA-LEAL, J. B. & BETTENCOURT-DIAS, M. 2011. Evolution: Tracing the origins of centrioles, cilia, and flagella. *J Cell Biol,* 194**,** 165-75.

CARVALHO-SANTOS, Z., MACHADO, P., BRANCO, P., TAVARES-CADETE, F., RODRIGUES-MARTINS, A., PEREIRA-LEAL, J. B. & BETTENCOURT-DIAS, M. 2010. Stepwise evolution of the centriole-assembly pathway. *J Cell Sci,* 123**,** 1414-26.

CHANG, C. W., HSU, W. B., TSAI, J. J., TANG, C. J. & TANG, T. K. 2016. CEP295 interacts with microtubules and is required for centriole elongation. *J Cell Sci,* 129**,** 2501-13.

CHEN, C., INABA, M., VENKEI, Z. G. & YAMASHITA, Y. M. 2016. Klp10A, a stem cell centrosome-enriched kinesin, balances asymmetries in Drosophila male germline stem cell division. *Elife,* 5.

CHEN, H. Y., WU, C. T., TANG, C. C., LIN, Y. N., WANG, W. J. & TANG, T. K. 2017. Human microcephaly protein RTTN interacts with STIL and is required to build full-length centrioles. *Nat Commun,* 8**,** 247.

CHEN, Z., INDJEIAN, V. B., MCMANUS, M., WANG, L. & DYNLACHT, B. D. 2002. CP110, a cell cycle-dependent CDK substrate, regulates centrosome duplication in human cells. *Dev Cell,* 3**,** 339-50.

CIZMECIOGLU, O., ARNOLD, M., BAHTZ, R., SETTELE, F., EHRET, L., HASELMANN-WEISS, U., ANTONY, C. & HOFFMANN, I. 2010. Cep152 acts as a scaffold for recruitment of Plk4 and CPAP to the centrosome. *J Cell Biol,* 191**,** 731-9.

COMARTIN, D., GUPTA, G. D., FUSSNER, E., COYAUD, E., HASEGAN, M., ARCHINTI, M., CHEUNG, S. W., PINCHEV, D., LAWO, S., RAUGHT, B., BAZETT-JONES, D. P., LUDERS, J. & PELLETIER, L. 2013. CEP120 and SPICE1 cooperate with CPAP in centriole elongation. *Curr Biol,* 23**,** 1360-6.

CONDUIT, P. T., BRUNK, K., DOBBELAERE, J., DIX, C. I., LUCAS, E. P. & RAFF, J. W. 2010. Centrioles regulate centrosome size by controlling the rate of Cnn incorporation into the PCM. *Curr Biol,* 20**,** 2178-86.

CONDUIT, P. T., FENG, Z., RICHENS, J. H., BAUMBACH, J., WAINMAN, A., BAKSHI, S. D., DOBBELAERE, J., JOHNSON, S., LEA, S. M. & RAFF, J. W. 2014a. The centrosome-specific phosphorylation of Cnn by Polo/Plk1 drives Cnn scaffold assembly and centrosome maturation. *Dev Cell,* 28**,** 659-69.

CONDUIT, P. T., RICHENS, J. H., WAINMAN, A., HOLDER, J., VICENTE, C. C., PRATT, M. B., DIX, C. I., NOVAK, Z. A., DOBBIE, I. M., SCHERMELLEH, L. & RAFF, J. W. 2014b. A molecular mechanism of mitotic centrosome assembly in Drosophila. *Elife,* 3**,** e03399.

CONDUIT, P. T., WAINMAN, A., NOVAK, Z. A., WEIL, T. T. & RAFF, J. W. 2015a. Re-examining the role of Drosophila Sas-4 in centrosome assembly using two-colour-3D-SIM FRAP. *Elife,* 4.

CONDUIT, P. T., WAINMAN, A. & RAFF, J. W. 2015b. Centrosome function and assembly in animal cells. *Nat Rev Mol Cell Biol,* 16**,** 611-24.

CORMIER, A., CLEMENT, M. J., KNOSSOW, M., LACHKAR, S., SAVARIN, P., TOMA, F., SOBEL, A., GIGANT, B. & CURMI, P. A. 2009. The PN2-3 domain of centrosomal P4.1-associated protein implements a novel mechanism for tubulin sequestration. *J Biol Chem,* 284**,** 6909-17.

COTTEE, M. A., MUSCHALIK, N., JOHNSON, S., LEVESON, J., RAFF, J. W. & LEA, S. M. 2015. The homo-oligomerisation of both Sas-6 and Ana2 is required for efficient centriole assembly in flies. *Elife,* 4**,** e07236.

COTTEE, M. A., MUSCHALIK, N., WONG, Y. L., JOHNSON, C. M., JOHNSON, S., ANDREEVA, A., OEGEMA, K., LEA, S. M., RAFF, J. W. & VAN BREUGEL, M. 2013. Crystal structures of the CPAP/STIL complex reveal its role in centriole assembly and human microcephaly. *Elife,* 2**,** e01071.

CRAIGE, B., TSAO, C. C., DIENER, D. R., HOU, Y., LECHTRECK, K. F., ROSENBAUM, J. L. & WITMAN, G. B. 2010. CEP290 tethers flagellar transition zone microtubules to the membrane and regulates flagellar protein content. *J Cell Biol,* 190**,** 927-40.

CUNHA-FERREIRA, I., BENTO, I., PIMENTA-MARQUES, A., JANA, S. C., LINCE-FARIA, M., DUARTE, P., BORREGO-PINTO, J., GILBERTO, S., AMADO, T., BRITO, D., RODRIGUES-MARTINS, A., DEBSKI, J., DZHINDZHEV, N. & BETTENCOURT-DIAS, M. 2013. Regulation of autophosphorylation controls PLK4 self-destruction and centriole number. *Curr Biol,* 23**,** 2245-2254.

CUNHA-FERREIRA, I., RODRIGUES-MARTINS, A., BENTO, I., RIPARBELLI, M., ZHANG, W., LAUE, E., CALLAINI, G., GLOVER, D. M. & BETTENCOURT-DIAS, M. 2009. The SCF/Slimb ubiquitin ligase limits centrosome amplification through degradation of SAK/PLK4. *Curr Biol,* 19**,** 43-9.

CZARNECKI, P. G. & SHAH, J. V. 2012. The ciliary transition zone: from morphology and molecules to medicine. *Trends Cell Biol,* 22**,** 201-10.

D'ANGIOLELLA, V., DONATO, V., VIJAYAKUMAR, S., SARAF, A., FLORENS, L., WASHBURN, M. P., DYNLACHT, B. & PAGANO, M. 2010. SCF(Cyclin F) controls centrosome homeostasis and mitotic fidelity through CP110 degradation. *Nature,* 466**,** 138-42.

DAMMERMANN, A., MULLER-REICHERT, T., PELLETIER, L., HABERMANN, B., DESAI, A. & OEGEMA, K. 2004. Centriole assembly requires both centriolar and pericentriolar material proteins. *Dev Cell,* 7**,** 815-29.

DAVID, A., AMARTELY, H., RABINOWICZ, N., SHAMIR, M., FRIEDLER, A. & IZRAELI, S. 2016. Molecular basis of the STIL coiled coil oligomerization explains its requirement for de-novo formation of centrosomes in mammalian cells. *Sci Rep,* 6**,** 24296.

DELATTRE, M., CANARD, C. & GONCZY, P. 2006. Sequential protein recruitment in C. elegans centriole formation. *Curr Biol,* 16**,** 1844-9.

DELATTRE, M., LEIDEL, S., WANI, K., BAUMER, K., BAMAT, J., SCHNABEL, H., FEICHTINGER, R., SCHNABEL, R. & GONCZY, P. 2004. Centriolar SAS-5 is required for centrosome duplication in C. elegans. *Nat Cell Biol,* 6**,** 656-64.

DELGEHYR, N., RANGONE, H., FU, J., MAO, G., TOM, B., RIPARBELLI, M. G., CALLAINI, G. & GLOVER, D. M. 2012. Klp10A, a microtubule-depolymerizing kinesin-13, cooperates with CP110 to control Drosophila centriole length. *Curr Biol,* 22**,** 502-9.

DIX, C. I. & RAFF, J. W. 2007. Drosophila Spd-2 recruits PCM to the sperm centriole, but is dispensable for centriole duplication. *Curr Biol,* 17**,** 1759-64.

DOBBELAERE, J., JOSUE, F., SUIJKERBUIJK, S., BAUM, B., TAPON, N. & RAFF, J. 2008. A genome-wide RNAi screen to dissect centriole duplication and centrosome maturation in Drosophila. *PLoS Biol,* 6**,** e224.

DOBBELAERE, J., SCHMIDT CERNOHORSKA, M., HURANOVA, M., SLADE, D. & DAMMERMANN, A. 2020. Cep97 Is Required for Centriole Structural Integrity and Cilia Formation in Drosophila. *Curr Biol,* 30**,** 3045-3056 e7.

DZHINDZHEV, N. S., TZOLOVSKY, G., LIPINSZKI, Z., ABDELAZIZ, M., DEBSKI, J., DADLEZ, M. & GLOVER, D. M. 2017. Two-step phosphorylation of Ana2 by Plk4 is required for the sequential loading of Ana2 and Sas6 to initiate procentriole formation. *Open Biol,* 7.

DZHINDZHEV, N. S., TZOLOVSKY, G., LIPINSZKI, Z., SCHNEIDER, S., LATTAO, R., FU, J., DEBSKI, J., DADLEZ, M. & GLOVER, D. M. 2014. Plk4 phosphorylates Ana2 to trigger Sas6 recruitment and procentriole formation. *Curr Biol,* 24**,** 2526-32.

DZHINDZHEV, N. S., YU, Q. D., WEISKOPF, K., TZOLOVSKY, G., CUNHA-FERREIRA, I., RIPARBELLI, M., RODRIGUES-MARTINS, A., BETTENCOURT-DIAS, M., CALLAINI, G. & GLOVER, D. M. 2010. Asterless is a scaffold for the onset of centriole assembly. *Nature,* 467**,** 714-8.

EDGAR, B. A. & O'FARRELL, P. H. 1990. The three postblastoderm cell cycles of Drosophila embryogenesis are regulated in G2 by string. *Cell,* 62**,** 469-80.

EILERS, P. H. & BOELENS, H. F. 2005. Baseline correction with asymmetric least squares smoothing. *Leiden University Medical Centre Report,* 1**,** 5.

ELIA, A. E., CANTLEY, L. C. & YAFFE, M. B. 2003a. Proteomic screen finds pSer/pThr-binding domain localizing Plk1 to mitotic substrates. *Science,* 299**,** 1228-31.

ELIA, A. E., RELLOS, P., HAIRE, L. F., CHAO, J. W., IVINS, F. J., HOEPKER, K., MOHAMMAD, D., CANTLEY, L. C., SMERDON, S. J. & YAFFE, M. B. 2003b. The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain. *Cell,* 115**,** 83-95.

FENG, Z., CABALLE, A., WAINMAN, A., JOHNSON, S., HAENSELE, A. F. M., COTTEE, M. A., CONDUIT, P. T., LEA, S. M. & RAFF, J. W. 2017. Structural Basis for Mitotic Centrosome Assembly in Flies. *Cell,* 169**,** 1078-1089 e13.

FODE, C., MOTRO, B., YOUSEFI, S., HEFFERNAN, M. & DENNIS, J. W. 1994. Sak, a murine protein-serine/threonine kinase that is related to the Drosophila polo kinase and involved in cell proliferation. *Proc Natl Acad Sci U S A,* 91**,** 6388-92.

FOE, V. E. & ALBERTS, B. M. 1983. Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis. *J Cell Sci,* 61**,** 31-70.

FRANZ, A., ROQUE, H., SAURYA, S., DOBBELAERE, J. & RAFF, J. W. 2013. CP110 exhibits novel regulatory activities during centriole assembly in Drosophila. *J Cell Biol,* 203**,** 785-99.

FRIKSTAD, K. M., MOLINARI, E., THORESEN, M., RAMSBOTTOM, S. A., HUGHES, F., LETTEBOER, S. J. F., GILANI, S., SCHINK, K. O., STOKKE, T., GEIMER, S., PEDERSEN, L. B., GILES, R. H., AKHMANOVA, A., ROEPMAN, R., SAYER, J. A. & PATZKE, S. 2019. A CEP104-CSPP1 Complex Is Required for Formation of Primary Cilia Competent in Hedgehog Signaling. *Cell Rep,* 28**,** 1907-1922 e6.

FU, J. & GLOVER, D. M. 2012. Structured illumination of the interface between centriole and peri-centriolar material. *Open Biol,* 2**,** 120104.

FU, J., LIPINSZKI, Z., RANGONE, H., MIN, M., MYKURA, C., CHAO-CHU, J., SCHNEIDER, S., DZHINDZHEV, N. S., GOTTARDO, M., RIPARBELLI, M. G., CALLAINI, G. & GLOVER, D. M. 2016. Conserved molecular interactions in centriole-to-centrosome conversion. *Nat Cell Biol,* 18**,** 87-99.

GALLETTA, B. J., FAGERSTROM, C. J., SCHOBORG, T. A., MCLAMARRAH, T. A., RYNIAWEC, J. M., BUSTER, D. W., SLEP, K. C., ROGERS, G. C. & RUSAN,

N. M. 2016a. A centrosome interactome provides insight into organelle assembly and reveals a non-duplication role for Plk4. *Nat Commun,* 7**,** 12476.
GALLETTA, B. J., JACOBS, K. C., FAGERSTROM, C. J. & RUSAN, N. M. 2016b. Asterless is required for centriole length control and sperm development. *J Cell Biol,* 213**,** 435-50.
GARTENMANN, L., WAINMAN, A., QURASHI, M., KAUFMANN, R., SCHUBERT, S., RAFF, J. W. & DOBBIE, I. M. 2017. A combined 3D-SIM/SMLM approach allows centriole proteins to be localized with a precision of approximately 4-5 nm. *Curr Biol,* 27**,** R1054-R1055.
GIBBONS, I. R. & GRIMSTONE, A. V. 1960. On flagellar structure in certain flagellates. *J Biophys Biochem Cytol,* 7**,** 697-716.
GILANI, S. 2018. *Generation of stable transgenic hTERT-RPE1 cell lines for imaging analysis of cilia tip proteins at high temporal and spatial resolution.*
GILULA, N. B. & SATIR, P. 1972. The ciliary necklace. A ciliary membrane specialization. *J Cell Biol,* 53**,** 494-509.
GODINHO, S. A. & PELLMAN, D. 2014. Causes and consequences of centrosome abnormalities in cancer. *Philos Trans R Soc Lond B Biol Sci,* 369.
GODINHO, S. A., PICONE, R., BURUTE, M., DAGHER, R., SU, Y., LEUNG, C. T., POLYAK, K., BRUGGE, J. S., THERY, M. & PELLMAN, D. 2014. Oncogene-like induction of cellular invasion from centrosome amplification. *Nature,* 510**,** 167-71.
GOETZ, S. C., LIEM, K. F., JR. & ANDERSON, K. V. 2012. The spinocerebellar ataxia-associated gene Tau tubulin kinase 2 controls the initiation of ciliogenesis. *Cell,* 151**,** 847-858.
GONZALEZ, C., TAVOSANIS, G. & MOLLINARI, C. 1998. Centrosomes and microtubule organisation during Drosophila development. *J Cell Sci,* 111 (Pt 18)**,** 2697-706.
GOSHIMA, G. & VALE, R. D. 2003. The roles of microtubule-based motor proteins in mitosis: comprehensive RNAi analysis in the Drosophila S2 cell line. *J Cell Biol,* 162**,** 1003-16.
GOSHIMA, G. & VALE, R. D. 2005. Cell cycle-dependent dynamics and regulation of mitotic kinesins in Drosophila S2 cells. *Mol Biol Cell,* 16**,** 3896-907.
GOSHIMA, G., WOLLMAN, R., GOODWIN, S. S., ZHANG, N., SCHOLEY, J. M., VALE, R. D. & STUURMAN, N. 2007. Genes required for mitotic spindle assembly in Drosophila S2 cells. *Science,* 316**,** 417-21.
GOTTARDO, M., CALLAINI, G. & RIPARBELLI, M. G. 2013. The cilium-like region of the Drosophila spermatocyte: an emerging flagellum? *J Cell Sci,* 126**,** 5441-52.
GOTTARDO, M., CALLAINI, G. & RIPARBELLI, M. G. 2015. The Drosophila centriole–conversion of doublets into triplets within the stem cell niche. *Journal of cell science,* 128**,** 2437-2442.
GOTTARDO, M., CALLAINI, G. & RIPARBELLI, M. G. 2016. Klp10A modulates the localization of centriole-associated proteins during Drosophila male gametogenesis. *Cell Cycle,* 15**,** 3432-3441.
GOTTARDO, M., PERSICO, V., CALLAINI, G. & RIPARBELLI, M. G. 2018. The "transition zone" of the cilium-like regions in the Drosophila spermatocytes and the role of the C-tubule in axoneme assembly. *Exp Cell Res,* 371**,** 262-268.
GRASER, S., STIERHOF, Y. D., LAVOIE, S. B., GASSNER, O. S., LAMLA, S., LE CLECH, M. & NIGG, E. A. 2007. Cep164, a novel centriole appendage protein required for primary cilium formation. *J Cell Biol,* 179**,** 321-30.

GRATZ, S. J., UKKEN, F. P., RUBINSTEIN, C. D., THIEDE, G., DONOHUE, L. K., CUMMINGS, A. M. & O'CONNOR-GILES, K. M. 2014. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics,* 196**,** 961-71.

GREENAN, G., BRANGWYNNE, C. P., JAENSCH, S., GHARAKHANI, J., JULICHER, F. & HYMAN, A. A. 2010. Centrosome size sets mitotic spindle length in Caenorhabditis elegans embryos. *Curr Biol,* 20**,** 353-8.

GREENAN, G. A., KESZTHELYI, B., VALE, R. D. & AGARD, D. A. 2018. Insights into centriole geometry revealed by cryotomography of doublet and triplet centrioles. *Elife,* 7.

GUDERIAN, G., WESTENDORF, J., ULDSCHMID, A. & NIGG, E. A. 2010. Plk4 trans-autophosphorylation regulates centriole number by controlling betaTrCP-mediated degradation. *J Cell Sci,* 123**,** 2163-9.

GUICHARD, P., CHRETIEN, D., MARCO, S. & TASSIN, A. M. 2010. Procentriole assembly revealed by cryo-electron tomography. *EMBO J,* 29**,** 1565-72.

GUICHARD, P., DESFOSSES, A., MAHESHWARI, A., HACHET, V., DIETRICH, C., BRUNE, A., ISHIKAWA, T., SACHSE, C. & GONCZY, P. 2012. Cartwheel architecture of Trichonympha basal body. *Science,* 337**,** 553.

GUICHARD, P., HACHET, V., MAJUBU, N., NEVES, A., DEMURTAS, D., OLIERIC, N., FLUCKIGER, I., YAMADA, A., KIHARA, K., NISHIDA, Y., MORIYA, S., STEINMETZ, M. O., HONGOH, Y. & GONCZY, P. 2013. Native architecture of the centriole proximal region reveals features underlying its 9-fold radial symmetry. *Curr Biol,* 23**,** 1620-8.

GUICHARD, P., HAMEL, V., LE GUENNEC, M., BANTERLE, N., IACOVACHE, I., NEMCIKOVA, V., FLUCKIGER, I., GOLDIE, K. N., STAHLBERG, H., LEVY, D., ZUBER, B. & GONCZY, P. 2017. Cell-free reconstitution reveals centriole cartwheel assembly mechanisms. *Nat Commun,* 8**,** 14813.

HABEDANCK, R., STIERHOF, Y. D., WILKINSON, C. J. & NIGG, E. A. 2005. The Polo kinase Plk4 functions in centriole duplication. *Nat Cell Biol,* 7**,** 1140-6.

HATCH, E. M., KULUKIAN, A., HOLLAND, A. J., CLEVELAND, D. W. & STEARNS, T. 2010. Cep152 interacts with Plk4 and is required for centriole duplication. *J Cell Biol,* 191**,** 721-9.

HATZOPOULOS, G. N., ERAT, M. C., CUTTS, E., ROGALA, K. B., SLATER, L. M., STANSFELD, P. J. & VAKONAKIS, I. 2013. Structural analysis of the G-box domain of the microcephaly protein CPAP suggests a role in centriole architecture. *Structure,* 21**,** 2069-77.

HEYDECK, W. 2016. Analysis of Two Centrin-Binding Proteins, Poc5 and Sfr1, in Tetrahymena Thermophila Basal Bodies.

HILBERT, M., ERAT, M. C., HACHET, V., GUICHARD, P., BLANK, I. D., FLUCKIGER, I., SLATER, L., LOWE, E. D., HATZOPOULOS, G. N., STEINMETZ, M. O., GONCZY, P. & VAKONAKIS, I. 2013. Caenorhabditis elegans centriolar protein SAS-6 forms a spiral that is consistent with imparting a ninefold symmetry. *Proc Natl Acad Sci U S A,* 110**,** 11373-8.

HIRAKI, M., NAKAZAWA, Y., KAMIYA, R. & HIRONO, M. 2007. Bld10p constitutes the cartwheel-spoke tip and stabilizes the 9-fold symmetry of the centriole. *Curr Biol,* 17**,** 1778-83.

HIRONO, M. 2014. Cartwheel assembly. *Philos Trans R Soc Lond B Biol Sci,* 369.

HODGES, M. E., SCHEUMANN, N., WICKSTEAD, B., LANGDALE, J. A. & GULL, K. 2010. Reconstructing the evolutionary history of the centriole from protein components. *J Cell Sci,* 123**,** 1407-13.

HOLLAND, A. J., FACHINETTI, D., ZHU, Q., BAUER, M., VERMA, I. M., NIGG, E. A. & CLEVELAND, D. W. 2012. The autoregulated instability of Polo-like kinase 4 limits centrosome duplication to once per cell cycle. *Genes Dev,* 26**,** 2684-9.

HOLLAND, A. J., LAN, W., NIESSEN, S., HOOVER, H. & CLEVELAND, D. W. 2010. Polo-like kinase 4 kinase activity limits centrosome overduplication by autoregulating its own stability. *J Cell Biol,* 188**,** 191-8.

HOUSDEN, B. E., MUHAR, M., GEMBERLING, M., GERSBACH, C. A., STAINIER, D. Y., SEYDOUX, G., MOHR, S. E., ZUBER, J. & PERRIMON, N. 2017. Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. *Nat Rev Genet,* 18**,** 24-40.

HSU, W. B., HUNG, L. Y., TANG, C. J., SU, C. L., CHANG, Y. & TANG, T. K. 2008. Functional characterization of the microtubule-binding and -destabilizing domains of CPAP and d-SAS-4. *Exp Cell Res,* 314**,** 2591-602.

HU, S., DANILOV, A. V., GODEK, K., ORR, B., TAFE, L. J., RODRIGUEZ-CANALES, J., BEHRENS, C., MINO, B., MORAN, C. A., MEMOLI, V. A., MUSTACHIO, L. M., GALIMBERTI, F., RAVI, S., DECASTRO, A., LU, Y., SEKULA, D., ANDREW, A. S., WISTUBA, II, FREEMANTLE, S., COMPTON, D. A. & DMITROVSKY, E. 2015a. CDK2 Inhibition Causes Anaphase Catastrophe in Lung Cancer through the Centrosomal Protein CP110. *Cancer Res,* 75**,** 2029-38.

HU, S., LU, Y., ORR, B., GODEK, K., MUSTACHIO, L. M., KAWAKAMI, M., SEKULA, D., COMPTON, D. A., FREEMANTLE, S. & DMITROVSKY, E. 2015b. Specific CP110 Phosphorylation Sites Mediate Anaphase Catastrophe after CDK2 Inhibition: Evidence for Cooperation with USP33 Knockdown. *Mol Cancer Ther,* 14**,** 2576-85.

HUANG, J. Y. & RAFF, J. W. 1999. The disappearance of cyclin B at the end of mitosis is regulated spatially in Drosophila cells. *The EMBO journal,* 18**,** 2184-2195.

HUANG, N., ZHANG, D., LI, F., CHAI, P., WANG, S., TENG, J. & CHEN, J. 2018. M-Phase Phosphoprotein 9 regulates ciliogenesis by modulating CP110-CEP97 complex localization at the mother centriole. *Nat Commun,* 9**,** 4511.

HUNG, L. Y., CHEN, H. L., CHANG, C. W., LI, B. R. & TANG, T. K. 2004. Identification of a novel microtubule-destabilizing motif in CPAP that binds to tubulin heterodimers and inhibits microtubule assembly. *Mol Biol Cell,* 15**,** 2697-706.

INABA, M., BUSZCZAK, M. & YAMASHITA, Y. M. 2015. Nanotubes mediate niche-stem-cell signalling in the Drosophila testis. *Nature,* 523**,** 329-32.

INOUE, S. & SALMON, E. D. 1995. Force generation by microtubule assembly/disassembly in mitosis and related movements. *Mol Biol Cell,* 6**,** 1619-40.

IZQUIERDO, D., WANG, W. J., URYU, K. & TSOU, M. F. 2014. Stabilization of cartwheel-less centrioles for duplication requires CEP295-mediated centriole-to-centrosome conversion. *Cell Rep,* 8**,** 957-65.

JAKOBSEN, L., VANSELOW, K., SKOGS, M., TOYODA, Y., LUNDBERG, E., POSER, I., FALKENBY, L. G., BENNETZEN, M., WESTENDORF, J., NIGG, E. A., UHLEN, M., HYMAN, A. A. & ANDERSEN, J. S. 2011. Novel asymmetrically localizing components of human centrosomes identified by complementary proteomics methods. *EMBO J,* 30**,** 1520-35.

JANKE, C. & MAGIERA, M. M. 2020. The tubulin code and its role in controlling microtubule properties and functions. *Nat Rev Mol Cell Biol,* 21**,** 307-326.

JIANG, K., TOEDT, G., MONTENEGRO GOUVEIA, S., DAVEY, N. E., HUA, S., VAN DER VAART, B., GRIGORIEV, I., LARSEN, J., PEDERSEN, L. B., BEZSTAROSTI, K., LINCE-FARIA, M., DEMMERS, J., STEINMETZ, M. O., GIBSON, T. J. & AKHMANOVA, A. 2012. A Proteome-wide screen for mammalian SxIP motif-containing microtubule plus-end tracking proteins. *Curr Biol,* 22**,** 1800-7.

JONES, D. T. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol,* 292**,** 195-202.

JONES, D. T. & COZZETTO, D. 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics,* 31**,** 857-63.

JORGENSEN, P., EDGINGTON, N. P., SCHNEIDER, B. L., RUPES, I., TYERS, M. & FUTCHER, B. 2007. The size of the nucleus increases as yeast cells grow. *Mol Biol Cell,* 18**,** 3523-32.

KELLER, L. C., GEIMER, S., ROMIJN, E., YATES, J., 3RD, ZAMORA, I. & MARSHALL, W. F. 2009. Molecular architecture of the centriole proteome: the conserved WD40 domain protein POC1 is required for centriole duplication and length control. *Mol Biol Cell,* 20**,** 1150-66.

KEMP, C. A., KOPISH, K. R., ZIPPERLEN, P., AHRINGER, J. & O'CONNELL, K. F. 2004. Centrosome maturation and duplication in C. elegans require the coiled-coil protein SPD-2. *Dev Cell,* 6**,** 511-23.

KERNAN, M. J. 2007. Mechanotransduction and auditory transduction in Drosophila. *Pflugers Arch,* 454**,** 703-20.

KHIRE, A., JO, K. H., KONG, D., AKHSHI, T., BLACHON, S., CEKIC, A. R., HYNEK, S., HA, A., LONCAREK, J., MENNELLA, V. & AVIDOR-REISS, T. 2016. Centriole Remodeling during Spermiogenesis in Drosophila. *Curr Biol,* 26**,** 3183-3189.

KIM, J., LEE, K. & RHEE, K. 2015a. PLK1 regulation of PCNT cleavage ensures fidelity of centriole separation during mitotic exit. *Nat Commun,* 6**,** 10076.

KIM, M. K., DUDOGNON, C. & SMITH, S. 2012. Tankyrase 1 regulates centrosome function by controlling CPAP stability. *EMBO Rep,* 13**,** 724-32.

KIM, S., LEE, K., CHOI, J. H., RINGSTAD, N. & DYNLACHT, B. D. 2015b. Nek2 activation of Kif24 ensures cilium disassembly during the cell cycle. *Nat Commun,* 6**,** 8087.

KIM, T. S., PARK, J. E., SHUKLA, A., CHOI, S., MURUGAN, R. N., LEE, J. H., AHN, M., RHEE, K., BANG, J. K., KIM, B. Y., LONCAREK, J., ERIKSON, R. L. & LEE, K. S. 2013. Hierarchical recruitment of Plk4 and regulation of centriole biogenesis by two centrosomal scaffolds, Cep192 and Cep152. *Proc Natl Acad Sci U S A,* 110**,** E4849-57.

KIRKHAM, M., MULLER-REICHERT, T., OEGEMA, K., GRILL, S. & HYMAN, A. A. 2003. SAS-4 is a C. elegans centriolar protein that controls centrosome size. *Cell,* 112**,** 575-87.

KIRSCHNER, M. W. & MITCHISON, T. 1986. Microtubule dynamics. *Nature,* 324**,** 621-621.

KISIELEWSKA, J., LU, P. & WHITAKER, M. 2005. GFP-PCNA as an S-phase marker in embryos during the first and subsequent cell cycles. *Biol Cell,* 97**,** 221-9.

KITAGAWA, D., VAKONAKIS, I., OLIERIC, N., HILBERT, M., KELLER, D., OLIERIC, V., BORTFELD, M., ERAT, M. C., FLUCKIGER, I., GONCZY, P. & STEINMETZ, M. O. 2011. Structural basis of the 9-fold symmetry of centrioles. *Cell,* 144**,** 364-75.

KLEBBA, J. E., BUSTER, D. W., NGUYEN, A. L., SWATKOSKI, S., GUCEK, M., RUSAN, N. M. & ROGERS, G. C. 2013. Polo-like kinase 4 autodestructs by generating its Slimb-binding phosphodegron. *Curr Biol,* 23**,** 2255-2261.
KLEYLEIN-SOHN, J., WESTENDORF, J., LE CLECH, M., HABEDANCK, R., STIERHOF, Y. D. & NIGG, E. A. 2007. Plk4-induced centriole biogenesis in human cells. *Dev Cell,* 13**,** 190-202.
KOBAYASHI, T., TSANG, W. Y., LI, J., LANE, W. & DYNLACHT, B. D. 2011. Centriolar kinesin Kif24 interacts with CP110 to remodel microtubules and regulate ciliogenesis. *Cell,* 145**,** 914-25.
KOCHANSKI, R. S. & BORISY, G. G. 1990. Mode of centriole duplication and distribution. *J Cell Biol,* 110**,** 1599-605.
KOHLMAIER, G., LONCAREK, J., MENG, X., MCEWEN, B. F., MOGENSEN, M. M., SPEKTOR, A., DYNLACHT, B. D., KHODJAKOV, A. & GONCZY, P. 2009. Overly long centrioles and defective cell division upon excess of the SAS-4-related protein CPAP. *Curr Biol,* 19**,** 1012-8.
KONG, D., FARMER, V., SHUKLA, A., JAMES, J., GRUSKIN, R., KIRIYAMA, S. & LONCAREK, J. 2014. Centriole maturation requires regulated Plk1 activity during two consecutive cell cycles. *J Cell Biol,* 206**,** 855-65.
KONG, D., SAHABANDU, N., SULLENBERGER, C., VASQUEZ-LIMETA, A., LUVSANJAV, D., LUKASIK, K. & LONCAREK, J. 2020. Prolonged mitosis results in structurally aberrant and over-elongated centrioles. *J Cell Biol,* 219.
KRATZ, A. S., BARENZ, F., RICHTER, K. T. & HOFFMANN, I. 2015. Plk4-dependent phosphorylation of STIL is required for centriole duplication. *Biol Open,* 4**,** 370-7.
LA TERRA, S., ENGLISH, C. N., HERGERT, P., MCEWEN, B. F., SLUDER, G. & KHODJAKOV, A. 2005. The de novo centriole assembly pathway in HeLa cells: cell cycle progression and centriole assembly/maturation. *J Cell Biol,* 168**,** 713-22.
LAGUNOFF, D. & CHI, E. Y. 1976. Effect of colchicine on rat mast cells. *J Cell Biol,* 71**,** 182-95.
LANE, H. A. & NIGG, E. A. 1996. Antibody microinjection reveals an essential role for human polo-like kinase 1 (Plk1) in the functional maturation of mitotic centrosomes. *J Cell Biol,* 135**,** 1701-13.
LARKIN, A., MARYGOLD, S. J., ANTONAZZO, G., ATTRILL, H., DOS SANTOS, G., GARAPATI, P. V., GOODMAN, JOSHUA L., GRAMATES, L S., MILLBURN, G., STRELETS, V. B., TABONE, C. J., THURMOND, J. & CONSORTIUM, F. 2020. FlyBase: updates to the Drosophila melanogaster knowledge base. *Nucleic Acids Research,* 49**,** D899-D907.
LAYCOCK, J. E., SAVOIAN, M. S. & GLOVER, D. M. 2006. Antagonistic activities of Klp10A and Orbit regulate spindle length, bipolarity and function in vivo. *J Cell Sci,* 119**,** 2354-61.
LEE, H. S., SIMON, J. A. & LIS, J. T. 1988. Structure and expression of ubiquitin genes of Drosophila melanogaster. *Mol Cell Biol,* 8**,** 4727-35.
LEE, K. S., GRENFELL, T. Z., YARM, F. R. & ERIKSON, R. L. 1998. Mutation of the polo-box disrupts localization and mitotic functions of the mammalian polo kinase Plk. *Proc Natl Acad Sci U S A,* 95**,** 9301-6.
LEE, M., SEO, M. Y., CHANG, J., HWANG, D. S. & RHEE, K. 2017. PLK4 phosphorylation of CP110 is required for efficient centriole assembly. *Cell Cycle,* 16**,** 1225-1234.
LEGG, K. 2019. *The spatio-temporal regulation of Polo-like kinase 1 in DNA repair.* University of Oxford.

LEIDEL, S., DELATTRE, M., CERUTTI, L., BAUMER, K. & GONCZY, P. 2005. SAS-6 defines a protein family required for centrosome duplication in C. elegans and in human cells. *Nat Cell Biol,* 7**,** 115-25.

LEIDEL, S. & GONCZY, P. 2003. SAS-4 is essential for centrosome duplication in C elegans and is recruited to daughter centrioles once per cell cycle. *Dev Cell,* 4**,** 431-9.

LETTMAN, M. M., WONG, Y. L., VISCARDI, V., NIESSEN, S., CHEN, S. H., SHIAU, A. K., ZHOU, H., DESAI, A. & OEGEMA, K. 2013. Direct binding of SAS-6 to ZYG-1 recruits SAS-6 to the mother centriole for cartwheel assembly. *Dev Cell,* 25**,** 284-98.

LEUNG, G. C., HUDSON, J. W., KOZAROVA, A., DAVIDSON, A., DENNIS, J. W. & SICHERI, F. 2002. The Sak polo-box comprises a structural domain sufficient for mitotic subcellular localization. *Nat Struct Biol,* 9**,** 719-24.

LEVY, D. L. & HEALD, R. 2010. Nuclear size is regulated by importin alpha and Ntf2 in Xenopus. *Cell,* 143**,** 288-98.

LI, J., D'ANGIOLELLA, V., SEELEY, E. S., KIM, S., KOBAYASHI, T., FU, W., CAMPOS, E. I., PAGANO, M. & DYNLACHT, B. D. 2013. USP33 regulates centrosome biogenesis via deubiquitination of the centriolar protein CP110. *Nature,* 495**,** 255-9.

LIN, Y. C., CHANG, C. W., HSU, W. B., TANG, C. J., LIN, Y. N., CHOU, E. J., WU, C. T. & TANG, T. K. 2013a. Human microcephaly protein CEP135 binds to hSAS-6 and CPAP, and is required for centriole assembly. *EMBO J,* 32**,** 1141-54.

LIN, Y. N., WU, C. T., LIN, Y. C., HSU, W. B., TANG, C. J., CHANG, C. W. & TANG, T. K. 2013b. CEP120 interacts with CPAP and positively regulates centriole elongation. *J Cell Biol,* 202**,** 211-9.

LINDSLEY, D. L. & ZIMM, G. G. 1992. *The genome of drosophila melanogaster*, Academic.

LIU, J., LEWELLYN, A. L., CHEN, L. G. & MALLER, J. L. 2004. The polo box is required for multiple functions of Plx1 in mitosis. *J Biol Chem,* 279**,** 21367-73.

LLAMAZARES, S., MOREIRA, A., TAVARES, A., GIRDHAM, C., SPRUCE, B. A., GONZALEZ, C., KARESS, R. E., GLOVER, D. M. & SUNKEL, C. E. 1991. polo encodes a protein kinase homolog required for mitosis in Drosophila. *Genes Dev,* 5**,** 2153-65.

LONCAREK, J., HERGERT, P. & KHODJAKOV, A. 2010. Centriole reduplication during prolonged interphase requires procentriole maturation governed by Plk1. *Curr Biol,* 20**,** 1277-82.

LOPES, C. A., JANA, S. C., CUNHA-FERREIRA, I., ZITOUNI, S., BENTO, I., DUARTE, P., GILBERTO, S., FREIXO, F., GUERRERO, A., FRANCIA, M., LINCE-FARIA, M., CARNEIRO, J. & BETTENCOURT-DIAS, M. 2015. PLK4 trans-Autoactivation Controls Centriole Biogenesis in Space. *Dev Cell,* 35**,** 222-35.

LOUKA, P., VASUDEVAN, K. K., GUHA, M., JOACHIMIAK, E., WLOGA, D., TOMASI, R. F., BAROUD, C. N., DUPUIS-WILLIAMS, P., GALATI, D. F., PEARSON, C. G., RICE, L. M., MORESCO, J. J., YATES, J. R., 3RD, JIANG, Y. Y., LECHTRECK, K., DENTLER, W. & GAERTIG, J. 2018. Proteins that control the geometry of microtubules at the ends of cilia. *J Cell Biol,* 217**,** 4298-4313.

LUO, M., CAO, L., CAO, Z., MA, S., SHEN, Y., YANG, D., LU, C., LIN, Z., LIU, Z., YU, Y., CAI, R., CHEN, C., GAO, H., WANG, X., CAO, M. & MA, X. 2019. Whole exome sequencing reveals novel CEP104 mutations in a Chinese patient with Joubert syndrome. *Mol Genet Genomic Med,* 7**,** e1004.

MADEIRA, F., PARK, Y. M., LEE, J., BUSO, N., GUR, T., MADHUSOODANAN, N., BASUTKAR, P., TIVEY, A. R. N., POTTER, S. C., FINN, R. D. & LOPEZ, R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res,* 47**,** W636-W641.

MARSHALL, W. F. 2015. How Cells Measure Length on Subcellular Scales. *Trends Cell Biol,* 25**,** 760-768.

MARTEIL, G., GUERRERO, A., VIEIRA, A. F., DE ALMEIDA, B. P., MACHADO, P., MENDONCA, S., MESQUITA, M., VILLARREAL, B., FONSECA, I., FRANCIA, M. E., DORES, K., MARTINS, N. P., JANA, S. C., TRANFIELD, E. M., BARBOSA-MORAIS, N. L., PAREDES, J., PELLMAN, D., GODINHO, S. A. & BETTENCOURT-DIAS, M. 2018. Over-elongation of centrioles in cancer promotes centriole amplification and chromosome missegregation. *Nat Commun,* 9**,** 1258.

MAURER, S. P., BIELING, P., COPE, J., HOENGER, A. & SURREY, T. 2011. GTPgammaS microtubules mimic the growing microtubule end structure recognized by end-binding proteins (EBs). *Proc Natl Acad Sci U S A,* 108**,** 3988-93.

MCCLELAND, M. L. & O'FARRELL, P. H. 2008. RNAi of mitotic cyclins in Drosophila uncouples the nuclear and centrosome cycle. *Curr Biol,* 18**,** 245-54.

MCLAMARRAH, T. A., BUSTER, D. W., GALLETTA, B. J., BOESE, C. J., RYNIAWEC, J. M., HOLLINGSWORTH, N. A., BYRNES, A. E., BROWNLEE, C. W., SLEP, K. C., RUSAN, N. M. & ROGERS, G. C. 2018. An ordered pattern of Ana2 phosphorylation by Plk4 is required for centriole assembly. *J Cell Biol,* 217**,** 1217-1231.

MCLAMARRAH, T. A., SPEED, S. K., RYNIAWEC, J. M., BUSTER, D. W., FAGERSTROM, C. J., GALLETTA, B. J., RUSAN, N. M. & ROGERS, G. C. 2020. A molecular mechanism for the procentriole recruitment of Ana2. *J Cell Biol,* 219.

MEEHL, J. B., BAYLESS, B. A., GIDDINGS, T. H., JR., PEARSON, C. G. & WINEY, M. 2016. Tetrahymena Poc1 ensures proper intertriplet microtubule linkages to maintain basal body integrity. *Mol Biol Cell,* 27**,** 2394-403.

MENNELLA, V., ROGERS, G. C., ROGERS, S. L., BUSTER, D. W., VALE, R. D. & SHARP, D. J. 2005. Functionally distinct kinesin-13 family members cooperate to regulate microtubule dynamics during interphase. *Nat Cell Biol,* 7**,** 235-45.

MITCHISON, T. & KIRSCHNER, M. 1984. Dynamic instability of microtubule growth. *Nature,* 312**,** 237-42.

MOORES, C. A. & MILLIGAN, R. A. 2006. Lucky 13-microtubule depolymerisation by kinesin-13 motors. *J Cell Sci,* 119**,** 3905-13.

MORITZ, M., BRAUNFELD, M. B., FUNG, J. C., SEDAT, J. W., ALBERTS, B. M. & AGARD, D. A. 1995. Three-dimensional structural characterization of centrosomes from early Drosophila embryos. *J Cell Biol,* 130**,** 1149-59.

MOUTINHO-SANTOS, T., SAMPAIO, P., AMORIM, I., COSTA, M. & SUNKEL, C. E. 1999. In vivo localisation of the mitotic POLO kinase shows a highly dynamic association with the mitotic apparatus during early embryogenesis in Drosophila. *Biol Cell,* 91**,** 585-96.

MOYER, T. C., CLUTARIO, K. M., LAMBRUS, B. G., DAGGUBATI, V. & HOLLAND, A. J. 2015. Binding of STIL to Plk4 activates kinase activity to promote centriole assembly. *J Cell Biol,* 209**,** 863-78.

MOYER, T. C. & HOLLAND, A. J. 2019. PLK4 promotes centriole duplication by phosphorylating STIL to link the procentriole cartwheel to the microtubule wall. *Elife,* 8.
NABAIS, C., PENEDA, C. & BETTENCOURT-DIAS, M. 2020. Evolution of centriole assembly. *Curr Biol,* 30**,** R494-R502.
NABAIS, C., PESSOA, D., DE-CARVALHO, J., VAN ZANTEN, T., DUARTE, P., MAYOR, S., CARNEIRO, J., TELLEY, I. A. & BETTENCOURT-DIAS, M. 2021. Plk4 triggers autonomous de novo centriole biogenesis and maturation. *J Cell Biol,* 220.
NAGAI, T., MUKOYAMA, S., KAGIWADA, H., GOSHIMA, N. & MIZUNO, K. 2018. Cullin-3-KCTD10-mediated CEP97 degradation promotes primary cilium formation. *J Cell Sci,* 131.
NAKAZAWA, Y., HIRAKI, M., KAMIYA, R. & HIRONO, M. 2007. SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole. *Curr Biol,* 17**,** 2169-74.
NCBI-RESOURCE-COORDINATORS 2018. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res,* 46**,** D8-D13.
NEUMANN, F. R. & NURSE, P. 2007. Nuclear size control in fission yeast. *J Cell Biol,* 179**,** 593-600.
NIGG, E. A. & HOLLAND, A. J. 2018. Once and only once: mechanisms of centriole duplication and their deregulation in disease. *Nat Rev Mol Cell Biol,* 19**,** 297-312.
NIGG, E. A. & RAFF, J. W. 2009. Centrioles, centrosomes, and cilia in health and disease. *Cell,* 139**,** 663-78.
NOVAK, Z. A., CONDUIT, P. T., WAINMAN, A. & RAFF, J. W. 2014. Asterless licenses daughter centrioles to duplicate for the first time in Drosophila embryos. *Curr Biol,* 24**,** 1276-82.
NOVAK, Z. A., WAINMAN, A., GARTENMANN, L. & RAFF, J. W. 2016. Cdk1 Phosphorylates Drosophila Sas-4 to Recruit Polo to Daughter Centrioles and Convert Them to Centrosomes. *Dev Cell,* 37**,** 545-57.
O'CONNELL, K. F., CARON, C., KOPISH, K. R., HURD, D. D., KEMPHUES, K. J., LI, Y. & WHITE, J. G. 2001. The C. elegans zyg-1 gene encodes a regulator of centrosome duplication with distinct maternal and paternal roles in the embryo. *Cell,* 105**,** 547-58.
ODA, T., CHIBA, S., NAGAI, T. & MIZUNO, K. 2014. Binding to Cep164, but not EB1, is essential for centriolar localization of TTBK2 and its function in ciliogenesis. *Genes Cells,* 19**,** 927-40.
OGAWA, T., NITTA, R., OKADA, Y. & HIROKAWA, N. 2004. A common mechanism for microtubule destabilizers-M type kinesins stabilize curling of the protofilament using the class-specific neck and loops. *Cell,* 116**,** 591-602.
OHTA, M., ASHIKAWA, T., NOZAKI, Y., KOZUKA-HATA, H., GOTO, H., INAGAKI, M., OYAMA, M. & KITAGAWA, D. 2014. Direct interaction of Plk4 with STIL ensures formation of a single procentriole per parental centriole. *Nat Commun,* 5**,** 5267.
OHTA, M., WATANABE, K., ASHIKAWA, T., NOZAKI, Y., YOSHIBA, S., KIMURA, A. & KITAGAWA, D. 2018. Bimodal Binding of STIL to Plk4 Controls Proper Centriole Copy Number. *Cell Rep,* 23**,** 3160-3169 e4.
OLIVEIRA, R. A., HAMILTON, R. S., PAULI, A., DAVIS, I. & NASMYTH, K. 2010. Cohesin cleavage and Cdk inhibition trigger formation of daughter nuclei. *Nat Cell Biol,* 12**,** 185-92.

PEARSON, C. G., OSBORN, D. P., GIDDINGS, T. H., JR., BEALES, P. L. & WINEY, M. 2009. Basal body stability and ciliogenesis requires the conserved component Poc1. *J Cell Biol,* 187**,** 905-20.
PEEL, N., STEVENS, N. R., BASTO, R. & RAFF, J. W. 2007. Overexpressing centriole-replication proteins in vivo induces centriole overduplication and de novo formation. *Curr Biol,* 17**,** 834-43.
PELLETIER, L., O'TOOLE, E., SCHWAGER, A., HYMAN, A. A. & MULLER-REICHERT, T. 2006. Centriole assembly in Caenorhabditis elegans. *Nature,* 444**,** 619-23.
PELLETIER, L., OZLU, N., HANNAK, E., COWAN, C., HABERMANN, B., RUER, M., MULLER-REICHERT, T. & HYMAN, A. A. 2004. The Caenorhabditis elegans centrosomal protein SPD-2 is required for both pericentriolar material recruitment and centriole duplication. *Curr Biol,* 14**,** 863-73.
PERSICO, V., CALLAINI, G. & RIPARBELLI, M. G. 2019. The Microtubule-Depolymerizing Kinesin-13 Klp10A Is Enriched in the Transition Zone of the Ciliary Structures of Drosophila melanogaster. *Front Cell Dev Biol,* 7**,** 173.
PINTARD, L. & ARCHAMBAULT, V. 2018. A unified view of spatio-temporal control of mitotic entry: Polo kinase as the key. *Open Biol,* 8.
PLOUGH, H. H. 1921. Further studies on the effect of temperature on crossing over. *Journal of Experimental Zoology,* 32**,** 187-202.
PORT, F., CHEN, H. M., LEE, T. & BULLOCK, S. L. 2014. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proc Natl Acad Sci U S A,* 111**,** E2967-76.
QIAO, R., CABRAL, G., LETTMAN, M. M., DAMMERMANN, A. & DONG, G. 2012. SAS-6 coiled-coil structure and interaction with SAS-5 suggest a regulatory mechanism in C. elegans centriole assembly. *EMBO J,* 31**,** 4334-47.
RADFORD, S. J., HARRISON, A. M. & MCKIM, K. S. 2012. Microtubule-depolymerizing kinesin KLP10A restricts the length of the acentrosomal meiotic spindle in Drosophila females. *Genetics,* 192**,** 431-40.
RAFF, J. W. 2019. Phase Separation and the Centrosome: A Fait Accompli? *Trends Cell Biol,* 29**,** 612-622.
RAMOS-VARA, J. A. 2005. Technical aspects of immunohistochemistry. *Vet Pathol,* 42**,** 405-26.
REITER, J. F., BLACQUE, O. E. & LEROUX, M. R. 2012. The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization. *EMBO Rep,* 13**,** 608-18.
REITER, J. F. & LEROUX, M. R. 2017. Genes and molecular pathways underpinning ciliopathies. *Nat Rev Mol Cell Biol,* 18**,** 533-547.
REYNOLDS, N. & OHKURA, H. 2003. Polo boxes form a single functional domain that mediates interactions with multiple proteins in fission yeast polo kinase. *J Cell Sci,* 116**,** 1377-87.
REZABKOVA, L., KRAATZ, S. H., AKHMANOVA, A., STEINMETZ, M. O. & KAMMERER, R. A. 2016. Biophysical and Structural Characterization of the Centriolar Protein Cep104 Interaction Network. *J Biol Chem,* 291**,** 18496-504.
RINGO, D. L. 1967. Flagellar motion and fine structure of the flagellar apparatus in Chlamydomonas. *J Cell Biol,* 33**,** 543-71.
RIPARBELLI, M. G., CALLAINI, G. & MEGRAW, T. L. 2012. Assembly and persistence of primary cilia in dividing Drosophila spermatocytes. *Dev Cell,* 23**,** 425-32.

ROBERTS, D. B. 1998. *Drosophila : a practical approach,* Oxford, IRL Press at Oxford University Press.
ROGALA, K. B., DYNES, N. J., HATZOPOULOS, G. N., YAN, J., PONG, S. K., ROBINSON, C. V., DEANE, C. M., GONCZY, P. & VAKONAKIS, I. 2015. The Caenorhabditis elegans protein SAS-5 forms large oligomeric assemblies critical for centriole formation. *Elife,* 4**,** e07410.
ROGERS, G. C., ROGERS, S. L., SCHWIMMER, T. A., EMS-MCCLUNG, S. C., WALCZAK, C. E., VALE, R. D., SCHOLEY, J. M. & SHARP, D. J. 2004. Two mitotic kinesins cooperate to drive sister chromatid separation during anaphase. *Nature,* 427**,** 364-70.
ROGERS, G. C., RUSAN, N. M., ROBERTS, D. M., PEIFER, M. & ROGERS, S. L. 2009. The SCF Slimb ubiquitin ligase regulates Plk4/Sak levels to block centriole reduplication. *J Cell Biol,* 184**,** 225-39.
ROQUE, H., WAINMAN, A., RICHENS, J., KOZYRSKA, K., FRANZ, A. & RAFF, J. W. 2012. Drosophila Cep135/Bld10 maintains proper centriole structure but is dispensable for cartwheel formation. *J Cell Sci,* 125**,** 5881-6.
SANTAMARIA, A., WANG, B., ELOWE, S., MALIK, R., ZHANG, F., BAUER, M., SCHMIDT, A., SILLJE, H. H., KORNER, R. & NIGG, E. A. 2011. The Plk1-dependent phosphoproteome of the early mitotic spindle. *Mol Cell Proteomics,* 10**,** M110 004457.
SATISH TAMMANA, T. V., TAMMANA, D., DIENER, D. R. & ROSENBAUM, J. 2013. Centrosomal protein CEP104 (Chlamydomonas FAP256) moves to the ciliary tip during ciliary assembly. *J Cell Sci,* 126**,** 5018-29.
SAURYA, S., ROQUE, H., NOVAK, Z. A., WAINMAN, A., AYDOGAN, M. G., VOLANAKIS, A., SIEBER, B., PINTO, D. M. & RAFF, J. W. 2016. Drosophila Ana1 is required for centrosome assembly and centriole elongation. *J Cell Sci,* 129**,** 2514-25.
SCHMIDT, T. I., KLEYLEIN-SOHN, J., WESTENDORF, J., LE CLECH, M., LAVOIE, S. B., STIERHOF, Y. D. & NIGG, E. A. 2009. Control of centriole length by CPAP and CP110. *Curr Biol,* 19**,** 1005-11.
SEONG, Y. S., KAMIJO, K., LEE, J. S., FERNANDEZ, E., KURIYAMA, R., MIKI, T. & LEE, K. S. 2002. A spindle checkpoint arrest and a cytokinesis failure by the dominant-negative polo-box domain of Plk1 in U-2 OS cells. *J Biol Chem,* 277**,** 32282-93.
SHANER, N. C., LAMBERT, G. G., CHAMMAS, A., NI, Y., CRANFILL, P. J., BAIRD, M. A., SELL, B. R., ALLEN, J. R., DAY, R. N., ISRAELSSON, M., DAVIDSON, M. W. & WANG, J. 2013. A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nat Methods,* 10**,** 407-9.
SHARMA, A., AHER, A., DYNES, N. J., FREY, D., KATRUKHA, E. A., JAUSSI, R., GRIGORIEV, I., CROISIER, M., KAMMERER, R. A., AKHMANOVA, A., GONCZY, P. & STEINMETZ, M. O. 2016. Centriolar CPAP/SAS-4 Imparts Slow Processive Microtubule Growth. *Dev Cell,* 37**,** 362-376.
SHARMA, A., GERARD, S. F., OLIERIC, N. & STEINMETZ, M. O. 2018. Cep120 promotes microtubule formation through a unique tubulin binding C2 domain. *J Struct Biol,* 203**,** 62-70.
SHARMA, A., OLIERIC, N. & STEINMETZ, M. O. 2021. Centriole length control. *Curr Opin Struct Biol,* 66**,** 89-95.
SHIMANOVSKAYA, E., QIAO, R., LESIGANG, J. & DONG, G. 2013. The SAS-5 N-terminal domain is a tetramer, with implications for centriole assembly in C. elegans. *Worm,* 2**,** e25214.

SHODA, T., YAMAZOE, K., TANAKA, Y., ASANO, Y. & INOUE, Y. H. 2021. Orbit/CLASP determines centriole length by antagonising Klp10A in Drosophila spermatocytes. *J Cell Sci*.

SHUKLA, A., KONG, D., SHARMA, M., MAGIDSON, V. & LONCAREK, J. 2015. Plk1 relieves centriole block to reduplication by promoting daughter centriole maturation. *Nat Commun,* 6**,** 8077.

SILLIBOURNE, J. E., TACK, F., VLOEMANS, N., BOECKX, A., THAMBIRAJAH, S., BONNET, P., RAMAEKERS, F. C., BORNENS, M. & GRAND-PERRET, T. 2010. Autophosphorylation of polo-like kinase 4 and its role in centriole duplication. *Mol Biol Cell,* 21**,** 547-61.

SLEVIN, L. K., NYE, J., PINKERTON, D. C., BUSTER, D. W., ROGERS, G. C. & SLEP, K. C. 2012. The structure of the plk4 cryptic polo box reveals two tandem polo boxes required for centriole duplication. *Structure,* 20**,** 1905-17.

SLEVIN, L. K., ROMES, E. M., DANDULAKIS, M. G. & SLEP, K. C. 2014. The mechanism of dynein light chain LC8-mediated oligomerization of the Ana2 centriole duplication factor. *J Biol Chem,* 289**,** 20727-39.

SONG, R., WALENTEK, P., SPONER, N., KLIMKE, A., LEE, J. S., DIXON, G., HARLAND, R., WAN, Y., LISHKO, P., LIZE, M., KESSEL, M. & HE, L. 2014. miR-34/449 miRNAs are required for motile ciliogenesis by repressing cp110. *Nature,* 510**,** 115-20.

SONG, S., GRENFELL, T. Z., GARFIELD, S., ERIKSON, R. L. & LEE, K. S. 2000. Essential function of the polo box of Cdc5 in subcellular localization and induction of cytokinetic structures. *Mol Cell Biol,* 20**,** 286-98.

SONNEN, K. F., GABRYJONCZYK, A. M., ANSELM, E., STIERHOF, Y. D. & NIGG, E. A. 2013. Human Cep192 and Cep152 cooperate in Plk4 recruitment and centriole duplication. *J Cell Sci,* 126**,** 3223-33.

SONNEN, K. F., SCHERMELLEH, L., LEONHARDT, H. & NIGG, E. A. 2012. 3D-structured illumination microscopy provides novel insight into architecture of human centrosomes. *Biol Open,* 1**,** 965-76.

SOROKIN, S. 1962. Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle cells. *J Cell Biol,* 15**,** 363-77.

SPEKTOR, A., TSANG, W. Y., KHOO, D. & DYNLACHT, B. D. 2007. Cep97 and CP110 suppress a cilia assembly program. *Cell,* 130**,** 678-90.

SROUR, M., HAMDAN, F. F., MCKNIGHT, D., DAVIS, E., MANDEL, H., SCHWARTZENTRUBER, J., MARTIN, B., PATRY, L., NASSIF, C., DIONNE-LAPORTE, A., OSPINA, L. H., LEMYRE, E., MASSICOTTE, C., LAFRAMBOISE, R., MARANDA, B., LABUDA, D., DECARIE, J. C., RYPENS, F., GOLDSHER, D., FALLET-BIANCO, C., SOUCY, J. F., LABERGE, A. M., MAFTEI, C., CARE4RARE CANADA, C., BOYCOTT, K., BRAIS, B., BOUCHER, R. M., ROULEAU, G. A., KATSANIS, N., MAJEWSKI, J., ELPELEG, O., KUKOLICH, M. K., SHALEV, S. & MICHAUD, J. L. 2015. Joubert Syndrome in French Canadians and Identification of Mutations in CEP104. *Am J Hum Genet,* 97**,** 744-53.

STEVENS, N. R., DOBBELAERE, J., BRUNK, K., FRANZ, A. & RAFF, J. W. 2010a. Drosophila Ana2 is a conserved centriole duplication factor. *J Cell Biol,* 188**,** 313-23.

STEVENS, N. R., ROQUE, H. & RAFF, J. W. 2010b. DSas-6 and Ana2 coassemble into tubules to promote centriole duplication and engagement. *Dev Cell,* 19**,** 913-9.

SWALLOW, C. J., KO, M. A., SIDDIQUI, N. U., HUDSON, J. W. & DENNIS, J. W. 2005. Sak/Plk4 and mitotic fidelity. *Oncogene,* 24**,** 306-12.

SYDOR, A. M., COYAUD, E., ROVELLI, C., LAURENT, E., LIU, H., RAUGHT, B. & MENNELLA, V. 2018. PPP1R35 is a novel centrosomal protein that regulates centriole length in concert with the microcephaly protein RTTN. *Elife,* 7.

TAKAO, D., WATANABE, K., KUROKI, K. & KITAGAWA, D. 2019. Feedback loops in the Plk4-STIL-HsSAS6 network coordinate site selection for procentriole formation. *Biol Open,* 8.

TANG, C. J., FU, R. H., WU, K. S., HSU, W. B. & TANG, T. K. 2009. CPAP is a cell-cycle regulated protein that controls centriole length. *Nat Cell Biol,* 11**,** 825-31.

TANG, C. J., LIN, S. Y., HSU, W. B., LIN, Y. N., WU, C. T., LIN, Y. C., CHANG, C. W., WU, K. S. & TANG, T. K. 2011. The human microcephaly protein STIL interacts with CPAP and is required for procentriole formation. *EMBO J,* 30**,** 4790-804.

TANOS, B. E., YANG, H. J., SONI, R., WANG, W. J., MACALUSO, F. P., ASARA, J. M. & TSOU, M. F. 2013. Centriole distal appendages promote membrane docking, leading to cilia initiation. *Genes Dev,* 27**,** 163-8.

TATES, A. D. 1971. Cytodifferentiation during spermatogenesis in Drosophila melanogaster: an electron microscope study. *Ph. D. Thesis. Rijksuniversiteit, Leiden*.

TINEVEZ, J. Y., PERRY, N., SCHINDELIN, J., HOOPES, G. M., REYNOLDS, G. D., LAPLANTINE, E., BEDNAREK, S. Y., SHORTE, S. L. & ELICEIRI, K. W. 2017. TrackMate: An open and extensible platform for single-particle tracking. *Methods,* 115**,** 80-90.

TIRNAUER, J. S. & BIERER, B. E. 2000. EB1 proteins regulate microtubule dynamics, cell polarity, and chromosome stability. *J Cell Biol,* 149**,** 761-6.

TSANG, W. Y., BOSSARD, C., KHANNA, H., PERANEN, J., SWAROOP, A., MALHOTRA, V. & DYNLACHT, B. D. 2008. CP110 suppresses primary cilia formation through its interaction with CEP290, a protein deficient in human ciliary disease. *Dev Cell,* 15**,** 187-97.

TSANG, W. Y. & DYNLACHT, B. D. 2013. CP110 and its network of partners coordinately regulate cilia assembly. *Cilia,* 2**,** 9.

TSOU, M. F., WANG, W. J., GEORGE, K. A., URYU, K., STEARNS, T. & JALLEPALLI, P. V. 2009. Polo kinase and separase regulate the mitotic licensing of centriole duplication in human cells. *Dev Cell,* 17**,** 344-54.

TSUCHIYA, Y., YOSHIBA, S., GUPTA, A., WATANABE, K. & KITAGAWA, D. 2016. Cep295 is a conserved scaffold protein required for generation of a bona fide mother centriole. *Nat Commun,* 7**,** 12567.

UNIPROT-CONSORTIUM 2021. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res,* 49**,** D480-D489.

VALENTE, E. M., ROSTI, R. O., GIBBS, E. & GLEESON, J. G. 2014. Primary cilia in neurodevelopmental disorders. *Nat Rev Neurol,* 10**,** 27-36.

VAN BREUGEL, M., HIRONO, M., ANDREEVA, A., YANAGISAWA, H. A., YAMAGUCHI, S., NAKAZAWA, Y., MORGNER, N., PETROVICH, M., EBONG, I. O., ROBINSON, C. V., JOHNSON, C. M., VEPRINTSEV, D. & ZUBER, B. 2011. Structures of SAS-6 suggest its organization in centrioles. *Science,* 331**,** 1196-9.

VAN BREUGEL, M., WILCKEN, R., MCLAUGHLIN, S. H., RUTHERFORD, T. J. & JOHNSON, C. M. 2014. Structure of the SAS-6 cartwheel hub from Leishmania major. *Elife,* 3**,** e01812.

VAN DER VAART, B., AKHMANOVA, A. & STRAUBE, A. 2009. Regulation of microtubule dynamic instability. *Biochem Soc Trans,* 37**,** 1007-13.

VOROBJEV, I. A. & CHENTSOV YU, S. 1982. Centrioles in the cell cycle. I. Epithelial cells. *J Cell Biol,* 93**,** 938-49.
VULPRECHT, J., DAVID, A., TIBELIUS, A., CASTIEL, A., KONOTOP, G., LIU, F., BESTVATER, F., RAAB, M. S., ZENTGRAF, H., IZRAELI, S. & KRAMER, A. 2012. STIL is required for centriole duplication in human cells. *J Cell Sci,* 125**,** 1353-62.
WALENTEK, P., QUIGLEY, I. K., SUN, D. I., SAJJAN, U. K., KINTNER, C. & HARLAND, R. M. 2016. Ciliary transcription factors and miRNAs precisely regulate Cp110 levels required for ciliary adhesions and ciliogenesis. *Elife,* 5.
WALENTEK, P., SONG, R. & HE, L. 2014. microRNAs and cilia. An ancient connection. *Cell Cycle,* 13**,** 2315-6.
WANG, W. J., SONI, R. K., URYU, K. & TSOU, M. F. 2011. The conversion of centrioles to centrosomes: essential coupling of duplication with segregation. *J Cell Biol,* 193**,** 727-39.
WEMMER, K. A. & MARSHALL, W. F. 2007. Flagellar length control in chlamydomonas--paradigm for organelle size regulation. *Int Rev Cytol,* 260**,** 175-212.
WINEY, M. & O'TOOLE, E. 2014. Centriole structure. *Philos Trans R Soc Lond B Biol Sci,* 369.
WOODRUFF, J. B., FERREIRA GOMES, B., WIDLUND, P. O., MAHAMID, J., HONIGMANN, A. & HYMAN, A. A. 2017. The Centrosome Is a Selective Condensate that Nucleates Microtubules by Concentrating Tubulin. *Cell,* 169**,** 1066-1077 e10.
WOODRUFF, J. B., WUESEKE, O. & HYMAN, A. A. 2014. Pericentriolar material structure and dynamics. *Philos Trans R Soc Lond B Biol Sci,* 369.
WOODRUFF, J. B., WUESEKE, O., VISCARDI, V., MAHAMID, J., OCHOA, S. D., BUNKENBORG, J., WIDLUND, P. O., POZNIAKOVSKY, A., ZANIN, E., BAHMANYAR, S., ZINKE, A., HONG, S. H., DECKER, M., BAUMEISTER, W., ANDERSEN, J. S., OEGEMA, K. & HYMAN, A. A. 2015. Regulated assembly of a supramolecular centrosome scaffold in vitro. *Science,* 348**,** 808-12.
YADAV, S. P., SHARMA, N. K., LIU, C., DONG, L., LI, T. & SWAROOP, A. 2016. Centrosomal protein CP110 controls maturation of the mother centriole during cilia biogenesis. *Development,* 143**,** 1491-501.
YAMAZOE, T., NAGAI, T., UMEDA, S., SUGAYA, Y. & MIZUNO, K. 2020. Roles of TOG and jelly-roll domains of centrosomal protein CEP104 in its functions in cilium elongation and Hedgehog signaling. *J Biol Chem,* 295**,** 14723-14736.
YANG, J., ANISHCHENKO, I., PARK, H., PENG, Z., OVCHINNIKOV, S. & BAKER, D. 2020. Improved protein structure prediction using predicted interresidue orientations. *Proc Natl Acad Sci U S A,* 117**,** 1496-1503.
YANG, T. T., CHONG, W. M., WANG, W. J., MAZO, G., TANOS, B., CHEN, Z., TRAN, T. M. N., CHEN, Y. D., WENG, R. R., HUANG, C. E., JANE, W. N., TSOU, M. B. & LIAO, J. C. 2018. Super-resolution architecture of mammalian centriole distal appendages reveals distinct blade and matrix functional components. *Nat Commun,* 9**,** 2023.
ZHANG, Y., WANG, Y., WEI, Y., LI, M., YU, S., YE, M., ZHANG, H., CHEN, S., LIU, W. & ZHANG, J. 2015. MiR-129-3p promotes docetaxel resistance of breast cancer cells via CP110 inhibition. *Sci Rep,* 5**,** 15424.
ZIBRAT, M. E. 2019. Regulation of the Distal Tip Complex Controls Centriole Length.
ZITOUNI, S., FRANCIA, M. E., LEAL, F., MONTENEGRO GOUVEIA, S., NABAIS, C., DUARTE, P., GILBERTO, S., BRITO, D., MOYER, T., KANDELS-LEWIS,

S., OHTA, M., KITAGAWA, D., HOLLAND, A. J., KARSENTI, E., LORCA, T., LINCE-FARIA, M. & BETTENCOURT-DIAS, M. 2016. CDK1 Prevents Unscheduled PLK4-STIL Complex Assembly in Centriole Biogenesis. *Curr Biol,* 26**,** 1127-37.

Appendix

A1. Automated measurements of GFP-Cep97 distancing

The GFP-Cep97 distancing algorithm was developed in collaboration with Mustafa Aydogan (Raff Lab, University of Oxford) and Felix Zhou (Lu Lab, University of Oxford). All computational work was performed by F. Zhou. The script can be found at https://github.com/RaffLab/Centriole_Distancing

I present here a summary of how the GFP-Cep97 distancing algorithm works, with figures to illustrate the process. Broadly, the algorithm has three main steps:

(1) Centriole detection and cropping from the original image to produce smaller image-patches for processing (**Figure A1 A**).

(2) Convolutional neural network (CNN) where input constitutes the image-patches (**Figure A1 B**). The CNN outputs enhanced image-patches of the same size that can be used to determine the locations of mother and daughter centrioles, through local signal peak detection.

(3) Pairing of detected individual mother and daughter peak centroids from the CNN enhanced output image, and measurement of the distance between them (**Figure A1 C**).

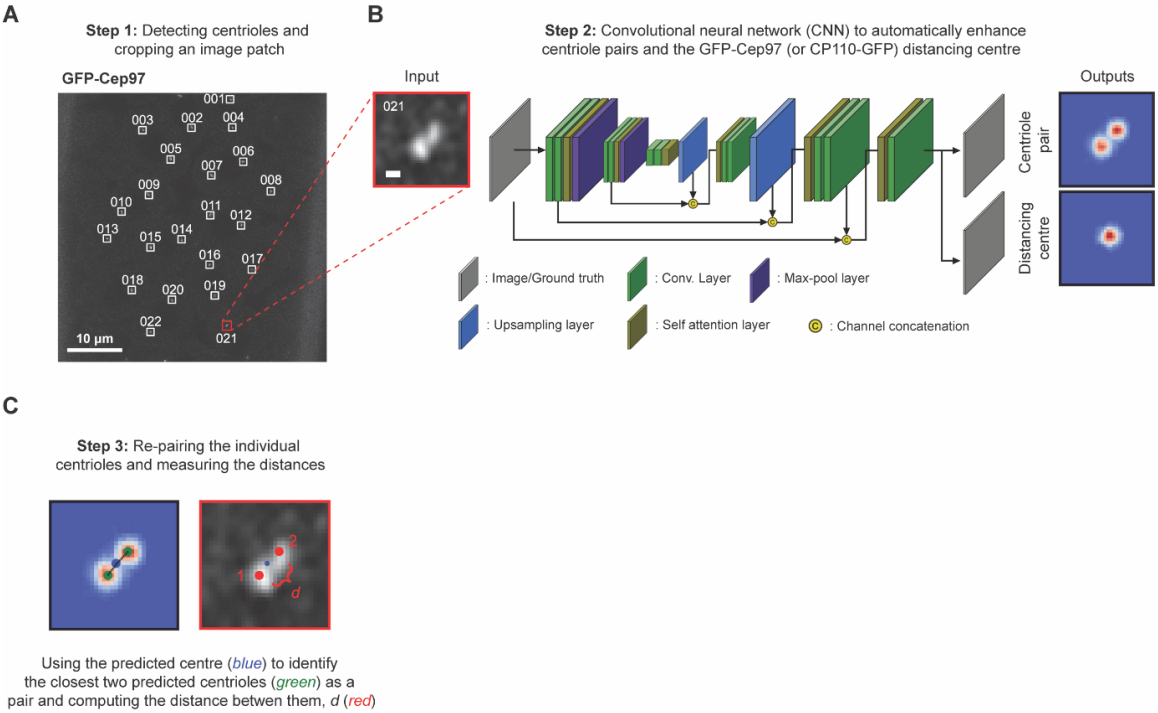


Figure A1 – Pipeline for automated measurement of GFP-Cep97 distancing in live embryos.

(A) A maximum intensity projection of an embryo expressing Ubq-GFP-Cep97 is used as the input for the centriole pair detection process. This produces image-patches (white squares with corresponding identifying numbers) predicted to contain centriole pairs. **(B)** The image-patches (scale bar represents 200 nm) produced in (A) are used as the input for the convolutional neural network (CNN). The CNN outputs are density heatmap-like images which are used to identify the locations of the centres of the mother and daughter centriole GFP-Cep97 signals (these centres are termed 'centroids') and the 'distancing centre' between centroids. A schematic shows the layers of the CNN; relative intermediate sizes illustrate successive downscaling of the image patch during processing. **(C)** Centroids identified in the CNN output (green dots) are paired so that the distance between them (d) can be measured. Figure courtesy of Felix Zhou for Hankins, Aydogan *et al.* (in preparation).

A1.1. Centriole detection

Images of embryos expressing one copy of Ubq-GFP-Cep97 (a marker for the distal-ends of the mother and daughter centriole MTs) over S-phase of cycle 12 were acquired using a Zeiss 880 Laser Scanning Microscope fitted with an Airyscan detector (zoom value =

24.41 pixels per μm). Images were acquired every 35 s as stacks of 5x 0.2 μm slices, with manual focus adjustment between acquisitions to maintain centrioles within the plane of imaging. Images were then Airy-processed in 3D with a processing strength of Auto (~6) or 6.5 using the microscope's software (ZEN black, Zeiss International). The first step in the analysis pipeline is to detect regions within these images that contain centrioles.

To do this, ImageJ was used to generate maximum intensity projections from the original micrographs (**Figure A2 A, Step 1**). Anisotropic filtering (Perona *et al.*, 1994) was then applied to these images using Trackpy (Allan *et al.*, 2018; $\delta t = 0.14$, $\kappa = 15$, *iterations* = 5) which enhances spot-like features in the image (these should correspond to centrioles) and decreases the intensity of the background (**Figure A2 A, Step 2**). Next, trackpy.locate (Allan *et al.*, 2018) was used (with *separation* = 5 pixels, *minmass* = 10) to detect Gaussian blob-like objects with widths of 25 pixels (1.04 μm) within the image (**Figure A2 A, Step 3**). These represent the initial centriole detections, which were cropped to generate an image-patch of 32x32 pixels (these will form the input images for the CNN, as illustrated in **Figure A1**). Choice of which width to apply to this detection step was determined by the microscope and imaging settings used; since these are kept constant in all our images, we keep the 25-pixel size as a constant parameter.

These initial detections were quality-controlled via a set of filtering steps:

- <u>Removal of edge detections</u>

  Any detections within 12.5 pixels (i.e. half a centriole width) of the edge of the image were removed (**Figure A2 A, Step 4**).

- Removal of detections in auto-fluorescent image regions

  Some detections were found to be in regions of auto-fluorescence, such as at the cortex of the embryo. Detections here are therefore unlikely to represent true centrioles. To remove these, regions in the overall image with high fluorescence intensity were detected (defined as regions of >1000 adjoining pixels (41.5 μm) with fluorescence intensity >20% of the image's maximum pixel intensity) and excluded (shown in yellow, **Figure A2 A, Step 5**).

- Removal of duplicate detections

  Sometimes, duplicate detections occurred. To remove these, detections within 10 pixels (0.42 μm) of each other were identified, and the second detection instance was discarded (**Figure A2 A, Step 6**).

- Removal of low fluorescence intensity detections

  The coefficient of variation (CV) was calculated by dividing *standard deviation of pixel intensities* by *mean of pixel intensities* within each detected 32x32 pixel image-patch. The value of the CV gives an indication of the signal:noise ratio of an image patch. Detections below a determined threshold were discarded (**Figure A2 A, Step 7**). Unlike the centriole width parameter described above, this CV threshold had to be set for each dataset. To determine this threshold, the CV value was varied and the resulting automated detections were compared to manual detections. Automated and manual detections were determined to 'match' if they were within 16 pixels of one another. This then allowed the *Precision*, *Recall* and *F score* performance at these various thresholds to be assessed:

Precision demonstrates how many relevant detections have been made via the automated approach:

$$Precision = \frac{\#(Matched\ automatic\ and\ manual\ detections)}{\#(Automatic\ detections)}$$

Recall demonstrates the proportion of manual detections that have been successfully identified via the automated approach:

$$Recall = \frac{\#(Matched\ automatic\ and\ manual\ detections)}{\#(Manual\ detections)}$$

F score takes the harmonic mean of *Precision* and *Recall* to assess how well the algorithm scores in both of these qualities:

$$\text{F score} = 2\frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}}$$

The peak *F score* was found when CV = 0.30 (**Figure A2 C**). Since the distance between centrioles increases over the course of S-phase, and since the degree of asymmetry between the GFP-Cep97 intensities on the mother and daughter centriole varies across this period, automated detection performance was also tested at early-, mid-, and late-S-phase. There was a slight propensity for over-prediction (i.e. more automated detections were made than manual ones, thus giving a lower *Precision* score) in all three S-phase periods, but overall *Precision* (0.86) and *Recall* (0.91) were high (**Figure A2 D**). There is also a trade-off

between *Precision* and *Recall* (**Figure A2 C**), though Average Precision (AP) for all possible CV cut-offs was high at 0.89. It therefore appeared that the script performed well at most CV thresholds; in practice, however, a range of CV thresholds (0.30-0.45 ± 0.05) were employed to optimise performance as far as possible.

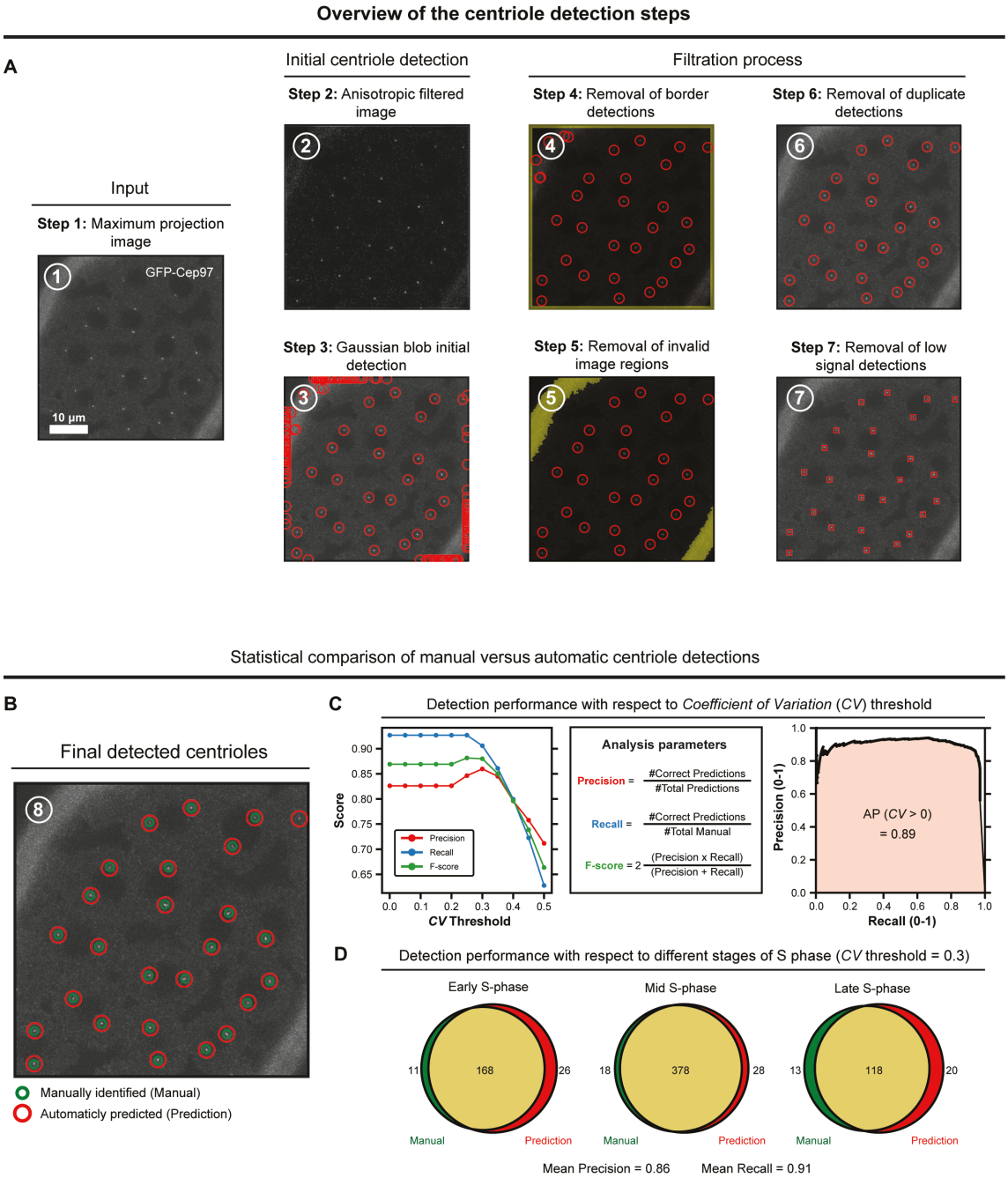

Overview of the centriole detection steps
A
Input
Step 1: Maximum projection image
GFP-Cep97
10 µm
Initial centriole detection
Step 2: Anisotropic filtered image
Step 3: Gaussian blob initial detection
Filtration process
Step 4: Removal of border detections
Step 5: Removal of invalid image regions
Step 6: Removal of duplicate detections
Step 7: Removal of low signal detections
Statistical comparison of manual versus automatic centriole detections
B
Final detected centrioles
Manually identified (Manual)
Automaticly predicted (Prediction)
C
Detection performance with respect to Coefficient of Variation (CV) threshold
Score
CV Threshold
Precision
Recall
F-score
Analysis parameters
Precision = #Correct Predictions / #Total Predictions
Recall = #Correct Predictions / #Total Manual
F-score = 2 (Precision x Recall) / (Precision + Recall)
Precision (0-1)
Recall (0-1)
AP (CV > 0) = 0.89
D
Detection performance with respect to different stages of S phase (CV threshold = 0.3)
Early S-phase
Mid S-phase
Late S-phase
11
168
26
18
378
28
13
118
20
Manual
Prediction
Mean Precision = 0.86
Mean Recall = 0.91

Figure A2 – Centriole detection.

(A) Illustration of the centriole detection process. *Step 1*: A maximum intensity projection of an embryo expressing Ubq-GFP-Cep97 is used as the input. *Step 2*: Anisotropic filtering is applied to the input image, enhancing spot-like features and decreasing the intensity of the background. *Step 3*: Gaussian blob-like objects with widths of 25 pixels are detected within the image. *Step 4*: Any detections within 12.5 pixels of the image edge are discarded. *Step 5*: Detections in areas of high auto-fluorescence are discarded. *Step 6*: Detections within 10 pixels of each other (duplicates) are identified, and the second detection instance is discarded. *Step 7*: Detections above the coefficient of variation (CV) threshold are kept. Final 32x32 pixel image-patches are indicated with red boxes. **(B)** Micrograph showing an embryo expressing Ubq-GFP-Cep97 where manually detected (green) and automatically detected (red) centrioles are circled. **(C)** Graph (*left*) showing the *Precision* (red), *Recall* (blue) and *F-score* (blue) attained by the centriole detection pathway at varying CV thresholds. Graph (*right*) showing a precision-recall curve composed of individual centriole predictions (small black dots); centrioles are ranked left-to-right according to CV, with those of the highest CV values on the right. The total area under the curve gives the Average Precision (AP). Total non-thresholded (CV>0) detections in plot = 1240. Total manual annotations = 716. **(D)** Venn diagrams showing how well automatic (red) and manual (green) detections overlap (yellow) at different S-phase stages when $CV \geq 0.3$. Numbers on the Venn diagrams indicate the number of centriole detections in each case. Figure courtesy of Felix Zhou and Mustafa Aydogan for Hankins, Aydogan *et al.* (in preparation).

A1.2. CNN to automatically predict centriole pairs

CNN architecture is illustrated in **Figure A1 B**. The CNN used in this analysis had the following architectural parameters:

Layer Name	Layer	Width	Padding	Nonlinearity	Output Size
Input	Input	-	-	-	1x32x32
Conv-1	Convolution (3x3)	32	'Same'	Scaled exponential linear units (SELU) (Klambauer *et al.*, 2017)	32x32x32
Conv-2	Convolution (3x3)	32	'Same'	SELU	32x32x32

Attn-1	Self Attention (latent dim=8) (Vaswani *et al.*, 2017, Zhang *et al.*, 2018)	-	-	-	32x32x32
Pool-1	Max-pooling (2x2)	-	-	-	32x16x16
Conv-3	Convolution (3x3)	64	'Same'	SELU	64x16x16
Conv-4	Convolution (3x3)	64	'Same'	SELU	64x16x16
Attn-2	Self Attention (latent dim = 16)	-	-	-	64x16x16
Pool-2	Max-pooling (2x2	-	-	-	64x8x8
Conv-5	Convolution (3x3)	128	'Same'	SELU	128x8x8
Conv-6	Convolution (3x3)	128	'Same'	SELU	128x8x8
Attn-3	Self Attention (latent dim = 32)	-	-	-	128x8x8
Upsample-1	Upsampling (2x2)	-	-	-	64x16x16
Concat-1	Concatenate (Conv-4, Upsample-1)	-	-	-	192x16x16
Attn-4	Self Attention (latent dim = 24)				192x16x16
Conv-7	Convolution (3x3)	64	'Same'	SELU	64x16x16
Conv-8	Convolution (3x3)	64	'Same'	SELU	64x16x16
Upsample-2	Upsampling (2x2)	-	-	-	64x32x32
Concat-2	Concatenate (Conv-2, Upsample-2)	-	-	-	96x32x32
Attn-5	Self Attention (latent dim = 8)	-	-	-	96x32x32

Conv-9	Convolution (3x3)	32	'Same'	SELU	32x32x32
Conv-10	Convolution (3x3)	32	'Same'	SELU	32x32x32
Concat-3	Concatenate (Input, Conv-10)	33	-	-	33x32x32
Attn-6	Self Attention (latent dim = 8)	-	-	-	33x32x32
Conv-11	Convolution (1x1)	2	'Same'	Rectified linear units (ReLU)	2x32x32

With the CNN architecture established, a training set of images needed to be generated. To do this, four embryos expressing Ubq-GFP-Cep97 in S-phase of cycle 12 were imaged at 35 s time intervals (this produced a total of 32 time points: early-S-phase = 10, mid-S-phase = 15, late-S-phase = 7). These images were processed for centriole detection, and a total of 616 centriole-containing regions (early-S-phase = 161, mid-S-phase = 340, late-S-phase = 115) were identified and cropped into 32x32 pixel image-patches as described in **A1.1**. The best-focused slice for each image-patch was then selected (this was defined as the slice where CV was highest). The mother and daughter centriole in each image-patch were then hand-annotated using red dots (**Figure A3 A, Step 1**). The central point ('distancing centre') between the two centrioles was also annotated using a blue dot. These dots were then replaced with Gaussians of width 1 (**Figure A3 A, Step 2**).

The same CNN was trained on the hand-annotations produced from either early-, mid-, or late S-phase. *Adam* (Kingma, 2014) was used for the training (learning rate λ=0.001, β_1 = 0.9, β_2 = 0.999). The regression loss function was mean-squared error (MSE) loss, and the magnitude of loss was enhanced by multiplying transformed annotations by 1000. The data was split 80%-20% into training and testing data, respectively. To increase the size of

the training set to ~10,000 image-patches for each stage of S-phase, the original training data image-patches were randomly adjusted in terms of their geometric, contrast and noise properties (**Figure A3 B**). Testing data was not adjusted, and was used to produce a validation MSE loss; this helped reduce overfitting to the training data. When loss of MSE did not continue decreasing for >15 epochs, training was stopped. CNN training was typically completed within 5-10 epochs (Nvidia GeForce GTX Titan Black 6GB graphics card).

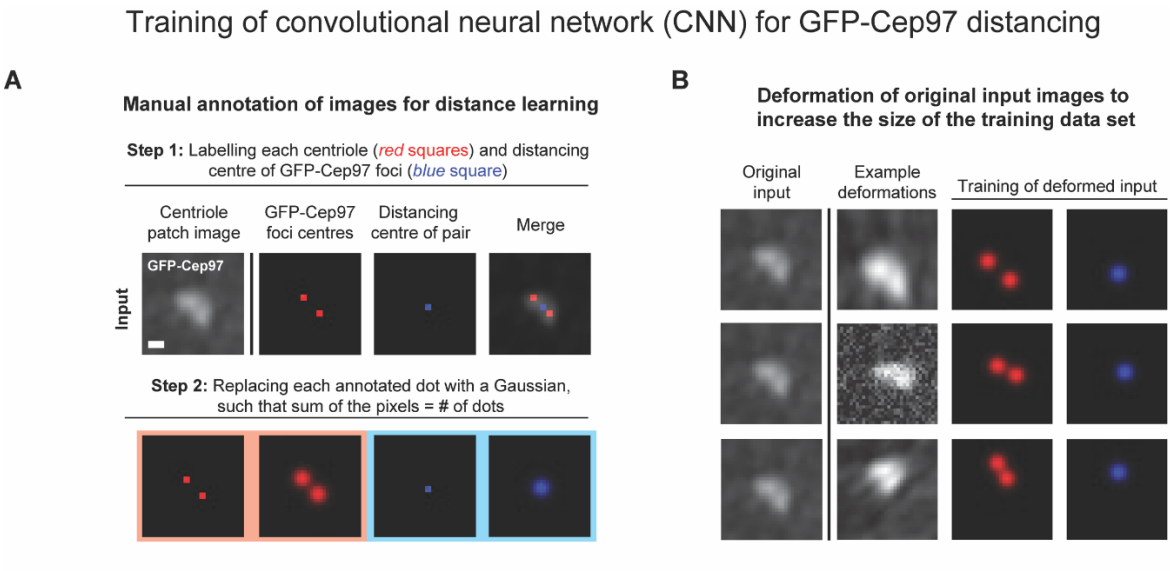


Figure A3 – Convolutional neural network (CNN) training.

(A) Illustration of the hand-annotation process used to produce the training dataset. *Step 1*: Image-patches of detected centrioles (generated as described in **A1.1**) are hand-annotated by the experimenter to mark, by eye, the centres of the mother and daughter centriole GFP-Cep97 signals (red dot) and the 'distancing centre' between them (blue dot). *Step 2*: The annotations are replaced with Gaussians (where sum of image pixel intensities = number of dot annotations). **(B)** To increase the size of the training set to ~10,000 images for each stage of S-phase, the training data were randomly adjusted in terms of geometric, contrast and noise properties. These deformations are shown in the 'Example deformations' column. If geometric deformations were applied, the matching annotation images were also deformed in the same way (examples are shown in the 'Training of deformed input' column). Figure courtesy of Felix Zhou and Mustafa Aydogan for Hankins, Aydogan *et al.* (in preparation).

A1.3. Pairing of individual mother and daughter centrioles and measuring the distance between the centres of the pairs

With training complete, the CNN could then be applied to our images of GFP-Cep97 distancing in different genotypes. Image-patches were produced as described in **A1.1** and the best-focussed slice of each image-patch was used as input. The image intensity $I(i,j)$ of input image-patch slices was normalised using min-max normalisation. This scaled pixel intensities to between 0 and 1. The resulting images therefore now had normalised image intensities ($I_{norm}(i,j)$) as described in the following equation:

$$I_{norm}(i,j) = \frac{I(i,j) - \min I(i,j)}{\max I(i,j) - \min I(i,j)}$$

CNN outputs were divided by 1000 (since this was the factor used during training to multiply the transformed annotations). These output images contained local maxima peaks which corresponded to predicted centres of the GFP-Cep97 signals (these centres are termed 'centroids') and to the 'distancing centre' between centroids. These peaks were located using the scikit-image's *skimage.feature.peak_local_max* function (Van der Walt *et al.*, 2014), where *minimum separation distance* = 1 pixel. Additionally, a threshold (determined by Otsu's method) was applied and centroids with intensity below this threshold were discarded.

Predicted centroids were then considered for pairing if >2 centroids were detected in the image-patch. The two centroids closest to the predicted distancing centre were selected (**Figure A4 A**) and the Euclidean distance between them ($d_{1\ and\ 2}$), the intensities of the

two centroids (I_1, I_2), and the pixel intensity ratio of the two centroids ($I_{\text{ratio}} = {}^{I_1}/_{I_2}$, where $I_1 > I_2$) were calculated. These values were used to determine if centroids were considered to be a pair: if the distance was <15 pixels and if the ratio between their intensities was ≥1/4, the centroids were successfully paired (i.e. these two centroids represent a mother and daughter centriole, and the distance between their respective GFP-Cep97 signal centres at this time point could now be calculated, as illustrated in **Figure A4 C**).

If the detected distance was >15 pixels or the ratio between centroid intensities was <1/4, the brighter of the two centroids was identified as one centriole. There were also cases where image-patches contained only one peak centroid could be detected from the CNN output. An example is shown in **Figure A4 B**. It is likely that there is actually a pair of centrioles present in these image-patches, but the CNN was unable to deconvolve the intensities of both centrioles (this could be due to one centriole having a low signal intensity/poor orientation, or it could be that, especially in early-S-phase, the centrioles are very close together). A Gaussian mixture model (GMM) was therefore used to resolve the single centriole as two centroids. This is because CNN output image intensity distributions are approximately 2D Gaussian ($\mathcal{N}(\mu, \Sigma)$, where μ = mean position vector, and Σ = covariance matrix). A single GFP-Cep97 signal can therefore be resolved as two overlapping signals by modelling the observed image intensity $I(i,j)$ as the 2D probability density of the pixel coordinates (i,j) from the weighted summation of two 2D Gaussian distributions:

$$I(i,j) = w_1 \underbrace{\mathcal{N}(\mu_1, \Sigma_1)}_{\text{Gaussian 1}} + w_2 \underbrace{\mathcal{N}(\mu_2, \Sigma_2)}_{\text{Gaussian 2}}$$

In the above GMM model equation, $I(i,j)$ = image pixel intensity in image-patch row i, column j. w_1 = proportion of Gaussian 1. w_2 = proportion of Gaussian 2.

To determine the parameters of the two 2D Gaussians described in the above equation, CNN output images were processed as follows:

- Min-max normalisation (to normalise image intensity to between 0 and 1, as described previously) was applied:

$$I_{norm}(i,j) = \frac{I(i,j) - \min I(i,j)}{\max I(i,j) - \min I(i,j)}$$

- An exclusion mask was generated to reduce the chance of sampling low intensity regions, thus improving the robustness of the fit. This mask was created using mean thresholding (i.e. this would discard any regions where pixel intensity is lower than or equal to mean pixel intensity) as follows:

$$M(i,j) = \begin{cases} 0, & I(i,j) \leq \text{mean}(I(i,j)) \\ 1, & I(i,j) > \text{mean}(I(i,j)) \end{cases}$$

  then, where $M(i,j) = 0$, all normalised intensities at (i,j) were set to 0:

$$I_{norm}(i,j) = \begin{cases} 0, & M(i,j) = 0 \\ I_{norm}(i,j), & M(i,j) = 1 \end{cases}$$

- The (i,j) pairs were sampled proportional to image intensity ($I(i,j)$). This was achieved by taking (i,j), $\lfloor n_{samples} \text{ x } I_{norm}(i,j) \rfloor$ a number of times.

  $n_{samples}$ = maximum number of times sampling of (i,j) for intensity $I(i,j)$ was

performed. $\lfloor \cdot \rfloor$ = the mathematical floor operator.

- All samples were then concatenated to form a long vector (S).

- S was fitted to the two-component GMM model described above using expectation-maximisation (Dempster *et al.*, 1977, Murphy, 2012) to extract the equation parameters ($w_1, w_2, \mu_1, \mu_2, \Sigma_1$ and Σ_2).

When fitted, the locations of two new centroids are now given by μ_1 and μ_2 (**Figure A4 B**, black dots). The covariances (Σ_1 and Σ_2) are shown as cyan and orange ellipses in **Figure A4 B**. Thus, two GFP-Cep97 centroids are now identified in the image and the distance between them can be calculated (as in **Figure A4 C**).

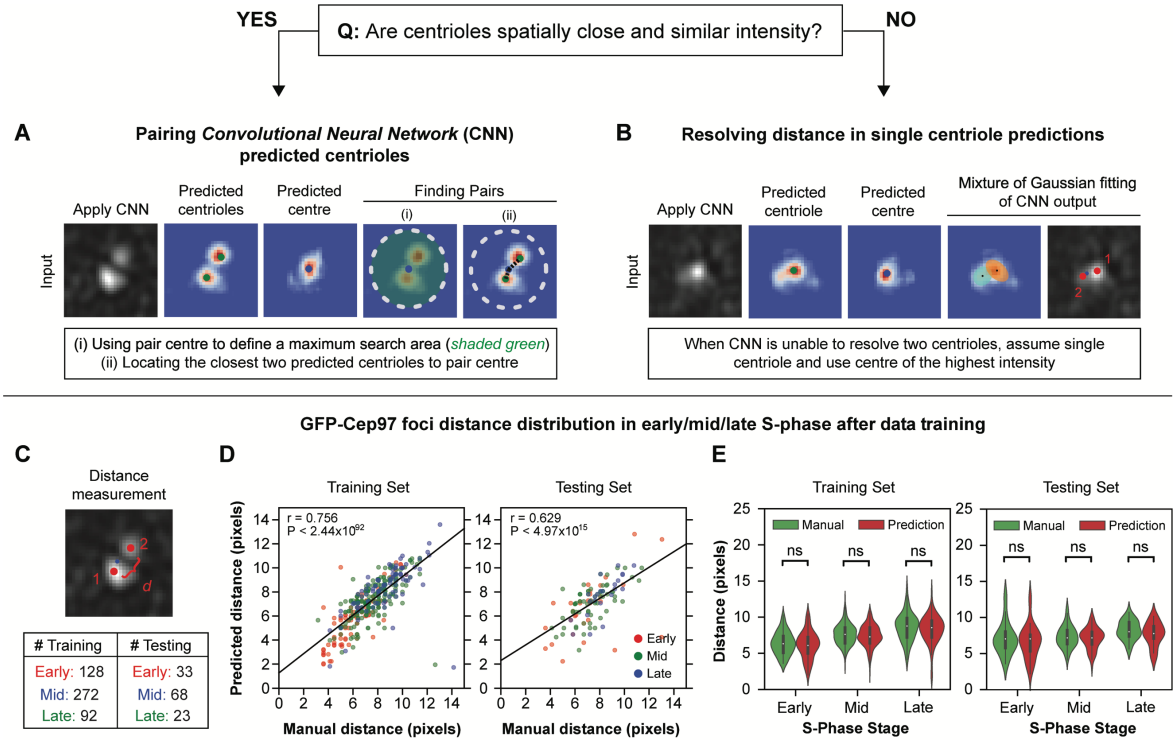


Figure A4 – Pairing of mother and daughter centrioles and assessment of algorithm performance.

(A) Panels showing the input image-patch (greyscale micrograph) and CNN output (blue backgrounds; green dots show centroid positions, blue dot shows 'distancing centre' between them). The two centroids closest to the distancing centre (search area around the distancing centre is indicated by a white dashed circle) were identified. These two centroids were within 15 pixels of each other and the ratio between their intensities was ≥ 1/4, so they were paired. Successful pairing is indicated by the black dashed line. **(B)** Panels showing an image-patch (greyscale micrograph) where the CNN output (blue backgrounds; green dot shows centroid position, blue dot shows 'distancing centre') gives just one predicted centriole. In this case, a Gaussian mixture model (GMM) was applied to the CNN output image in order to resolve the single centriole as two centroids. The positions of the two new centroids are given by μ_1 and μ_2 (black dots). Their covariances (Σ_1 and Σ_2) are shown as cyan and orange elipses. Once fitted, the mean centre the two Gaussians becomes the new distancing centre between the centroids. Thus, two GFP-Cep97 centroids are now identified in the image. **(C)** Micrograph showing a pair of mother and daughter centrioles where the Euclidean distance between their centres (d) is indicated. The table below indicates the number of hand-annotated centriole pairs (excluding those produced via deformation as in **Figure A3 B**) used in the training and testing sets for the CNN. **(D)** Scatter plots for both the training and testing sets are shown, to demonstrate the positive correlation between the distances measured from the manual annotations and from the GFP-Cep97 distancing algorithm predictions. *r* values (stating Pearson's correlation) and *p*-values are shown on the graphs, and linear regressions have been fitted. Measurements are shown for early- (red dots) mid- (green dots) and late-S-phase (blue dots). **(E)** Violin plots showing comparisons of distances measured from the manual annotations (green) and the GFP-distancing algorithm predictions (red) for early-, mid-, and late-S-phase. Data are presented as median (white dot) ± interquartile range (IQR, thick black lines). The whiskers indicate ± 1.5x IQR either side of the lower and upper quartiles. Wilcoxon signed-rank tests were performed to assess if the differences between datasets are statistically significant; the resulting *p*-values indicated that these differences were not significant (ns). Figure courtesy of Felix Zhou and Mustafa Aydogan for Hankins, Aydogan *et al.* (in preparation).

A1.4. Validating the GFP-Cep97 distancing output

To assess whether there was any overfitting by the GFP-Cep97 distancing algorithm, scatter plots were produced to compare the distances measured from the manual annotations with those measured from the GFP-distancing algorithm predictions on the training set (**Figure A4 D**). This demonstrated that these two sets of distances were similar, suggesting no significant overfitting. To assess the overall performance of the GFP-Cep97 distancing algorithm in measuring distances between mother and daughter centriole signals, scatter plots were produced to compare the distances measured from manual annotations with those measured from the GFP-distancing algorithm predictions on the testing set (**Figure A4 D**). This again demonstrated a positive correlation between these two sets of distancing values. T-tests revealed that there was also no significant difference between the distribution of distances calculated from either manual annotations or from predictions; this data is presented as violin plots in **Figure A4 E**.

A1.5. Fitting the resulting distancing data

Extensive trials fitting different regression models against the raw distancing data were performed, and, as explained in **Chapter 3**, we reasoned that we could fit this data using the same approach previously described for modelling cartwheel growth (Aydogan *et al.*, 2018). This was formally confirmed by fitting the data against the following regression models using GraphPad Prism 8 to compare fit:

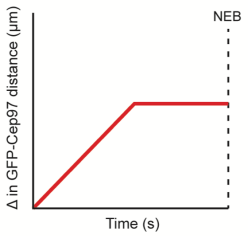
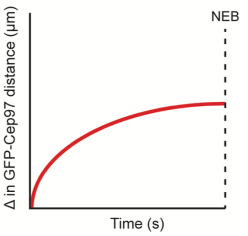
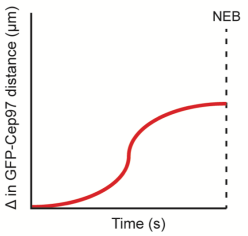
(1) Linear / Linear increase followed by a plateau

(2) One-phase association

(3) Sigmoidal

The models selected above represent those which had given the most promising fits during the initial process of trialling different models. All functions above are described in GraphPad Prism except for *Linear increase followed by a plateau*; this is an in-house function described in Aydogan *et al.* 2018. Fitting was performed on data from individual embryos by considering the mean y value of each time point. R^2, absolute sum-of square values and Runs tests were used to assess goodness-of-fit. Comparisons between fits were performed using the sum-of-squares *F* test; the simpler model was selected unless $p<0.05$. The models in *(1)* fitted the data best (**Figure A5 A**), so all distancing data presented in this thesis were regressed with these models. An example of this fitting is shown in **Figure A5 B**.

A

N = 16 embryos	Linear / Linear followed by a plateau	One phase association (Parabolic)	Sigmoidal
			
Deg. of freedom	7-10	6-11	6-11
R^2 value	0.86 ± 0.07	0.87 ± 0.06	0.56 ± 0.36
% of fit preference	75%	20%	5%

B

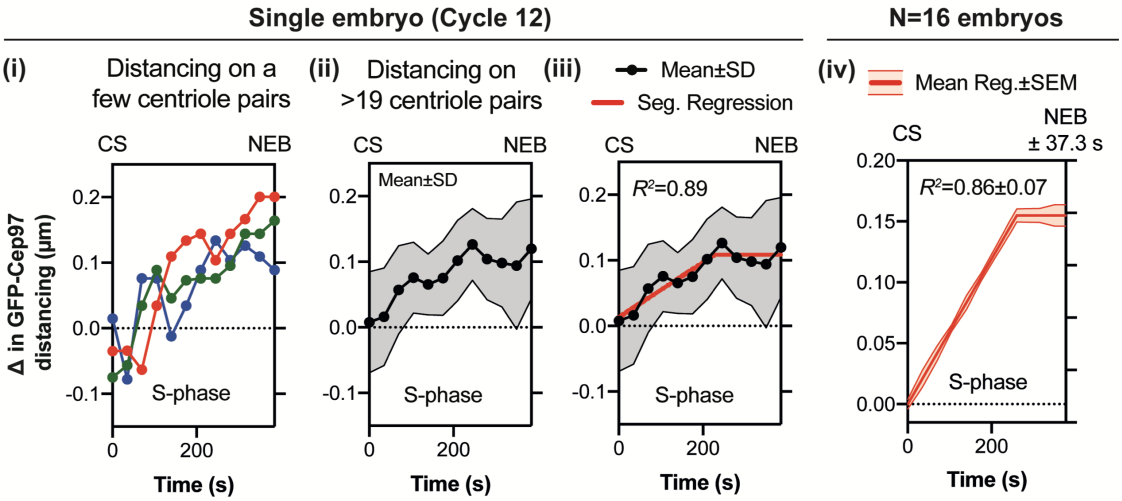


Figure A5 – Centriole MT growth can best be modelled with linear growth.

(A) Distancing data were fitted against three separate regression models (illustrated in red): *(1) Linear / Linear increase followed by a plateau (2) One-phase association (3) Sigmoidal.* The majority of the graphs preferentially fitted *(1)*, so this model was selected to fit all distancing data. R^2 values were also assessed, since these give an indication of the goodness-of-fit. **(B)** Graphs showing the change in distance between mother and daughter GFP-Cep97 signals plotted as a function of time. This is shown for a few individual centriole pairs **(i)**, and the mean change in distance (± SD) for >19 centriole pairs **(ii)** in WT embryos. The *Linear increase followed by a plateau* function **(iii)** was fitted to the data shown in (ii). **(iv)** An average curve showing GFP-Cep97 distancing during cycle 12 in 16 WT embryos, where data are presented as mean ± SEM. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). R^2 values are shown to indicate the goodness-of-fit. The data presented in this figure were acquired and analysed as part of a collaboration with Felix Zhou and Mustafa Aydogan; final figure by M. Aydogan for Hankins, Aydogan *et al.* (in preparation).

A1.6. Extracting the GFP-Cep97 oscillation data from images used for distancing measurements

In **Chapter 3**, we showed that there was no significant correlation between the amplitude of the GFP-Cep97 oscillation and the growth rate of the centriole MTs. In order to do this, it was necessary to extract the fluorescence intensity of the GFP-Cep97 signal from the images used to measure the increasing distance between these signals. This would allow us to compare the amplitude of the GFP-Cep97 oscillation with the growth rate of the centriole MTs within the same embryo. Fluorescence intensity was measured by calculating the signal:noise ratio (SNR) of the image-patches used in the GFP-Cep97 distancing algorithm. The SNR was calculated by dividing the *mean pixel intensity* of the image-patch in question by the *mean intensity, excluding centrioles*, of the original image from which the image-patch was cropped (**Figure A6 A**).

As described in **Chapter 3**, we have found that GFP-Cep97 oscillation data can be fitted using the Lorentzian function; GraphPad Prism 8 was therefore used to fit this non-linear regression to the oscillation in each individual embryo (**Figure A6 B**). The values of the peaks in each case were then extracted from the fitting parameters.

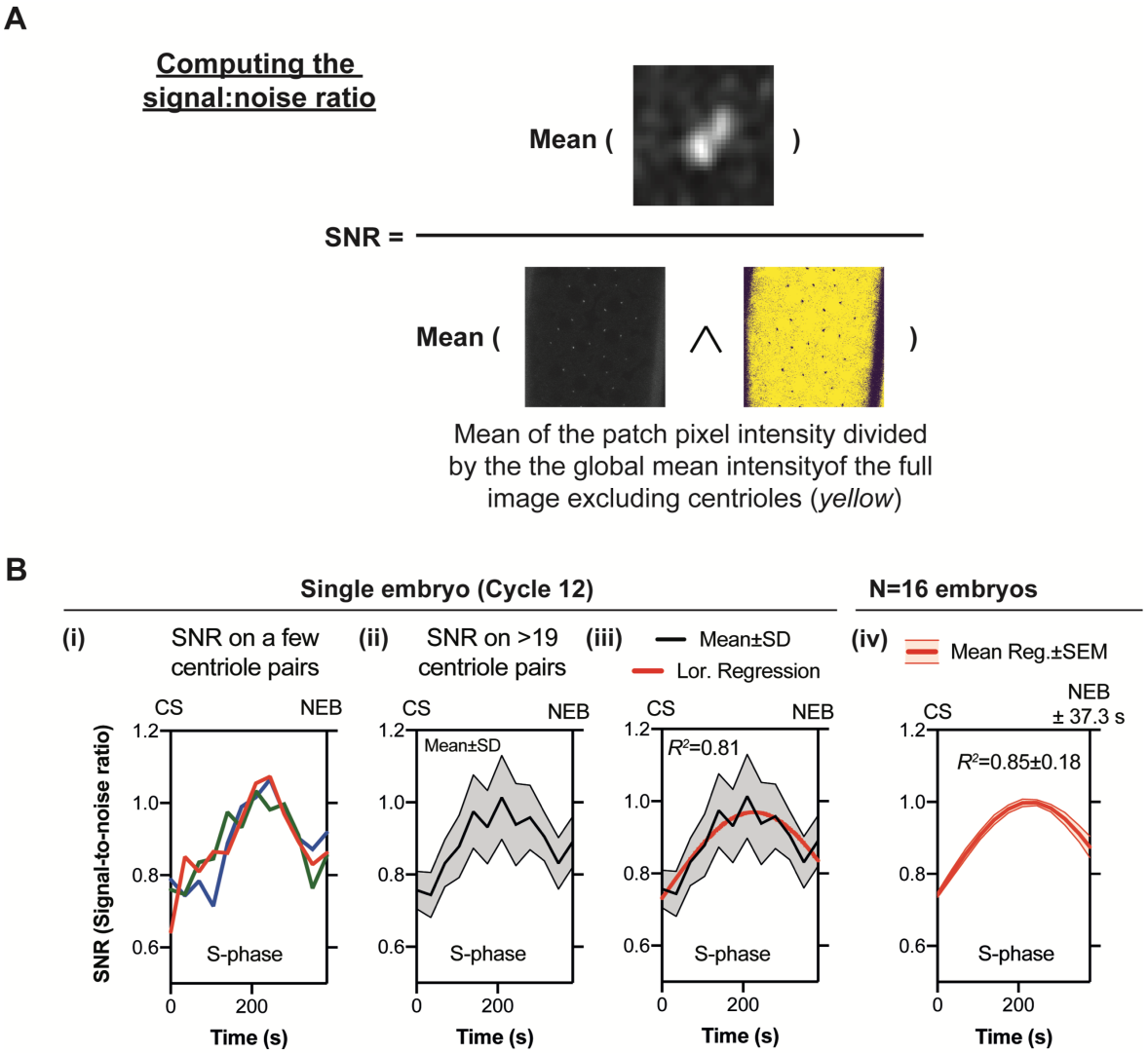


Figure A6 – Calculating the signal:noise ratio (SNR) of the GFP-Cep97 signal.

(A) Illustration showing how the signal:noise ratio (SNR) of each image-patch was calculated. As shown, the *mean pixel intensity* of the image-patch in question was divided by the *mean intensity, excluding centrioles*, of the original image from which the image-patch was cropped. The region used to calculate the latter value is indicated in yellow. **(B)** Graphs showing SNR plotted as a function of time. This is shown for a few individual centriole pairs **(i)**, and the mean SNR (± SD) for >19 centriole pairs **(ii)** in WT embryos. A Lorentzian function **(iii)** was fitted to the data shown in (ii). **(iv)** An average curve showing the GFP-Cep97 SNR oscillation during cycle 12 in 16 WT embryos, where data are presented as mean ± SEM. CS = centrosome separation (marking the start of S-phase), NEB = nuclear envelope breakdown (marking the start of mitosis). R^2 values are shown to indicate the goodness-of-fit. Figure courtesy of Felix Zhou and Mustafa Aydogan for Hankins, Aydogan *et al.* (in preparation).

Supplemental References

ALLAN, D. B., CASWELL, T., KEIM, N. C. & VAN DER WEL, C. M. 2018. trackpy: Trackpy v0.4.1 (Version v0.4.1). *Zenodo, April 21* http://doi.org/10.5281/zenodo.1226458.

AYDOGAN, M. G., WAINMAN, A., SAURYA, S., STEINACKER, T. L., CABALLE, A., NOVAK, Z. A., BAUMBACH, J., MUSCHALIK, N. & RAFF, J. W. 2018. A homeostatic clock sets daughter centriole size in flies. *J Cell Biol,* 217**,** 1233-1248.

DEMPSTER, A. P., LAIRD, N. M. & RUBIN, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological),* 39**,** 1-22.

KINGMA, D. P. B., J. 2014. Adam: A Method for Stochastic Optimization. *(available at* https://arxiv.org/abs/1412.6980*)*.

KLAMBAUER, G., UNTERTHINER, T., MAYR, A. & HOCHREITER, S. 2017. Self-normalizing neural networks. *arXiv preprint arXiv:1706.02515*.

MURPHY, K. P. 2012. *Machine learning: a probabilistic perspective*, MIT press.

PERONA, P., SHIOTA, T. & MALIK, J. 1994. Anisotropic diffusion. *Geometry-driven diffusion in computer vision.* Springer.

VAN DER WALT, S., SCHÖNBERGER, J. L., NUNEZ-IGLESIAS, J., BOULOGNE, F., WARNER, J. D., YAGER, N., GOUILLART, E. & YU, T. 2014. scikit-image: image processing in Python. *PeerJ,* 2**,** e453.

VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, L. & POLOSUKHIN, I. 2017. Attention Is All You Need. *(available at* https://arxiv.org/abs/1706.03762*)*.

ZHANG, H., GOODFELLOW, I., METAXAS, D. & ODENA, A. 2018. Self-Attention Generative Adversarial Networks. *(available at* https://arxiv.org/abs/1805.08318*)*.