SENIOR EXAMINER'S REPORT 2021
NATURAL SCIENCES TRIPOS, PART II GENETICS AND BBS GENETICS

Examiners: Dr Marisa Segal (Senior), Professor Richard Durbin (Internal), Sara Goodacre (External, University of Nottingham)

1. STRUCTURE OF THE COURSE AND EXAMINATION

i) Part II Genetics

A new course structure has been introduced this year, with concomitant changes to the weight of individual exam components as follows:

Literature Review (9% of final mark) - completed during the Christmas vacation. *Research project performance and report* (17% of final mark) - undertaken during the Lent term. *End-of-year open book online written papers* of 24 hrs (74% of final mark) – a) Integrated paper (10% of final mark) and b) Papers 1 through 4 (each 16% of final mark) corresponding to one of four modules, respectively.

The External Examiner gave each of 27 students a 15-minute *viva* by Zoom. *Viva* performance is not an assessment that contributes to marks, but is taken into account when considering borderline candidates.

ii) BBS

Candidates taking Genetics as their major subject (BBS course code 414): *Four written papers*, each linked to one of four modules (64% of final mark), sat alongside Part II Genetics candidates. *Dissertation* (20% of final mark), all but one conducted in their major subject. *Minor subject* (16% of final mark), three students took Bioinformatics, one took Applied clinical research and the other Plant signalling networks in growth and development. BBS students are not invited for *vivas.*

Papers taken by BBS candidates as minor subjects in Genetics:
Paper 2 (BBS course code 83): 2 students
Paper 4 (BBS course code 84): 3 students

iii) Module 4 (Evolutionary Genetics & Adaptation) –*a shared paper*. Module 4 is also offered as a Lent term option (aka ZL5) to those taking NST2 Zoology and NST2 Plant Sciences, and as a major subject option to students taking Part II BBS Zoology. The conduct of examination was identical to that observed for Genetics candidates, but the scripts from these students were processed by the Department of Zoology.

2. NUMBER OF CANDIDATES SITTING GENETICS PAPERS:

27 NST2GN candidates (not including one who withdrew before the examinations)
5 NST2BBS major (Genetics)
5 other NST2BBS students taking a Genetics minor subject
27 NST2 Zoology (19), NST2Plants (1) NST2BBS with Zoology as major subject (7)

Table 1: Numbers of candidates sitting Genetics Papers

Paper	NST2GN	BBS Major Genetics	BBS Minor in Genetics	Others	Total registered*	Sitting elsewhere	Total in Part II Room
1	27	5	0	0	32	32	0
2	27	5	2	0	34	34	0
3	27	5	0	0	32	32	0
4	27	5	3	27*	62	62	0
IP	27	0	0	0	27	27	0

* excluding one candidate who withdrew before the exams began
* scripts processed by the Department of Zoology

3. SETTING THE WRITTEN PAPERS

A Form & Conduct notice and a set of sample papers reflecting the changes to the examination structure were submitted in Michaelmas Term (19 October 2020). The tentative rubric retained the overall format from 2019-20 (online, open book, word limits, file naming, upload and cover sheet instructions). One significant change introduced as a result of the new course structure was that candidates were asked to answer one question *out of seven* in 2 hours for the Integrated Paper and three questions *out of seven* in 3 hours for Papers 1 through 4.

The information was shared with the class on a brief and also posted on Moodle. Two further updates were communicated to the class (29 October 2020 and 5 March 2021) regarding the finalised framework for assessment —relative to the initial F&C notice, the only significant change was that assessment would be conducted over 24 hours.

The call for suggestions was circulated to all lecturers in *early January*. Question setters were also required to supply outline answers. Additionally, Module organisers were asked to assist in proposing integrative and data handling questions. The Senior Examiner arranged for problem questions to be double-checked by colleagues with the relevant academic background. The Genetics papers were drafted by the examiners in consultation with each Module Organiser. Revisions were requested from some lecturers. Those approached responded quickly and effectively. Drafts of the Genetics papers were shared with the External Examiner in late March, who approved the questions subject to minor changes. Rubrics based on the sample exam set were revised following confirmation that assessment would be conducted online over a 24 h period. All paper rubrics listed the NST2GN subject code and the NST2BBS code. In addition, Paper 4 also carried codes for NST2ZO and NST2PL (but only the Genetics paper code i.e. "Paper 4 Evolutionary Genetics & Adaptation"). The finalised Paper4/ZL5 was shared with the Part II Zoology Senior Examiner.

The Senior Examiner and the Exam Administrator were nominated as Departmental contacts for submission of examination papers via *Teams*.

4. CONDUCT OF ASSESSMENTS

i) Coursework:

NST2GN Literature Review and Project report. The deadline for the Literature Review was 20 January 2021 and for the Project Report 29 March 2021. NST2GN candidates were asked: (i) to submit the title and scope of their Literature Review, as agreed in a meeting with the PI hosting their research project, by the end of Michaelmas Full term; and (ii) to arrange a feedback meeting with the PI, to discuss their Project Report, by the end of the Lent Full Term. ***BBS Dissertation.*** Deadlines were as set by the Faculty of Biology: Title/proposal approved - 9 November 2020; last day for title change - 19 March 2021; submission - 30 April 2021.

All coursework was double-marked, and differences of >10% were resolved by discussion between the original assessors. All coursework was submitted to Turnitin and the output files reviewed by the senior examiner. No cases of plagiarism, or poor scholarship requiring action, were identified by (or independent of) Turnitin. One student required a literature review deadline extension and two students required project report deadline extensions for health reasons.

ii) Written papers

The written papers took place online with papers being downloaded from and scripts uploaded to Moodle within 24-hour assessment periods starting at noon on 27, 28 May and 1, 3, 7 June 2021. After polling the students to determine their time zones (one UTC+2, one UTC+3, one UTC+8, all others UTC+1) it was agreed that the internal examiners would be available via email for the first two hours of each exam to answer any queries. In practice continuous cover was provided (except between 1 - 6 am). Queries related to issues with submissions —delays due to internet problems or errors in the uploading procedure. In three instances, scripts were submitted via email to the Exam Administrator (Miss Amy

Glover) ensuring that anonymity was preserved, in parallel to email submission to the Exams Office. Otherwise, script submission and use of G Suite form for coversheets proceeded smoothly.

iii) Marking and return of scripts

At the end of each exam period, scripts were downloaded from Moodle and together with those received via email, sorted by the Senior Examiner. Yet, scripts delivered by email to the Exams Office only arrived after three days. Assessors generally felt that type-written scripts were much easier to mark. Almost all assessors were prompt, returning marked scripts within the prescribed 48-hour turnaround period. However, there were a few issues raised with marking taking place over weekends. The internal examiners undertook checks targeted to inexperienced Part II assessors. The External Examiner was fully informed about the overall marking process.

iv) Shared paper (Paper4/ZL5)

For the first time, scripts from candidates who sat Paper4/ZL5 as part of NST2ZO, NST2PL and NST2BBS with Zoology as major subject were processed by the Department of Zoology and assessed separately. Senior examiners from all three Departments interacted before and during the conduct of the exam. Support during the 24-hour period was offered uniformly to all candidates by Genetics in concert with Zoology. Genetics collected exam e-covers recording the questions answered by all candidates and submitted the relevant spreadsheet to the Department of Zoology.

v) External Examiner and *vivas*

The External Examiner carried out *vivas* for the 27 NST2GN candidates by Zoom teleconference on 14 and 15 June 2021. She was provided in advance with all scripts, coursework and assessment material of both the NST2GN and BBS candidates.

5. PERFORMANCE OF CANDIDATES

i) Written papers

As in previous years, the number of answers to individual questions was highly variable (Table 2). Answers were received for all questions, with four attempted by only two candidates. Seven questions were each attempted by ≥20 candidates.

Table 2: Question choice & average mark per question in written papers (all candidates)

Paper 1		**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Overall**
	N	5	8	28	21	12	2	20	96
	Mean	78.8	66.8	64.1	64.0	64.3	76.0	65.8	65.7
	SD	11.0	7.3	10.0	7.2	8.9	26.9	7.0	9.4
Paper 2		**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Overall**
	N	19	2	14	20	10	17	19	101
	Mean	64.3	71.5	66.3	62.6	65.4	63.2	65.4	64.5
	SD	8.9	4.9	4.0	6.7	8.7	17.5	5.7	9.5
Paper 3		**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Overall**
	N	6	11	18	12	10	14	24	95
	Mean	70.2	65.9	67.7	59.6	68.6	67.1	68.2	66.7
	SD	8.2	9.6	5.2	13.2	7.7	4.6	6.9	8.2
Paper 4		**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Overall**
	N	15	27	9	24	9	9	12	105
	Mean	61.9	66.1	59.2	68.6	62.7	63.8	68.9	65.3
	SD	3.8	7.5	9.8	6.6	5.3	5.4	5.2	7.1
Integrated Paper		**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Overall**
	N	7	2	2	3	3	6	4	27
	Mean	68.7	77.5	67.5	71.3	64.7	65.7	72.0	68.9
	SD	4.5	10.6	3.5	5.8	2.5	8.3	7.8	6.7

ii) Coursework
The average mark for the Literature Review 68.0% (SD 5.7) and for the Project Report was 68.9% (SD 5.6), and. For the BBS Dissertation, the average mark was 65.9% (SD 5.5).

6. THE FINAL EXAMINERS MEETING FOR PART II GENETICS was held on Wednesday 16 June 2021. The mark distribution across all papers was reviewed. All marks were awarded on a scale of 100 according to the marking scheme and class boundaries approved by the Committee of Management for the Natural Sciences Tripos. There was some discussion regarding the highest-ranked 2-i candidates: three candidates were moved into the first classification by unanimous agreement of the examiners. The 2-ii candidates were also thoroughly discussed. The class distribution was compared to the mean from the previous three years before the pandemic (2017, 2018 and 2019) and cohort equity was achieved by setting the Class I/Class II.i grade boundary to 69.6/69.5. The External Examiner accepted the results as robust. The Thoday Prize was awarded to the best overall performance. The grade roster was uploaded to CamSIS on 17 June 2020 and the mark book and class list also sent to the Student Registry.

7. OUTCOME OF THE PART II GENETICS EXAMINATION

i) NST2GN class distribution

First Class:	33% (n=9)
Upper Second Class:	56% (n=15)
Lower Second Class:	11% (n=3)
Third Class:	0%

ii) BBS (Genetics major subject)
The completed mark book was submitted to the Faculty Office on 21 June 2021. Of the 5 candidates sitting Part II BBS with Genetics as their major subject the outcome was as follows:

First Class:	0 % (0/5)
Upper Second Class:	80 % (4/5)
Lower Second Class:	20 % (1/5)
Third Class:	0%

8. STATISTICS FOR THE EXAM RESULTS

i) Class composition
27 NST2GN candidates: 17 NST (8 male, 9 female), 10 MVST (7 male, 3 female).

5 BBS Genetics major: 3 NST (0 male, 3 female), 2 MVST (1 male, 1 female); plus 5 BBS students taking a single Genetics Paper as their minor subject.

27 Zoology/Zoology BBS/Plants Sciences students taking shared Paper 4/ZL5

ii) Distribution of classes and movement between classes (NST2GN only)
No information available as the Part IB class was subject to formative assessment in 2019-20.

iii) Performance in written papers by Part II subject

Paper 1	**Gen**	**BBS**	**All**
N scripts	81	15	96
Mean	66.4	61.8	65.7
SD	9.4	8.2	9.4
Paper 2	**Gen**	**BBS**	**All**
N scripts	80	21	101
Mean	65.2	61.9	64.5
SD	9.9	7.3	9.5
Paper 3	**Gen**	**BBS**	**All**
N scripts	80	15	95
Mean	66.9	65.8	66.7
SD	8.6	6.1	8.2
Paper 4	**Gen**	**BBS**	**All**
N scripts	81	24	105
Mean	66.1	62.7	65.3
SD	7.0	6.7	7.1
Integrated Paper	**Gen**	**BBS**	**All**
N scripts	27		27
Mean	68.9		68.9
SD	6.7		6.7
Greatest difference between averages	2.8	4	4.4

iv) Mean marks in each exam component by gender (NST2GN only)

N=	12		15	
	Female		Male	
	Mean	SD	Mean	SD
Literature Review	67.4	4.8	68.5	6.5
Project	69.3	5.0	68.3	5.9
Examinations	67.7	6.0	65.3	8.7
Total	67.8	4.3	65.9	6.5

v) Mean marks in each exam component by background (NST2GN only)

N=	17		10	
	NST		MVST	
	Mean	SD	Mean	SD
Literature Review	69.1	4.9	66.2	6.8
Project	70.3	3.9	66.1	6.9
Examinations	68.0	7.7	63.6	6.9
Total	68.3	5.1	64.1	5.6

vi) Representation in classes by IB background (NST2GN only)

N=	17		10	
	NST		MVST	
	Count	%	Count	%
1	5	29.4	1	10
2i	11	64.7	7	70
2ii	1	5.9	2	20
3	0	0.0	0	0

vii) Representation in classes by gender (NST2GN only)

N=	12		15	
	Female		Male	
	Count	%	Count	%
1	2	16.7	4	26.7
2i	9	75.0	9	60.0
2ii	1	8.3	2	13.3
3	0	0.0	0	0.0

9. REPRESENTATIONS & APPEALS

No representations were received. The Senior examiner responded to four information-only requests for various details including marks for individual exam questions or verification of marks. One application for an allowance was received and processed.

10. ACKNOWLEDGEMENTS

I would like to thank my colleagues, Prof. Richard Durbin and Prof. Sara Goodacre, the Exam Administrator, Amy Glover, and the Departmental Administrator, Casey Mein, for their support and hard work in connection with the exams. Special thanks are due to Prof. Gos Micklem (Senior Examiner in 2019-20) for invaluable help and advice in establishing the conduct of online examinations, to Dr. David Summers for his continued presence and insight into best adapting and communicating the framework of assessment in this especially difficult year and to Dr. Christine Farr for her assistance in preparing the documentation to notify the new Part II Genetics examination structure and adaptations. Finally, I would also like to thank all participating assessors and project supervisors.

Dr. Marisa Segal (Senior Examiner)
August 2021

APPENDIX A: SENIOR EXAMINER'S REPORT 2021, Part II GENETICS AND BBS GENETICS, NATURAL SCIENCES TRIPOS

IP: Integrated Paper

Paper 1: Genomes, Chromosomes and the Cell Cycle

Paper2: Human Genetics, Genomics & Systems Biology

Paper 3: Developmental Genetics

Paper 4: Evolutionary Genetics & Adaptation

Marker	Paper	Q#	What are you looking for in the answer?	How well was the question answered?
David Summers	IP	1	A summary of the consequences of plasmids for the fate of humanity must necessarily begin a wide-ranging consideration of their effects and uses. The question required these to be sorted (with justifications) into good or bad for humanity. "Good" might include the uses of plasmids in research leading to medical advances or the development of bacterial cell factories leading to industrial benefits. Their application in the human genome project might be used as another good example. On the "bad" side of the argument we have the role of plasmids in the dissemination of antibiotic resistance and pathogenicity. A good critical assessment doesn't stop at listing the factors on the two sides of the equation but should try to ask more searching questions such as why we pin so much blame on plasmids for AbR when chromosomal resistance is the norm in some species (e.g. tuberculosis) and phages are well-known to transfer resistance (e.g. the origins of TB). Could we have developed and practiced gene cloning without plasmids? Are the categories of good and bad too simplistic here?	There were really no bad answers to the question but some were better than others. The weaker answers dwelled too much on plasmids and their properties, forgetting that the focus should be on their impact on human happiness. Weaker answers were sometimes limited in the range of plasmid effects on humans. Some answers made no mention of pathogenicity for example. At other times when discussing the ills of antibiotic resistance, some candidates seemed to forget that it is not invariably plasmid-associated. The best answers dug deeper into these aspects and provided a more genuinely critical appraisal of the question.
Marisa Segal	IP	2	Biological diversity and models in any area, (drawing from the course syllabus). A discussion illustrating the many sources of inspiration to innovative approaches and their significance in advancing research and knowledge.	Two excellent answers that differed in scope and approach. Both effectively integrated material across the course.
John Welch	IP	3	The question is broad, but some obvious topics include the estimation and interpretation of narrow-sense heritability, and connected GWAS methods. The historical association with eugenics would be an obvious thing to mention, and the widespread, but erroneous assumption that high heritability traits imply immutability. There is also the possibility of discussing more philosophical topics, concerning free will, just punishment etc.	There were only two answers. Both contained the key technical ideas, clearly explained, and discussions of eugenics.
Christine Farr	IP	4	Burden posed by monogenic disease versus common ills. Monogenic disorders – highly penetrant, no/few options to ameliorate symptoms and no cures. Western/ UK healthcare systems but also worldwide. Consider GT drug successes and failures – efficacy, safety. Costs of research & development and final high costs of drug. One-off treatments? Cost-benefit analyses. Economic burden. Financial constraints, limited resources. Right to healthcare? Explore possible alternative payment options. Overall conclusion.	Generally well argued answers, that tackled a wide range of complex issues.
Aylwyn Scally	IP	5	This is a challenging question and needs broad and deep understanding to answer well. Looking for a discussion of different gene concepts, both in a molecular context and in evolutionary theory. Not expecting a single consistent definition, but illustration that the term's meanings have changed/increased over time. Perhaps discussion of confusion between 'gene' and 'allele', and between abstract (evolutionary) & physical definitions. If some concepts no longer essential, which are they and why (with examples, ideally including theoretical aspects such as of polygenicity, not just molecular complexity in translation & regulation). If concept(s) still essential, which are they and what would we not be able to understand/explain without it/them?	Answers did a good job of showing why simplistic molecular definitions have broken down since the 60s, often with good examples. None really discussed the parallel concept of the gene in evolution and/or quantitative genetics, e.g. the concept of a replicator or what 'unit of inheritance' means, and to what extent this is still essential or valid. In places, answers focused too much on molecular details of examples. All answers concluded that despite confusion the _word_ gene is central to scientific & lay discourse - and that this meant the concept should evolve to remain relevant. But this implies that the previous concept has indeed become inessential. To argue that some core gene concept remains essential, need to elucidate what this is even if it is poorly understood.
Felix Day	IP	6	Candidates should have noted that there was benefit both for the individual and medical science. The individual might be a in a position to prevent any diseases that they might be at risk for (since they would have this information) early in life. And medical science would get access to a very large data set that they would be able to use to link more genetic variation to disease and help to understand more biology. There was also the possibility that sequencing everyone might reduce the lack of diversity in genetic data (though this would only be within a country). Candidates might also have noted that there were other, non-medical uses of the data, such a policing that might also be improved. Candidates might have contrasted the idea of sequence data with the existing tests for specific diseases. There were also some potential challenges to this sort of data collection - in particular the idea of informed consent is impossible with a new-born. Candidates might also have talked about the role of insurance companies (though they might have limited the economic discussion to national health services). There should also have been mention of data privacy - potentially of the risk of malicious or discrimination by those that hold the data. Finally, the candidate would, ideally, have come to a considered opinion of the best way to proceed rather than repeating the arguments in the concluding section.	All the candidates covered the medical uses of sequencing, but many could have been clearer about the possibility that this might focus attention on prevention rather than treatment. Some of the candidates made specific reference to the fact that the data was collected in new-borns, but some treated the question as though sequencing was simply being used for diagnosis. Non -medical uses of genetic data were sometimes ignored by the candidates. There was usually a question of the idea of consent, but again many candidates did not make clear the specific issues with informed consent if the data was taken at birth. Most candidates mentioned that there was a ethical question of the use of the data by insurance companies, but given that many essays did not cover non-medical uses of the data, the possibility of genetic data being used as part of discrimination was rarely discussed. Many candidates suggested that society or politicians would be the ones to make the decision, without always saying were they themselves saw the balance of evidence. Others did with good conclusions in favour or against the idea; some even cam up with a decision tree that could be applied to see if sequencing was appropriate.

Felipe Karam Teixeira	IP	7	Essays may discuss the role of technology in scientific discovery in general, a historical perspective of how abstraction and imagination has played a key role on the advent of major concepts in genetics, the impulse that technologies such as sequencing, imaging, cell culture, etc, have imparted to a variety of fields, and how the access to large amounts of data may critically shape the scientific questions being addressed	Despite the diversity of argumentation expected in integrative papers, most essays were well constructed, using a variety of examples and concepts learned in different Part II modules to build a solid argumentation on the balance between data generation, abstraction, and imagination on scientific discovery, and the current and future challenges brought by the ease in generating large amounts of data.
Marisa Segal	1	1	With due consideration to experimental support: - CDK(s) promoting order in the cell cycle - quali and quantitative models - Multiple Cdks and cyclins: redundancy versus specificity - Network motifs, stage-specific cyclins and the dynamics of cell cycle transitions - Temporal and spatial patterns linked to cyclins and their significance	Overall received excellent to outstanding answers. The range of marks depended on two key considerations: a) Show of command by narrating AND interpreting experimental evidence. b) A demonstration of the ability to integrate cell cycle concepts and paradigms, in particular, to address the second part of the question i.e. why multiple cyclins? The best answers drew from the course more broadly and achieved remarkable synthesis.
Marco Geymonat	1	2	1) Description of the spatial regulation of MEN (role of Kin4 and Lte1) 2) Importance of the Meta to Ana transition for the activation of the FEAR to de-P some MEN components 3) Feedback loop on Cdc14 4) Experiment set up to alter the localisation of the spindle in metaphase or anaphase 5) Experiment set up to alter CDK activity in metaphase or anaphase. 6) Experiment set up to monitor spindle position and MEN activation	I was pleased to see that the question was answered quite well by the 8 candidates that choose it. 2 answers were very good (one quite outstanding). Generally speaking the part better answered was the description of the MEN and its regulation in time and space. The part less well answered was the experimental set up. Also the role of the FEAR in the de-P of some MEN components was poorly described/understood (hence the feedback loop of Cdc14 was often under looked).
Rosana Collepardo	1	3	1. A definition of what the 30-nm fiber is. 2. Discussion of the experimental (and optionally computer modelling) evidence in favour of the 30-nm fiber in vitro, highlighting the reasons why in vitro experimental conditions bias the results towards ordered fibers. An excellent answer would discuss why were such experiments so important, and what can we learn from them. 3. A discussion of the key chemical and physical heterogeneity of chromatin in vivo versus reconstituted fibers (nucleosome breathing is a plus). 4. A discussion of the new experimental techniques that can interrogate the structure of chromatin at nucleosome resolution in cells. At least cryo-EM, STORM, and ChromEMT should be discussed. An excellent answer would describe they key features of the techniques, and potential drawbacks. 5. A discussion of how the consensus is shifting towards the idea of a disordered, or liquid-like organisation, and its biological implications	Q3 was very popular, with 28 students submitting an answer. 12 students produced excellent essays covering all the key points, and going beyond what was discussed in the lectures; clearly demonstrating that they had read relevant research articles critically, and had acquired a profound understanding of the ideas. 9 gave very good answers, describing the key ideas very well, but failing to demonstrate understanding going beyond what was covered in the lectures. 3 gave good answers but with some misunderstandings or errors, and 4 gave answers that were brief and missed the point of discussing the key experiments that can interrogate chromatin at the nucleosome level and provide evidence against the 30-nm fiber inside cells.
Christine Farr	1	4	Traditionally regarded as heterochromatic, but now known not to be transcriptionally silent, with eu-/ heter-/ and specialised chromatin at both structures. Consider both roles of the act of transcription itself and of the transcripts originating from these regions. Influence on localisation of centromere/ telomere proteins, establishment & maintenance of the various chromatin domains, any other aspects. Draw out any parallels/ differences between centromeres & telomeres in this regard	Generally well attempted. Better answers used a wide range of sources, differentiated between the act of transcription and the role of the RNA transcripts and explored parallels/ differences between the two functional elements.
Christine Farr	1	5	Introduce key features of nuclear organisation: membraneless nuclear bodies – concentrations of molecules associated with particular functions/ processes. Highly dynamic exchanges of molecules in these various organelles with those in the nucleoplasm. Often highly dynamic entities that can dissolve and reform along with change in activity status (e.g. nucleolus as cells enter/exit M phase). One such nuclear body is the nucleolus. Introduce the key features & function of this membraneless organelle and its relationship to the genome/ chromosomes. Then consider why the nucleolus has been described as a paradigm/ model of nuclear organisation. Nucleolus illustrates self-organisation & functional compartmentalisation. Dynamic, rapid mixing, yet stable organisation maintained. Role and mechanism of liquid-liquid phase separation. Informs about self-organisation of other nuclear subcompartments. Also consider functional partitioning and role in heterochromatin organisation.	Most candidates recognised the need to integrate material from more than one part of the module. Generally decent discussions of role of LLPS in the organisation of the nucleolus, with extrapolation of this to other nuclear bodies & nuclear organisation generally.
Marisa Segal	1	6	- Understanding of APC/C regulation along the cell cycle. Role of Cdh1. - Concept of biosensor. - Understanding of multisite phosphorylation driven by multiple Cdk-cyclin complexes - Concept of network motifs	One outstanding answer dealt with all points with very explicit and clear reasoning and justification. A second answer was on target but at points was conceptually confused.
David Summers	1	7	An appropriate way to approach this question would be via comparative genetic analysis of the two systems. Can counterparts of mutant phenotypes predicted by Pritchard's model be found in the extensive analysis that has been performed with P1? In short, the answer is yes: the replication-promoting factor in the model maps to P1 Rep while the replication-inhibiting factor maps to incA. More subtle considerations suggest that plasmid handcuffing is also consistent with the philosophy of the Pritchard model, even though the molecular components are very different. Not everything goes Pritchard's way, however, and some aspects of P1 control are more consistent with the ideas proposed in Sompayrac & Maaloe's alternative autorepressor model for the control of DNA replication.	Given the open-book nature of the exam some of the lower-scoring answers were surprising. Weaker candidates sometimes displayed poor understanding of the inhibitor dilution model or failed to investigate seriously how far the principles in the model could be found in the reality of P1 molecular biology. Mid-range answers recognised the repressor role of incA but the best went beyond this and explored how higher-level phenomena (handcuffing, Rep dimerization) might be interpreted in terms of the model. The best answers sometimes also recognised the alternative "autorepressor" model for plasmid replication control in some aspects of the P1 system. Weaker answers confused the autorepressor and inhibitor dilution models.

Gos Micklem	2	1	Non-exhaustive list of possible techniques: Genome annotation: - find genes and their exon/intron structure by - aligning known mRNA or mRNA from related organisms - aligning known protein sequences - aligning models to genome - align suitably distanced genomes to identify putative conserved regulatory motifs - multiple genome alignment can help with gene annotation and regulatory motif identification - Transcription factor ChIP-seq Genome architecture: - align Hi-C data to identify regions of the genome that are close together in space - ATAC-seq align sequences generated next to transposon insertions to assay chromatin accessibility Epigenome: - bisulphite sequencing: align modified sequence to determine methylation status - ChIP with histone modification antibodies: align pulled down DNA to genome to determine epigenetic state of bound chromatin Transcriptomics - alignment of sequenced cDNA to genome to determine gene expression levels - alignment of barcodes in e.g. split-seq as one approach to single-cell sequencing - use to annotate gene intron/exon structure Genome assembly can be viewed as sequence alignment.	There was a lot of variation in the way answers were approached, with some students not identifying distinct techniques or treating sequence alignment itself as a technique rather than the specified underpinning. The best answers wove complementary techniques, for instance transcriptomics, ChIP-seq and Hi-C, into a story about gene regulation. There was a tendency to skip the compare and contrast part of the question.
Gos Micklem	2	2	Many human genes are known to be involved in glycosylation. Examining gene expression data from mature B-cells could help select the ones most likely to be relevant. With cost being no option, resynthesis of genes will make sense. This will also allow codon optimisation of the incoming human genes and the addition of yeast promoters/translation signals. Human genes could be integrated into an existing chromosome, or a neo-chromosome constructed with its own centromere and telomeres. Either way, marker swapping as for Sc2.0 will likely be needed for the stepwise construction. Inclusion of loxPsym sites after each human gene will allow SCRaMbLE-style shuffling of the genes to create diversity of gene expression and the opportunity to screen for the best expression of the correctly glycosylated antibody while preserving strain fitness. Some work-around will be needed for the endogenous system: some genes will be non-essential and can be deleted. Others can be substituted for by the analogous human genes. However this is unlikely to be possible for all genes, so it may be necessary to make their expression repressible at both the transcription and protein (destabilisation) levels.	This was not a popular question with only two answers, both good. Given the open book nature of the exam one answer managed to dig out a great amount of molecular detail rather than approaching the problem from first principles.
Michael Imbeault	2	3	Some discussion about statistical power, where WES has a significant advantage given the higher number of samples the same money can afford. However, this should be counterbalanced by a discussion about how all the statistical power in the world won't help if what you are looking for (causal variant) isn't found where you are looking, which should lead a good answer to discuss the non-coding genome and the need to study better its biological functions to understand how mutations found within can influence biology / cause disease. Many possible examples, but transposable elements, gene promoters, distal enhancers, long non coding RNAs, etc are all good ones. No best answer as to which one is the best choice, however top answers would probably advocate for a balanced approach, using the cheaper, more statistically powered method to find variants in the exome, then fully sequence other patients.	Overall very well, only a few students wrote short texts without good detail / depth - I was pleased to see many diverse examples from the literature I didn't mention, although some of them might be from other lectures. The vast majority of the student appropriately tackled the question with varied opinions about the conclusion, which is nice to see.
Michael Imbeault	2	4	Mainly a description of the normal cycle of DNA methylation and its establishment and maintenance, with the actors involved. Then a discussion of the 'epigenetic clock' phenomenon, with a follow-up paragraph about mechanisms that can lead to this - essentially centered around entropy, but the best answers would make direct links, such as DNMT1 needing a template to maintain methylation, and if it was to make a mistake it could lead to the future propagation of it in daughter cells, which can lead to memory and aberrant depletion / enrichment at particular loci. The best answers would mention that this can slowly accumulate at specific loci in more and more cells of a particular tissue at a constant rate. Many other answers possible. As for the consequences of this on the biology of aging, a choice between general dysregulation of developmental genes, stem cell genes, oncogenes, etc leading to cell death, inflammation and so on. Also dysregulation of transposable elements leading to poor genomic stability. Ways to 'treat' it could include a discussion of targeting Cas9-DNMT3A to loci that need to be changed, but then quickly followed by a discussion of how impossible this task is - amount of cells to be treated, different patterns to be reestablished between different cell types, offtarget effects, etc.	Overall very well, although a few students (2) got very confused between DNA methylation and histone methylation mediated heterochromatin. Few people got exactly right what I had in mind as an easy answer with DNMT1-mediated maintenance, but many discussed more intricate mechanisms involving polycomb, or overall levels of DNMTs that change their expression over time. Many students also discussed other implications of the potential 'treatment' of aging, which was nice to see.
Richard Durbin	2	5	. discussion that the difference is likely due to choices made in handling repeats/duplicates, which is the major challenge of genome assembly . show how metrics suggest assembly 1 is inflated by including haplotypic duplication, while assembly 2 is overcollapsed by excluding paralogous copies . potential measures include trio sequencing, using coverage information to identify under/over-collapsed sequence, changing matching stringency, and using scaffolding methods (e.g. HiC or 10X) to phase contigs	Generally this was well answered, with only a couple of answers being really strong. Most people understood the measures in the table given in the question, and could explain the main reason for the differences, but there were differences of understanding and articulating the detail. The "what strategies could improve" section was in most cases less interestingly answered, relying on methods from the lectures. One person found the data set the question was based on, but it didn't obviously help them (not the highest score).

Richard Durbin	2	6	. four possible inheritance patterns out of X-linked, autosomal recessive, autosomal dominant (de novo in germline), mitochondrial, parent-of-origin effect . discussion of what these are and their recurrence risk across two brothers . discussion of relative likelihood of the potential cases and how that depends on other factors such as family history of the disease (if any), consanguinity, population history, parental age, parental exposure to mutagens	About half the answers were essentially complete and correct and made good use of material from lectures, with varying amounts of additional material. Others showed errors of understanding or application of logic. Most people could apply the logic required to prioritise standard modes of inheritance in terms of likelihood. Responses varied in how well they they addressed the second half of the question on factors that influence relative likelihood.
Steve Russell	2	7	A discussion of the contention that technologies have been the driver in our understanding of the regulatory of the genome. Arguments for and against accepted. illustration of how technological approaches have improved the understanding of gene regulation, these would include genomics (RNA-seq, ChIP-seq, ATAC-seq, HiC), imaging, computational. Material or insights not presented in the lectures expected for high scoring essays.	A range as expected. The best answers were comprehensive and wide ranging, illustrating a direct link between technology and gene regulation as well as an understanding of limitations. Weaker answers either ignored the link to the regulatory code, were incomplete or simply reproduced material presented in lectures.
Ben Steventon	3	1	A clear delineation of the concepts 'lineage and 'trajectory', evidence of wider reading and critical description of a range of research methodologies. Opportunity for exceptional essays that would demonstrate critical thinking beyond what it currently in the primary research literature.	It was answered very well, with some essays showing an exceptional level of understanding and critical thinking. It also revealed some cases where central concepts taught across the module were not fully grasped.
Ben Steventon	3	2	Clear description of the mechanism by which the actomyosin cytoskeleton can elicit a range of effects at both the single cell and tissue level. It provided an opportunity ti bring in examples from across the module, that some essays achieved. It also provided ample opportunity to bring in detailed molecular information from primary research material- also achieved.	It was well answered with opportunities for evidencing wider reader was taken up. There was also opportunity to demonstrate a deeper grasp of the complexities of interpreting gene function across multiple length scales (i.e. from intra-cellular function to whole organism phenotypes)- this was not achieved on this occasion.
Daniel St Johnston	3	3	I was expecting the students to start by defining a morphogen and the criteria need to prove that a molecule is a morphogen. A morphogen must be distributed in a gradient and the response to it must be graded. Most importantly, one needs to also prove that the morphogen directly induces at least two different responses at different threshold concentrations. This can be done by creating a uniform distribution of the putative morphogen and then showing that the response changes as the concentration increases. In the case of the Dorsal nuclear gradient on would compare the phenotypes produced by a dorsal null mutant, a Toll cactus double mutant and TollD. An alternative approach is to show that the morphogen regulates two target genes directly at different threshold concentrations. This requires showing direct binding to the regulatory regions of the target genes with different affinities (e.g Bicoid binding to otd and hb). The threshold can be set as above by the affinity of the morphogen for its DNA-binding sites in different target genes. It can also be set by cooperative binding to multiple sites (Bicoid) or cooperative binding to another ubiquitous transcription factor that binds at an adjacent site (eg Cut and Dri binding to AT rich sequence elements in zen. Students who read my references might also add that chromatin accessibility plays a key role. I was expecting most students to use Bicoid and Dorsal as examples as their roles as morphogens were extensively discussed in my lectures. However, other good examples would be Dpp in the fly wing, Activin in the Xenopus embryo (covered in 1B) and Shh in the vertebrate neural tube (covered in Ben Steventon's lectures).	This question was answered well by almost all students, perhaps because it was an open book exam and the lecture material directly dealt with this topic. Almost everyone understood the evidence required to prove that a substance acts as a morphogen and most provided to different ways to do this and presented the evidence provided in the lectures. The best answers included the latest work on how chromatin helps set the responsiveness to Bicoid and thereby determine the threshold. The worst answers contained irrelevant information and failed to address the question directly.
Julie Ahringer	3	4	. The E blastomere clonally gives rise to the entire intestine. The E fate is specified by maternal factors: 1. The SKN-1 transcription factor 2. A wnt signaling pathway Info from outside the lectures: END-1 and END-3 are functionally redundant txn factors important for intestine development and which have the expression pattern of TFX-1. END-1 and END-3 are are activated by transcription factor MED-1, which is both maternally and zygotically provided. Additional (novel) potential regulatory mechanisms could be proposed. . To investigate whether known pathways regulate TFX-1 expression, TFX-1 expression should be examined in different mutant backgrounds (skn-1 and wnt pathway mutants); anticipated results and their interpretation should be described and discussed. Experiments to explore other proposed mechanisms should be briefly described. . A genetic screen would be the best method to find novel genes that control TFX-1 expression. Since tfx-1 is a zygotically active gene, its expression must be transcriptionally regulated, though other mechanisms may also be involved. The regulators could be maternal factors or zygotic factors. An RNAi screen would be most efficient. The screen could be genome-wide, to be unbiased and try to find direct and indirect regulators, or could be directed (e.g, knockdown all transcription factors) to try to find a transcriptional regulator. Before the screen is conducted, it would be beneficial to analyse the expression of TFX-1; in particular, a reporter gene should be constructed to test whether the tfx-1 promoter is sufficient to drive the TFX-1 expression pattern. This would give evidence of a transcription control mechanism and identify the relevant sequence. The expression of the reporter gene could be used as an assay for the screen. The assay could alternatively involve staining using the antibody to TFX-1. This would be more work, but could uncover mechanisms that involve regulation of TFX-1 protein. Brief description of some downstream analysis of identified genes could be included.	Answers were spread across the range of marks. Some answered the first part well but struggled with the experimental section and vice versa. A few answered all three parts well.
Eric Miska	3	5	Evolution of RNAi, Argonaute proteins. TEs, viruses, genome structure. Fungi, killer, etc. Then miRNA and piRNA on their different paths. Plants, animals, specialisation of miRNAs.	I received 10 answers. It was a really interesting spread of answers. The somewhat integrative question really challenged the students to think through and then think beyond the material of the lecture and to bring in material form other lectures in the course. I was very happy over all.

Felipe Karam Teixeira	3	6	1) Answers are expected to provide a general description of the mechanisms of PGC specification in both specifies and the importance of the maternally inherited germ plasm. 2) Compare both species in terms of core components of germ plasm (Vasa, nanos, piwi, RNA, etc) and its assembly into granules (similarities and differences) and discuss the relationship between the establishment of embryo polarity and germ plasm assembly. 3) Based on the similarities in terms of maternal inheritance, granule formation, or PGC specification, provide a mechanistic rationale for the function of the uncovered gene.	All essays provided a description of germ plasm assembly and most of them also presented similarities and differences in terms of polarity establishment between Drosophila and C elegans. Many essays were also focused on the mechanisms of transcriptional quiescence, which would not be necessarily required for the answer answer. While some essays focused on the latter aspect (transcriptional regulation) to focus their experimental design regarding the gene of interest, others preferred to construct their essays focusing on more conserved features between the two systems (such as role of RNA binding proteins on PGC specification) to build their experimental design.
Marta Shahbazi Alonso	3	7	A complete description and a critical analysis of embryological and molecular evidence that supports the role of plasticity.	Excellent. I am extremely satisfied with the outcome. Most of the students had supervisions with me before the exam, and I think this really helped them. I think they were very well prepared, and most of them did an excellent job.
John Welch	4	1	The question was an attempt to allow candidates to explore the difficulties of testing theories in evolutionary biology. Some points to discuss would include (1) the difficulties in finding populations that differ solely in the variable of interest (i.e. level of sexual vs asexual reproduction); (2) the question of timescales (short-term vs. long-term effects); (3) the strengths and weaknesses of broad-scale comparative approaches (with limited control, and possible confounding factors) vs. detailed and controlled study of single populations, where results may not generalise; (4) the many difficulties in scoring Darwinian fitness in wild populations etc.	This was quite a difficult question, and it was not well answered in general. The majority of answers briefly summarised the two lectures on the evolution of sex, in the same order presented in the lectures. This meant that there were very few attempts to answer the question, and fewer examples of thought about the difficulties of testing theories. Some weak arguments recurred multiple times, e.g., (1) Muller's ratchet requires small populations, but this is difficult to test because most populations are large, and (2) Theory X is difficult to test because the mathematical model on which it was based contains simplifying assumptions.
Chris Jiggins	4	2	I was looking for an understanding of the complexity of the sources of similarity in genetic changes, as well as the major reasons for repeatability in particular - mutational hotspots, pathway position, low pleiotropy, single function genes	In general all answers addressed the main points and provided some relevant examples from the lectures. In some ways the question was too straightforward and proved very popular! Relatively few answers went beyond this and introduced a diversity of different sources and some novel ideas, but a few that did were really excellent
Nick Mundy	4	3	The principles and brief summary of dN/dS methodology, and its application to particular lineages and sites, with appropriate examples, including at genomic scale. The main issues with the methodology (only for coding, low power in some cases, difficulty linking to function). Other methods for non-coding sequences, such as rate variation, convergence. A brief comparison with the scope of these methods compared with population genomic methods.	Overall, answers were quite weak. Most included dN/dS methodology, but the description and inclusion of good examples was variable. A few answers included a good critique of these methods. Few answers considered methods for non-coding sequences, and a couple inappropriately spent a lot of time on population genomic methods, outside the scope of the question.
Emilia Santos	4	4	A description of the experimental design and the use of phylostratigraphy and synteny methods in a complementary fashion, which a critical discussion of the pros and cons of each of them. Mention how we can use syntenic alignments and species trees to find enabler mutations that transform a non-coding sequence into a protein coding gene. Include a comparison between the transcription versus ORF first models; calculating how often do these genes emerge, how complex is their gene structure and expression profiles Mention the identification of particular genomic features that are associated with and may facilitate emergence of de novo genes. Finally, the student should have discussed how to assess and validate the function of these genes. Perhaps, also discuss how these genes become functional (e.g. do they gain novel functions or do they get overlapping functions with pre-existing genes).	The answers were generally of high quality. The majority of the students included all the key topics mentioned above. About half did not reach the conclusion that both methods should be used in a complementary fashion. Sometimes they tended to list facts instead of critically presenting the dynamics and patterns of de novo gene formation (e.g. why should we expect that genes will emerge preferentially through an ORF first or transcription first mechanism). We are also supposed to mark the students' originality and extra reading, I found that most material was taken directly from the lecture and/or lecture notes. Finally very few discussed how such genes gain functions, and whether they play a major role in lineage specific adaptations or if they get integrated into conserved pre-existing functions.
Aylwyn Scally	4	5	a) Discussion of genealogical and genetic ancestry and the distinction between the two (i.e. pedigree and ancestry of inherited DNA); the exponential growth of genealogical ancestors and the fact that even one cross-population mating or migration is sufficient to merge genealogies. Some indication of where the figure of 5000 years for the genetic isopoint might come from (i.e. from modelling the amount of genetic contact globally). In the second part, explanation that DNA is not uniformly inherited because the number of paths through the genealogy to a given ancestor is not uniform. Thus e.g. an individual in Africa will thus derive more genomic segments from an African ancestor 20kya than an Asian (indeed may derive none from the latter). A strong answer could also mention differences due to selection, inbreeding an other genealogical parameters, and archaic admixture. b) Correct calculation and clear derivation. Then in final part, discussion could include: goodness-of-fit or hypothesis testing of the HW model e.g. using chi-squared test (given a number of samples or other details of the data); model comparison, e.g. nested likelihood or BIC/AIC approaches. Also perhaps further experimental investigation, eg. looking for other evidence of inbreeding such as homozygosity & IBD.	On the whole the question was not terribly well answered. Discussing genealogical ancestry, most answers were largely correct but none really explained the 5 kyr figure. In discussing genetic ancestry all but one answer missed the key point that such ancestry is indeed equally likely in the absence of population structure, so the reason genetic ancestry differs is that there is an unequal distribution of genealogical ancestors. The calculation in b was straightforward and most answers got it all correct. However the final part was also badly answered in general. Only a few answers really discussed it as a model comparison or goodness of fit issue, or suggested any sort of statistical test.

Bill Amos	4	6	Good well-argued answers that address the question set	Generally disappointing. This was quite a tough question and contained more than one elephant trap (e.g. is a phenomenon rare just because there are few examples? or because it is evolutionarily transient?). Sadly, most students fell straight in. Overall, too much listing of examples without proper connection to the question but on the plus side, evidence of plenty of outside reading.
Frank Jiggins	4	7	This is a broad question that could be answered in multiple ways. Core material os the role of selection vs drift, and transposable elements/levels of selection. It would also have been possible to discuss material from beyond the lectures about many other features of genomes that are shaped by natural selection or other materials.	Many answers drew heavily on the lectures with rather limited material from outside them or wider reading. In the lectures there is discussion of both selection versus drift, and selfish genetic elements, most motably transposable elements. Many answers tended to pick one of these topics. Some answers did not pay much attention to the word 'genomes' writing about single genes or phenotypes.