New genomic resources in watercress (Nasturtium officinale, R.) enabling molecular breeding for improved anti-cancer properties, flavour and nutrient density

Gail Taylor, Annabelle Damerum, Nikol Voutina, Daniel Kliebenstein, Mark Coldwell, Graham Packham, Mark Chapman, Tingting Zhu, Mingcheng Lou

Plant and Animal Genome Meeting XXVIII
January 11 2020
Aims of this research

To develop new genomic resources in watercress to underpin a molecular breeding program

- germplasm collection
- mapping population
- genome sequence
- Phenotyping and genotyping
### 17 Nutrients of Public Health Importance

17 nutrients of public health importance per the Food and Agriculture Organization of the United Nations and Institute of Medicine (i.e., potassium, fiber, protein, calcium, iron, thiamin, riboflavin, niacin, folate, zinc, and vitamins A, B6, B12, C, D, E, and K)

<table>
<thead>
<tr>
<th>Vegetable</th>
<th>Nutrient Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watercress</td>
<td>100%</td>
</tr>
<tr>
<td>Chinese cabbage</td>
<td>91.99</td>
</tr>
<tr>
<td>Chard</td>
<td>89.27</td>
</tr>
<tr>
<td>Beet green</td>
<td>87.08</td>
</tr>
<tr>
<td>Spinach</td>
<td>86.43</td>
</tr>
<tr>
<td>Chicory</td>
<td>73.36</td>
</tr>
<tr>
<td>Leaf lettuce</td>
<td>70.74</td>
</tr>
<tr>
<td>Parsley</td>
<td>65.59</td>
</tr>
<tr>
<td>Romaine lettuce</td>
<td>64.48</td>
</tr>
<tr>
<td>Collard green</td>
<td>62.49</td>
</tr>
<tr>
<td>Turnip green</td>
<td>62.12</td>
</tr>
<tr>
<td>Mustard green</td>
<td>61.39</td>
</tr>
<tr>
<td>Endive</td>
<td>60.44</td>
</tr>
<tr>
<td>Chive</td>
<td>54.80</td>
</tr>
<tr>
<td>Kale</td>
<td>49.07</td>
</tr>
<tr>
<td>Dandelion leaves</td>
<td>46.34</td>
</tr>
<tr>
<td>Red pepper</td>
<td>41.26</td>
</tr>
</tbody>
</table>

Watercress was the *first* fast food

A long cultural heritage
- no breeding and selection
- no active crossing
- no molecular biology
- no genome sequence
Watercress has been the past but does it have a new future?
Brassicaceae crops and health benefits

• Research suggests that Brassicaceae consumption is linked to a reduced risk of: cancer, cardiovascular disease, diabetes, asthma, Alzheimer’s disease, and have antimicrobial activity

(For review: Manchali et al. 2012)
Introduction to Watercress

- *Nasturtium officinale* (Rorippa nasturtium-aquaticum)
- Perennial dicot herb
- *Brassicaceae* family
- Aquatic or semi-aquatic habitat
- Traditional crop of southern England
• Watercress is the new kale

• Dioscorides in 70 AD

• Detoxified tobacco carcinogens
  (Hecht et al. 1995)

• Suppressed cancer *in vitro*
  (Zhang et al. 2003, Rose et al. 2005)

• Reduced DNA damage & increase in blood antioxidant
  (Gill et al. 2007, Fogatry et al 2013)

• Decreased of inflammation
  (Sadeghi et al. 2013)

• No.1 ‘powerhouse’ fruit or vegetable
  (Di Noia 2014)
Anti-cancer properties of watercress: Glucosinolates and Isothiocyanates

- Most studied compound of Brassicaceae
- Glucosinolates are hydrolyzed by myrosinase
- Myrosinase is released by cell damage
- Producing isothiocyanate
- Secondary metabolite & plant defense
- Chemopreventative properties

Gluconasturtiin

![Reaction](image)

phenethyl isothiocyanate (PEITC)

Peppery taste!
Chemopreventive Pathways

Inducing apoptosis, limiting cell growth, and inhibiting angiogenesis

Graham Packham, Cancer Genetics, University of Southampton, UK
Watercress: the most nutrient intense leafy crop

**Daily Express**
Latest weapon in the fight against cancer: WATERCRESS

**Spicy leaves are a superfood, say boffins**

**Watercress fights cancer**

By Andrea Hamnett

Move over blueberries – WATERCRESS is the new superfood of 2007. It’s so rich in antioxidants that it could help to reduce the risk of cancer.

**Daily Mail**

**Good news: Watercress can cut risk of cancer.**
Bad news: You need to eat a bagful every day

**Telegraph**

Watercress is the latest super food to combat cancer

**Times**

Eating raw watercress every day may reduce risk of cancer

**The Guardian**

A bagful of cress a day may keep cancer cells at bay, study suggests

**The Sunday Telegraph**

Watercress ‘fights cancer’ ... if you eat a bunch every day

**Evening Standard**

Watercress may stop growth of cancer cells

**Superfood prevents disease say doctors**

Eating a portion of watercress every day could help protect against cancer, according to new research. Scientists at Southampton University found that volunteers who ate 80 grams of watercress a day were less likely to develop cancer.
Identifying phenotypic variation

- Germplasm collection established at the UoS
- Assessed phenotypic variation within the collection across environments
- Evaluated gene expression in three accessions of interest
Capturing natural genetic variation

- Plants grown on watercress farm
- Phenotyped at 7 weeks (Payne et al 2015)
- Five high/low: RNASeq and GLS quantification
First watercress transcriptome study

We sequenced the expressed part of the watercress genome (RNASeq) and gene expression between samples with high and low health benefits.

**Outputs:**
- 323,827,923 reads produced
- Transcriptome assembled
- Catalogue of 50,000 genes created
- Differential Expression (DE) analysis for AO and GLS
- Polymorphic markers developed
- Candidate gene mining
- Publication

**Statistics of watercress assembly**

<table>
<thead>
<tr>
<th>Total transcripts</th>
<th>87,844</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total ‘genes’</td>
<td>48,732</td>
</tr>
</tbody>
</table>

We sequenced the expressed part of the watercress genome (RNASeq) and gene expression between samples with high and low health benefits.
Breeding targets in watercress?

Secondary plant metabolites:
- Antioxidants (AO)
- Glucosinolates (GLS)

Improved forkability, different flavour
- Consumer preferences?

Resource-use efficiency:
- Less fertilizer

Peppery taste!

Myrosinase
A watercress mapping population

We created the first watercress mapping population and RILs
First mapping population

- 260 F2 offspring grown in **ECR & on the farm**
- DNA extracted & sequenced (DDRAD Ed Buckler)
- Phenotyped
- Seed collected for future work

877 billion bases of watercress sequenced!!
Watercress genetic map development

• GBS data for 260 lines - DDRAD
• SNPs identified de novo using Stacks software
• Genetic map construction
  • **Map V1**: rqtl: 321 markers into **21 linkage groups** with an overall length of 1733.2 cM, 82.5 average length/LG, average 15.2 markers/LG, average marker spacing of 5.8 cM
  • **Map V2**: JoinMap (Max. Likelihood): 259* markers into **16 linkage groups**, overall length 1764.7, 110 average length/LG, average 16.1 markers/LG, marker spacing = 6.5 cM
  • **Map V3**: in progress, using Wx genome assembly to call SNPs in Stacks
The first linkage map and QTL for watercress

The linkage map contains 320 markers and found 17 significant QTL explaining health benefits and morphology in watercress.

- Dr Joost Keurentjes and Dr Erik Winjker
First watercress genome assembly

• 10X linked read sequencing (Supernova v2.1.1) + PacBio long reads (Falcon, PBBioconda), merged using Quickmerge

• Assembly stats:

<table>
<thead>
<tr>
<th></th>
<th>10X</th>
<th>PB</th>
<th>Merged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number contigs</td>
<td>1.48E+04</td>
<td>664</td>
<td>626</td>
</tr>
<tr>
<td>Total assembled bases</td>
<td>2.66E+08</td>
<td>3.65E+08</td>
<td>3.595E+08</td>
</tr>
<tr>
<td>Percent.N</td>
<td>4.541</td>
<td>0</td>
<td>0.692</td>
</tr>
<tr>
<td>GC content</td>
<td>0.3611</td>
<td>0.3743</td>
<td>0.3745</td>
</tr>
<tr>
<td>N50</td>
<td>3,436,423</td>
<td>2,381,132</td>
<td>3,685,503</td>
</tr>
</tbody>
</table>
Bionano – improved assembly

• 1 flow cell = 122 Gb molecules (>150 kb) = ~300X coverage

• Assembly:
  • N50 = 10.35 Mb
  • longest scaffold = 23.2 Mb – likely whole chromosome
  • 80% genome in 28 scaffolds

<table>
<thead>
<tr>
<th></th>
<th>PacBio + Bionano alignment</th>
<th>PacBio + Bionano resolved alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td># Sequences</td>
<td>664</td>
<td>738</td>
</tr>
<tr>
<td>Max length</td>
<td>9,884,970</td>
<td>9,881,647</td>
</tr>
<tr>
<td>Total size</td>
<td>364,712,898</td>
<td>360,521,812</td>
</tr>
<tr>
<td>Sequence N50 N%</td>
<td>2,381,132</td>
<td>2,329,984</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Scaffolds

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

536

23,202,078

361,128,634

10,350,430

1.45

Ming Cheng Lou, Tingting Zhu
Resolving mis-assemblies in merged 10X/PacBio assembly

Tingting Zhu
Chromosome-scale assembly

- Anchored 59 scaffolds, totaling 312 Mb onto 16 linkage groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Length (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.9</td>
</tr>
<tr>
<td>2</td>
<td>16.8</td>
</tr>
<tr>
<td>3</td>
<td>21.9</td>
</tr>
<tr>
<td>4</td>
<td>14.3</td>
</tr>
<tr>
<td>5</td>
<td>21.4</td>
</tr>
<tr>
<td>6</td>
<td>16.5</td>
</tr>
<tr>
<td>7</td>
<td>22.2</td>
</tr>
<tr>
<td>8</td>
<td>20.9</td>
</tr>
<tr>
<td>9</td>
<td>13.0</td>
</tr>
<tr>
<td>10</td>
<td>22.0</td>
</tr>
<tr>
<td>11</td>
<td>18.4</td>
</tr>
<tr>
<td>12</td>
<td>18.5</td>
</tr>
<tr>
<td>13</td>
<td>18.6</td>
</tr>
<tr>
<td>14</td>
<td>35.7</td>
</tr>
<tr>
<td>15</td>
<td>15.7</td>
</tr>
<tr>
<td>16</td>
<td>15.2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>312.0</strong></td>
</tr>
</tbody>
</table>
Mapping population on the farm – and in ECR
Population phenotyping

- Stem length (cm)
- Stem diameter (mm)
- Number of nodes
- Mean internode distance (cm)
- Specific leaf area (cm²/g)

[Graphs showing frequency distributions for different plant traits with red and blue lines indicating Parent A and Parent B, respectively.]

[Images of plants showing variability in leaf and branch patterns.]
lines – a first

Decreasing watercress sap ul/ml

Metabolic activity (% of untreated)

log(Concentration)
Cell cytotoxicity assays

Phenotyped 1,000 watercress samples for cancer-fighting ability!
Transgressive segregation in this trait as well.
QTL for anti-cancer traits

- QTL for IC50 mapped alongside those for long chain glucosinolates and Phenethyl isothiocyanate (PEITC) – good target for future breeding efforts
Summary

• Watercress is a powerhouse leaf of high value – plants can deliver enhanced phytonutrients for health and may have significant ‘value added’; suitable for indoor ag.

• We have developed the first mapping population, molecular genetic map and transcriptome of watercress, to underpin future breeding.

• We have provided the first genome assembly, significantly improved through optical mapping.

• We have identified a candidate QTL for anti-cancer properties for future breeding.